

# Why So Many Errors ?: Use of Articles by Japanese Learners of English

Tomoko Kaneko

## Abstract

English articles are not easy for Japanese learners to master. Although they are categorized as local errors, it is often said that article errors are fossilized for most Japanese. Nevertheless, understanding the system of article use is a key to learn the way native speakers of English realize the world in language. In this paper, the actual use of English articles by Japanese university students will be analyzed using ICLE Error-Tagged Japanese Sub-corpus and reasons for the erroneous use will be explored. In addition, some ideas on more effective ways for teaching Japanese students how to use articles appropriately will be proposed.

## 1. The ICLE Error-Tagged Japanese Sub-Corpus

The International Corpus of Learner English (ICLE) contains over 2 million words of argumentative essays from 19 different mother tongue backgrounds. Japanese sub-corpus has been collected by a team at Showa Women's University. The writing in the corpus has been contributed by advanced learners of English as a foreign language rather than as a second language. For the purpose of the ICLE project, "the advanced learners of English" is broadly defined as university students of English in their 3rd or 4th year of study. The ICLE error-tagged corpus has also been compiled. The original purpose of the error-tagged corpus is to "draw up a list of the most typical grammatical, lexical and lexico-grammatical errors in each of the L1 varieties, which could serve as a basis for an L1-differentiated or common core set of pedagogical materials" (Granger & Meunier, 2003). More than 10 countries in the world, including Japan, have completed error-tagged sub-corpus already. As for the ICLE Error-Tagged Japanese Sub-Corpus, one hundred and three files had been randomly selected from the ICLE Japanese Sub-Corpus and error-tagged. The error-tagged corpus contains 70,507 tokens and 5,096 types. The standardized type/token ratio is 34.79, which means that the students used on average about 35 different words in the text of each 100 words.

Chart 1 is a list of frequency of errors in the corpus. Grammar is the cause of a large proportion of errors. Among the grammatical errors, articles were the most erroneous items, even more erroneous than verbs.

Chart 1. Frequency of Errors in ICLE Error-Tagged Japanese Sub-Corpus

Tag		Error Rate	Tag		Error Rate
Form		6.25%	Lexis		16.25%
Grammar		42.53%		Single Words	9.32%
Articles		15.48%		Phrases	3.54%
	Nouns	8.39%		Connecting Words	3.40%
Pronouns		5.10%	Word		8.29%
Adjectives		0.40%		Redundant	4.28%
Adverbs		0.73%		Missing	3.38%
Verbs		10.50%		Word Order	0.64%
Parts of Speech		1.93%	Punctuation		7.32%
Lexicogrammar		8.17%	Register		0.78%
Conjunctions		2.13%	Style		10.39%
	Prepositions	4.23%	Total Errors		100.00%
	Countable/Uncountable	1.82%	Total Errors/Total Words		14.04%

## 2. Studies on Article Errors by Japanese Learners of English

Izumi et al. (2004) studied Japanese learners' use of English articles based on the NICT JLE Corpus (The National Institute of Information and Communications Technology Japanese Learner English Corpus). The data for the NICT JLE Corpus was compiled from the Standard Speaking Test (SST) interview. Learner errors in articles are divided into three groups; substitution, redundancy, and missing errors. Their finding is that missing errors are the most frequent in the erroneous use of English articles among the Japanese learners and that they have a tendency to overuse definite articles. They also state that the context of the article errors made by the learners will tell us more on how they should learn articles than a mere error frequency list. The present study has been originally motivated by their study.

It is not plausible that the learners use articles randomly. If there is a rule in the use of articles at each level of their interlanguage, finding the rule will shed light on a better way of teaching and learning the article system to Japanese learners. One of the most widely used models for classifying noun phrase environment in English article acquisition studies is the one presented by Huebner (1983, 1985). In his model, English noun phrases are grouped into four types as follows using two semantic functions HK (hearer known) and SR (specific referent).

Type 1	-SR/+HK	<i>the, a, or zero</i>	<i>Ø Lions are beautiful.</i>
Type 2	+SR/+HK	<i>the</i>	<i>Ask <u>the</u> man over there.</i>

Type 3	+SR/−HK	<i>a</i> or zero	<i>She gave me <u>a</u> present.</i>
Type 4	−SR/−HK	<i>a</i> or zero	<i>He's <u>a</u> nice man.</i>

The above categorization seems useful in finding out the interlanguage system in article use and course of article acquisition by learners.

A thorough study on the causes of learners' difficulty in using articles properly was conducted by Butler (2002). She administered a structured interview to 80 Japanese college students on why the learners chose the target article immediately after a fill-in-the-article test. The study tried to examine the difference of the metalinguistic knowledge on the English article system of the learners with different English proficiency. She found that the higher the learners' proficiency levels were, the more target like usage they could achieve, while that there remained a large gap in the use of articles between the native English speakers and even the most advanced Japanese learners. She states that Japanese learners "found it hard to determine which circumstances or conditions would make a reference identifiable to the hearer", which leads to the misuse of articles and concludes that "considering the number of the reference is a prerequisite for understanding" the English article system.

T. Ionin (2003) examines the acquisition of English articles by Russian and Korean speakers in her dissertation. The learners with these language backgrounds were selected because no articles are used in their native languages. She collected data from a series of elicitation tasks, as well as from a collection of written L2 production data. It is shown that L2-English learners' errors in article use are systematic and that they reflect the degree of learners' access to the universal semantic distinction of definiteness and specificity. The results show that the learners overuse *the* in [+specific] indefinite contexts, and overuse *a* in [−specific] definite contexts. In contrast, the learners are highly accurate in using *the* with [+specific] definite contexts, and using *a* with [−specific] indefinites. It is concluded that L2-learners fluctuate between two linguistic possibilities: specifying *the* as [+definite] and specifying it as [+specific].

Although article errors are categorized as local errors which do not disturb the communication in general, it is true that even advanced level learners, especially in the case of Japanese learners, have difficulty in mastering the system. Part of the complexity, Andersen (1984) points out, lies in the fact that the English article system does not consist of one-to-one form and meaning relationships. In addition, articles are especially problematic for Japanese learners of English because Japanese does not have an article system, where definiteness/indefiniteness and countable/uncountable are indicated by different linguistic means. Asano (1996) developed a format by which his students would be able to find the correct article to be used for Japanese junior college students. The trial was not a great success. However, his realization that correct article use is crucial for

Japanese learners of English because the use of articles reflect a part of the native speakers' concept is shared by many Japanese teachers of English.

Thus, the questions addressed in the present paper are the following:

1. Which article error type (substitution, redundancy, missing) is the most frequent among the Japanese learners ?
2. Are there any learner differences in the way the learners make article errors ?, and
3. In what language context are the article errors most frequent ?

### 3. Analysis

Frequency and concordance analyses were undertaken to ascertain the extent to which learners used the articles correctly and to determine the learner difference in the correctness of the article use. In addition, in order to check the context where the article errors are most frequent, the language context of article use were manually checked.

#### 3-1. Frequency of article errors

Chart 2 shows the frequency of article errors in the sub-corpus. The first three rows in the article column show the frequency of indefinite articles. The 4th row shows the use of the definite article. The "obligatory context" column shows the total frequency calculated by adding up the frequencies of correct article use, tagged article errors, and tagged other errors with articles as corrections. "Incorrect use" consists of three types; substitution, redundancy, and missing. Substitution errors are the ones substituted by erroneous articles or other words/phrases. For example, the number 68 in the indefinite article row, *a*, under "substitution", *the*, in the "incorrect use" column, shows the frequency of erroneous use of *the* in the place of *a*. Redundancy errors are the ones where the students put unnecessary articles. Missing errors are the ones where the students missed a certain article.

To make the comparison easier, Chart 3 sets the obligatory context frequency as 10,000. The average frequency of the correct use of indefinite articles is 7,435, while that of the

Chart 2. Frequency of Article Errors in ICLE Error-Tagged JSC

Articles	Obligatory Context	Correct Use	Incorrect Use								
			Substitution					Redundancy	Missing	Total	
			<i>a</i>	<i>an</i>	<i>the</i>	others	Sub-Total				
I.	<i>a</i>	1341	1014	/	0	68	5	73	23	231	327
	<i>an</i>	199	131	9	/	9	1	19	4	45	68
	Total	1540	1145	9	0	77	6	92	27	276	395
D.	<i>the</i>	3100	2444	30	8	/	36	74	194	388	656
Total		4640	3589	39	8	77	42	166	221	664	1051

Chart 3. Frequency of Article Errors per 10000 Words

Articles	Obligatory Context	Correct Use	Incorrect Use								
			Substitution					Redundancy	Missing	Total	
			<i>a</i>	<i>an</i>	<i>the</i>	others	Sub-Total				
I.	<i>a</i>	10000	7562	/	0	507	37	544	172	1723	2438
	<i>an</i>	10000	6583	452	/	452	50	955	201	2261	3417
	Ave.	10000	7435	58	0	500	39	597	175	1792	2565
D.	<i>the</i>	10000	7884	97	26	/	116	239	626	1252	2116
Average		10000	7735	147	39	166	91	358	476	1431	2265

definite article is 7,884, which shows that the definite article is used more accurately than the indefinite articles. Missing errors are more frequent in both indefinite articles and the definite article. Finally, the least frequent error type in indefinite articles is redundancy errors, while that in the definite article is substitution errors.

### 3-2. Learner differences

Although the frequency of the obligatory context for indefinite and definite article use depends on the content of the essays, there seemed to exist great learner differences in the profiles of the learners' article errors. Thus, after all the frequency of correct and erroneous article use of all the files being checked, the worst article users and the best article users were picked up as shown in the following two charts. Chart 4 shows the profile of the worst 12 article users. The files listed in Chart 4 are selected based on the ratio of incorrect use. The highest incorrect ratio was 60.5% (file No. 80) and the 11th highest was 27.3% (file No. 13). Although the ratio of the incorrect article use is not high, the file No. 96 is included in this chart. It is because the file has all of the three types of indefinite article errors in one file. Compared to the total ratio of the erroneous use of indefinite articles (47.6%), that of the definite article (35.5%) is much lower. Both in definite and indefinite articles, there are more missing errors than the other two types of errors among the worst article users. Nevertheless, some students made more errors with the definite article, and others made more with the indefinite articles.

Chart 5 is a list of the best 12 article users. There were three students who made no errors in articles in the 103 files. Among the three types of errors, missing errors were still the most frequent, although the difference from the other two types of errors is very small. The difference of the ratio of erroneous use between the definite article and indefinite articles is narrower than that of the worst users' list. This means that the students lessen the frequency of substitution and redundancy errors more than that of missing errors, which infers that they have progressed in the use of indefinite articles more quickly than

Chart 4. Profiles of Frequent Error Makers

File No.	<i>a, an</i>						<i>the</i>						Total			
	Oblig. Cxt.	Correct Use	Incorrect Use				Oblig. Cxt.	Correct Use	Incorrect Use				Oblig. Cxt.	Correct Use	Incorrect Use %	
			Subs.	Red.	Miss.	Total			Subs.	Red.	Miss.	Total				
4	7	5	0	0	2	2	40	22	2	0	16	18	47	27	20	42.6
13	9	8	0	0	1	1	79	56	1	3	19	23	88	64	24	27.3
28	13	7	0	0	6	6	40	19	0	3	18	21	53	26	27	50.9
59	17	7	0	0	10	10	20	12	1	4	3	8	37	19	18	48.6
61	9	3	1	0	5	6	50	34	0	10	6	16	59	37	22	37.3
80	21	7	1	0	13	14	17	8	0	0	9	9	38	15	23	60.5
81	7	6	0	0	1	1	72	51	2	9	10	21	79	57	22	27.8
91	19	13	3	0	3	6	47	24	5	12	6	23	66	37	29	43.9
96	29	21	1	1	6	8	13	13	0	0	0	0	42	34	8	19.0
99	8	4	1	0	3	4	60	33	8	15	4	27	68	37	31	45.6
100	15	4	0	0	11	11	37	33	0	0	4	4	52	37	15	28.8
102	12	2	0	0	10	10	9	7	0	0	2	2	21	9	12	57.1
Total	166	87	7	1	71	79	484	312	19	56	97	172	650	399	251	
%	100.0	52.4	4.2	0.6	42.8	47.6	100.0	64.5	3.9	11.6	20.0	35.5	100.0	61.4	38.6	

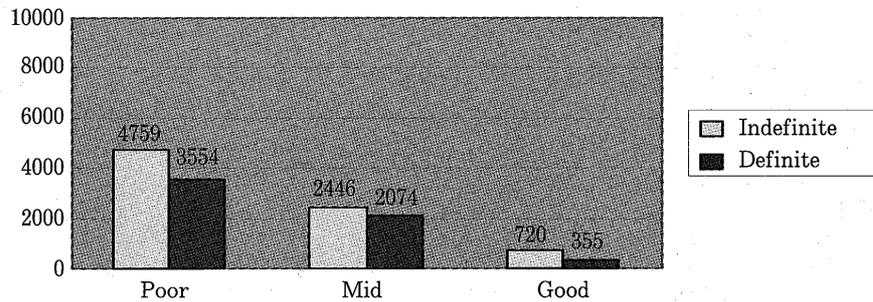
Chart 5. Profiles of Infrequent Error Makers

File No.	<i>a, an</i>						<i>the</i>						Total			
	Oblig. Cxt.	Correct Use	Incorrect Use				Oblig. Cxt.	Correct Use	Incorrect Use				Oblig. Cxt.	Correct Use	Incorrect Use %	
			Subs.	Red.	Miss.	Total			Subs.	Red.	Miss.	Total				
3	9	9	0	0	0	0	38	35	0	3	0	3	47	44	3	6.4
12	4	4	0	0	0	0	68	68	0	0	0	0	72	72	0	0.0
19	9	9	0	0	0	0	31	31	0	0	0	0	40	40	0	0.0
20	23	23	0	0	0	0	28	27	1	0	0	1	51	50	1	2.0
34	17	16	1	0	0	1	19	18	0	0	1	1	36	34	2	5.6
48	8	7	0	0	1	1	4	3	0	0	1	1	12	10	2	16.7
55	7	5	1	0	1	2	14	13	1	0	0	1	21	18	3	14.3
69	11	11	0	0	0	0	20	18	0	0	2	2	31	29	2	6.5
71	12	9	2	0	1	3	48	48	0	0	0	0	60	57	3	5.0
83	8	7	0	0	1	1	12	11	0	0	1	1	20	18	2	10.0
85	5	5	0	0	0	0	29	29	0	0	0	0	34	34	0	0.0
89	6	5	0	0	1	1	27	25	0	0	2	2	33	30	3	9.1
Total	125	110	4	0	5	9	338	326	2	3	7	12	457	436	21	
%	100.0	88.0	3.2	0.0	4.0	7.2	100.0	96.4	0.6	0.9	2.1	3.6	100.0	95.4	4.6	

Chart 6. Comparison of Error Frequency by the 3 Groups

Levels	<i>a, an</i>					<i>the</i>					Average	
	Correct Use/10000	Incorrect Use/10000				Correct Use/10000	Incorrect Use/10000				Correct Use/10000	Incorrect Use/10000
		Subs.	Red.	Miss.	Total		Subs.	Red.	Miss.	Total		
Poor Users	5241	422	60	4278	4759	6446	393	1157	2004	3554	6138	3862
Mid Users	7554	1591	208	647	2446	7926	1244	752	234	2074	7794	2206
Good Users	8800	320	0	400	720	9644	57	89	207	355	9540	460

Graph 1. Comparison of Erroneous Article Use by the 3 Groups



the definite article from the “poor” to the “good” article users’ levels.

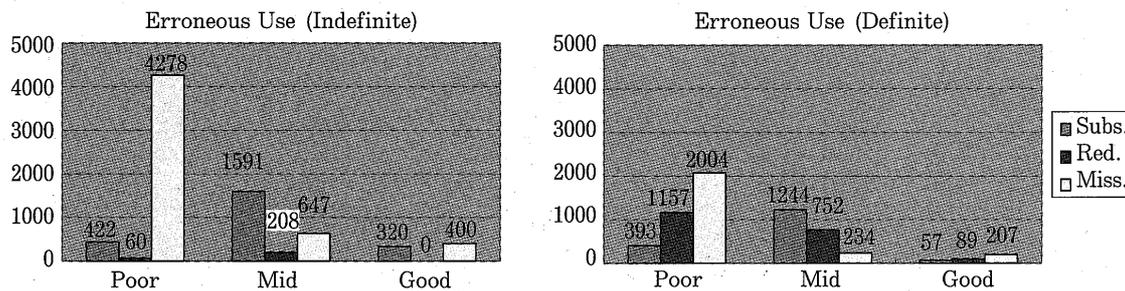
After the worst 12 and the best 12 articles users were picked out from the data, the rest of the files were grouped as “mid” level users. Chart 6 shows a comparison of error frequency of the three groups by setting the obligatory context frequency as 10,000. The frequency of correct use smoothly increases from “poor” to “mid” to “good” groups. However, although the frequency of total erroneous use of all the types does follow the reversed pattern, there are a few error types which do not follow the same pattern.

Similarities and differences of the profiles of the “poor”, “mid”, and “good” article users give some interesting insight. Graph 1 shows how many differences and similarities there are among the three groups in the frequency of indefinite and definite article errors. Indefinite articles were more erroneous than definite articles in all the groups.

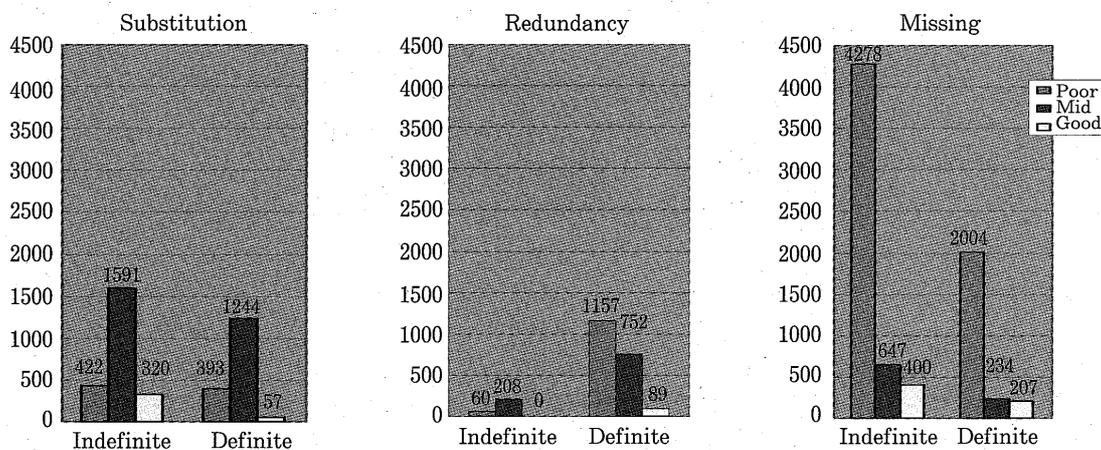
Graph 2 compares the use of indefinite and definite articles by the three groups. The typical profile of the “poor” group is that it shows a lot of missing errors in both indefinite and definite articles and redundancy errors in definite articles. The “mid” group is different from the other two groups in that it shows a lot of substitution errors both in indefinite and definite articles.

Graph 3 also compares the three level groups, but this time, the result is shown according to the three types of article errors. The graph on the far right shows that the frequency of missing errors suddenly decreases in the “mid” group both in indefinite and definite articles. On the other hand, the chart on the far left shows that in the case of

Graph 2. Comparison of Definite & Indefinite Article Errors



Graph 3. Comparison of 3 Types of Article Errors



substitution errors, the “mid” group made more errors than the “poor” group both in indefinite and definite articles. As for redundancy errors, the frequency of errors increased in the “mid” group only in indefinite articles.

### 3-3. Language context

Finally, in order to find out the language context where the article errors are most frequent, the data was analyzed based on three different language contexts. The first context is its grammatical functions. Huebner’s (1983, 1985) categories of the English article system explained earlier seems very useful in studying the interlanguage development. Nevertheless, since the present study is more interested in finding out the relationship of article errors and grammar teaching in classrooms, I will instead use the regular grammatical categorization of articles, which has been used in grammar books students usually use, in the present study. Based on several grammar books (Biber et al. 1999, Petersen 2004, Watanuki 2006), functions of indefinite and definite articles are grouped into three; basic, extended and exceptional. Some examples from the present data are listed in Chart 7-1. Each erroneous use of the articles in the selected corpus was grouped based on the

three categories by the researcher and a native speaker of English. When there was inconsistency in the first judgment, a thorough negotiation led to an agreement.

#### Chart 7-1. Language Context (1): Grammatical Functions of the Articles

##### Indefinite Articles

###### Basic Function:

*I borrowed a book from the library yesterday.*

*Once upon a time, there was a tiger who lived in a cave.*

###### Extended Function:

*I stayed in London for a week. / Take this medicine twice a day.*

###### Exceptional Function:

*My cellphone is a Sony. / He invented a new plastic.*

##### Definite Article

###### Basic Function:

*Mary has a car. The car is red. / It's cold in here. Please shut the window.*

*The moon travels around the earth. / It is the fastest animal on land.*

###### Extended Function:

*The dolphin is a marine mammal. / The learned are not necessarily scholars.*

###### Exceptional Function:

*A police officer caught him by the arm. / Meat is sold by the pound.*

The second context to be checked in order to find whether it affects the use of articles is the distance between the article and the following noun as explained in Chart 7-2. Each erroneous article was grouped whether it was directly used in front of a noun or indirectly used in front of an adjective(s) + a noun.

#### Chart 7-2. Language Context (2): Distance between the Articles and Nouns

##### Direct Use:

*They had an illusion about the economy of Japan.*

*Nothing will happen during the pregnancy.*

##### Indirect Use:

*It is a very useful tool.*

*The greatest invention of the twentieth century is the cellphone.*

The last context to be checked is whether the article is in the part of idioms or lexical phrases as shown in Chart 7-3.

#### Chart 7-3. Language Context (3): Learned as Item vs. Learned as System

##### Learned as Item:

*It takes a couple of days to arrive somewhere.*

*There are so many people who use cellphones around the world.*

Learned as System:

*I have an idea to solve the problem.*

*The Maori are opposed to the individualism of Western society.*

The notion of “learned as item” versus “learned as system” is based on Schmitt, et al. (2004). They state that “some L1 acquirers do learn and use formulaic sequences before they have mastered the sequences’ internal makeup” (p. 11). They explain that the acquisition of formulaic sequences might depend to some extent on whether children are referential or expressive learners, that is, whether they are ‘system learners’ or ‘item learners’. They refer to Nelson (1973) and describe that children who had referential preferences (naming things or activities and dealing with individual word items) usually learned more single words, particularly nouns. Conversely, children who had more expressive tendencies (having interactional goals; focusing on the social domain) were more likely to learn whole expressions which were not segmented. Schmitt and Carter (2004) further states that “the reason for these preferences may be psycholinguistic in nature, or may only reflect what the child ‘supposes the language to be useful for’: predominantly naming things in the world or engaging in social interaction” (p. 11). Thus, in the present paper, the term “item” is used to mean learning based on chunks, idioms, or lexical phrases, while, “system” is used to mean learning based on grammatical rules.

Among the worst 12 article users, the five worst indefinite article users and the five worst definite article users were selected. Charts 8-1, 8-2, & 8-3 show the results of the

**Chart 8. Effects of Language Contexts on Indefinite Articles**

**8-1. Grammatical Functions**

Errors \ Roles	Basic Use		Extended Use		Total	
	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect
Substitution	32	2	13	1	45	3
Redundancy		1		0		1
Missing		42		11		53
Total	32	45	13	12	45	57

**8-2. Distance**

D/ID \ C/IC	Correct	Incorrect	Total
Direct	28 30.1%	33 35.5%	61 65.6%
Indirect	9 9.7%	23 24.7%	32 34.4%
Total	37 39.8%	56 60.2%	93 100.0%

**8-3. Learning Types**

	Correct	Incorrect	Total
Idioms	8 7.8%	1 1.0%	9 8.8%
Non-idioms	37 36.3%	56 54.9%	93 91.2%
Total	45 44.1%	57 55.9%	102 100.0%

• Idioms are not included here.

Chart 9. Effects of Language Contexts on Definite Articles

9-1. Grammatical Functions

Errors \ Roles	Basic Use		Extended Use		Total	
	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect
Substitution	104	15	2	0	106	15
Redundancy		29		1		30
Missing		45		8		53
Total	104	89	2	9	106	98

9-2. Distance

C/IC \ D/ID	Correct	Incorrect	Total
	Direct	80 40.2%	67 33.7%
Indirect	21 10.6%	31 15.6%	52 26.1%
Total	101 50.8%	98 49.2%	199 100.0%

9-3. Learning Types

	Correct	Incorrect	Total
Idioms	5 2.5%	0 0.0%	5 2.5%
Nonidioms	101 49.5%	98 48.0%	199 97.5%
Total	106 52.0%	98 48.0%	204 100.0%

· Idioms are not included here.

effect of language contexts on indefinite article use of the five worst indefinite article users. There was no example of exceptional function use in the files.

Charts 9-1, 9-2, & 9-3 above show the results of the effect of language contexts on definite article use of the five worst definite article users.

4. Results and Discussion

To summarize the results of the present study, answers to the research questions raised will be presented with some discussions in order.

4-1. Frequency of errors

The answer to the research question 1, "which article error type is the most frequent among the Japanese learners?" is missing errors. It is clear from Chart 3. In addition, it is also shown that definite article use is more accurate than indefinite article use.

The details of the types of the errors also show some interesting facts. As for the substitution errors, there was no example of erroneous use of *an* instead of *a*, although there were a lot of examples of erroneous use of *a* instead of *an*. And there were more examples of erroneous use of *the* instead of *a*, *an* than erroneous use of *a*, *an* instead of *the*. Thus, *an* seems to be the most marked article, then *a* follows, and finally *the* follows. They have a tendency to use *the* erroneously most often and use *a* erroneously more often than *an* whether it is grammatically correct or not. As for redundancy errors, there was more erroneous use in the definite article than in the indefinite articles. As for the missing

errors, there was more erroneous use in the indefinite articles than in the definite article.

Thus in addition to the fact that missing errors was the most frequent, it is clear from the charts that the students have a tendency to underuse indefinite articles and overuse definite article. Therefore teachers who teach Japanese university students English need to pay more attention to students' missing and substitution errors in indefinite articles and redundancy errors in definite article in classrooms.

These findings support Izumi et al.'s (2004) descriptive study, which showed exactly the same frequency order of erroneous article use by Japanese learners of English.

#### 4-2. Learner differences

As for the learner differences, Graphs 1, 2, & 3 show that it is really difficult for Japanese students to completely master correct article use. There are only three students who had no article errors among the 103 files. The error rate varies from 0% to over 60%. There are some similarities in the profiles of the three groups. The students in all levels made more indefinite article errors than definite article errors. "Poor" and "good" users made missing errors more often than the other types of errors both in indefinite and definite articles. However, there are some differences, too. The typically different profiles of each group were as follows:

1. The "poor" users had extremely frequent missing errors both in indefinite and definite articles.
2. The "mid" group increased the substitution errors both in indefinite and definite articles and decreased redundancy errors in definite articles, however, the frequency of redundancy errors in indefinite articles increased.
3. The "good" users decreased the error frequency a lot both in indefinite and definite articles, nevertheless they still had more missing and substitution errors in indefinite articles than in definite article.

Missing errors decrease as the levels go up, both in definite and indefinite articles. The frequency of errors is on the downside. However, missing errors continue to exist even at the "good" users' level. Substitution errors increase at the "mid" level once and decrease again in the upper level in both definite and indefinite articles. The frequency of the substitution errors shows a bell-shape curve. As for redundancy errors, the students made more errors in the definite article than in indefinite articles at all levels.

Based on the above observation, the progress of article learning can be inferred. It is probable that the existence of many missing errors is a possible sign of the learners' indifference to the existence of articles, because the missing errors decrease group by group from "poor" level to "good" level. Furthermore, the emergence of frequent substitution errors in the "mid" group seems to be a sign of learners' noticing of the existence of articles. It is because substitution errors take place when the students notice that some

words should be placed in the spot. The learners start to use various words including articles in front of more nouns when they get to the “mid” level.

#### 4-3. Language contexts

Surprisingly, the most erroneous use took place in the basic functions. However, it is a relief that in the case of the definite article, there was more correct use than erroneous use. Since the most errors took place in the basic functions, the most plausible reason for their erroneous use seems to be that they have difficulty in distinguishing countable and uncountable nouns. They also seem to have difficulty in having a clear idea of indefiniteness. This issue has also been pointed out by Butler (2002) and Ionin (2003). In addition, since the students learn indefinite articles later than definite articles, they made more errors when they use indefinite articles in basic function than when they use definite articles in the same function.

Distance also matters. Charts 8-2 and 9-2 show that the effect of distance is clearer in definite article use than in indefinite article use. Students made fewer indefinite article errors in direct use than in indirect use. This is shown by the fact that correct use outnumbered when the definite article was used directly in front of a noun, while erroneous use outnumbered when it was used with some other words in between.

Also it is insightful that the students made fewer indefinite article errors when the article was used within idioms or lexical phrases than without.

### 5. In Teaching Articles

Based on the findings on the effect of language context on the use of articles, the following two things seem to be important when teaching articles to Japanese students.

Firstly, Japanese is a so-called “High Context” language (Hall, 1976). Because the difference between countable and uncountable nouns does not work as a key factor in using Japanese language, the concept of countable/uncountable needs to be learned explicitly. Also, from the same reason, the difference between indefiniteness and definiteness needs to be learned. As Ishida (2002) explains, since the grammatical concept of indefiniteness/definiteness and general concept of those are sometimes different and it makes the Japanese learners confused. He compared the sentences “I want a dog.” and “I have a dog.” to show that grammatically both sentences use indefinite article *a*, “a dog” in the former sentence means an indefinite dog, while the one in the latter sentence means a definite dog in our general concept. Lyons (1999) states that “in Japanese a noun phrase marked with *wa* can only be rendered into English as definite or generic; noun phrases marked with *ga*, on the other hand, can in principle be construed as definite or indefinite. It does not follow, of course, that *wa* is a definite article, or even that a category of definiteness exists in Japanese.” In addition, demonstratives such as *sono* (the item nearer to the hearer), and

*ano* (the item far from both the speaker and the hearer) do not have one-to-one relationship with the English article system.

Secondly, item learning seems effective for correct article use, especially for beginners, because it prevents errors. On the other hand, the learners need a lot of chances to use the articles in their actual communication when they are learned as a system. When we use English, we can not add articles after we decide which noun we are going to use. Although in writing, it may be possible to monitor after completing the sentence, it is definitely impossible in speaking. Thus, as suggested by Petersen (2004), it will be a useful practice for students to decide articles first and then decide which noun they are going to use following that article. In this way, the students will have more chances to notice the existence of articles in English.

## 6. Conclusion

In this study, the frequency of article errors in the ICLE Error-Tagged Japanese Sub-Corpus was shown as descriptive data first, and, based on the features of article errors made by the three-level groups, the development of article use was suggested. Finally, some reasons for article errors by Japanese learners of English were explored. What has been pointed out in the present study are the following:

1. Missing errors were the most frequent among the three types of errors in all levels of groups. Extremely frequent missing errors in indefinite articles in the "poor" group shows the students' indifference to the existence of articles. Substitution errors increased in the "mid" group, although other types of errors smoothly decreased as the levels go up. The emergence of frequent substitution errors in the "mid" group may be a sign of students' noticing of the existence of articles.
2. Students show a tendency to overuse the definite article and underuse indefinite articles. For the students, *the* is the least marked, then *a*, and finally *an*. The learners seem to start from  $\emptyset$  article use. Then they start using the definite article *the* first, and then go on to use indefinite articles. At the second stage, they start using various words in front of nouns, including erroneous articles, adjectives, possessive forms and so on. Finally, they narrow down to the appropriate use of articles.
3. Most errors took place in the basic functions of the articles. The concept of countable versus uncountable and indefiniteness versus definiteness seems to be the main cause of the learners' confusion in using correct articles. Thus the concepts need to be taught explicitly. On the other hand, articles learned as a part of an item or a lexical phrase prevents errors. Learning idioms and lexical phrases which include an article is highly recommended. It is because the learners do not fluctuate in whether and which article should be used in the chunk because the set has been memorized as it is.

There are some limitations that need to be removed to improve the present study. It is very important to note that the frequency of errors differs depending on the task performed. The present study is based on the written corpus. This suggests that learners had chances to avoid using articles when they felt it was difficult. It is advisable to add objective tasks as a part of the data. Defining the language context of the erroneous article use was another problem. Some contexts are not easy for the researcher to judge. In addition, the contexts have been checked using only the ten files because of the limitation of the time for the data analysis. All 103 files should be analyzed to make this study more reliable. Despite such limitations, the finding that, in the end, the concepts of definiteness/indefiniteness and countable/uncountable is a key to improve article use for the Japanese learners of English is insightful. For Japanese learners of English to use articles properly, they cannot get by with avoiding to know the way native speakers change their perception of an entity depending on the context. One way to get through this problem at least for beginning level learners seems to be to learn articles as a part of idioms or lexical phrases. This finding again will shed light to the various functions idioms and lexical phrases take on.

#### References

- Asano, T. (1996). Characteristics of article acquisition in junior college students. *CASELE (the Chugoku Academic Society of English Language Education) Research Bulletin* 26, 269-274.
- Andersen, R. W. (1984). The one to one principle of interlanguage construction. *Language Learning* 34, 77-95.
- Biber, D. et al. (1999). *Longman Grammar of Spoken and Written English*. Longman.
- Butler, Y. G. (2002). Second language learners' theories on the use of English articles: An analysis of the metalinguistic knowledge used by Japanese students in acquiring the English article system. *Studies in Second Language* 24, 451-480.
- Granger, S. & F. Meunier. (2003). *Error Tagging Project Revised Guidelines*. Center for English Corpus Linguistics, Louvain Catholic University.
- Hall, E. T. (1976). *Beyond Culture*. Doubleday.
- Huebner, T. (1983). *A Longitudinal Analysis of the Acquisition of English*. Karoma Publishers.
- Huebner, T. (1985). System and variability in interlanguage syntax. *Language Learning* 35, 141-163.
- Ionin, T. (2003). Article semantics in second language acquisition, Ph. D. thesis, MIT. Distributed by *MIT Working Papers in Linguistics*.
- Ionin, T., H. Ko & K. Wexler. (2004). Article semantics in L2 acquisition: The role of specificity. *Language Acquisition* 12 (1), 3-69.
- Ishida, H. (2002). *EOIGO-KANSHI KOUGI*. Taishukan.
- Izumi, E., K. Uchimoto & H. Isahara. (2004). NIHON-JIN EIGO GAKUSHU-SHA NO EIGO KANSHI SHUTOKU KEIKOU NO BUNSEKI, in Izumi, et al. (eds.) *NIHON-JIN 1200 NIN NO EIGO Speaking Corpus*. National Institute of Information and Communications Technology.
- Lyons, C. (1999). *Definiteness*. Cambridge University Press.
- Nelson, K. (1973). Structure and strategy in learning to talk, in *Monographs of the Society for Research in Child Development*, Serial no. 149, nos. 1-2.

- Petersen, M. (2004). *NIHON-JIN NO EIGO*. Iwanami.
- Schmitt, N. & R. Carter. (2004). Formulaic sequence in action, in N. Schmitt (ed.) *Formulaic Sequences: Acquisition, Processing and Use*. John Benjamins.
- Takahashi, T. (2000). The effect of count/uncount distinction on the choice of the indefinite English article measured by sentence verification test. *ARELE (Annual Review of English Language Education in Japan) 11*, 1-10.
- Watanuki, Y. (2006). *The Royal English Grammar for Practical Expressiveness*. Obunsha.

(金子 朝子 英語コミュニケーション学科)