Japanese Learners' Characteristic Errors of the Chinese Numeral Classifier Sequence *yi ge*: Learning and/or Unlearning?

Xiaochun Teng, Yasuo Ogawa, Jun Yamada

Hiroshima University, Japan

Abstract This study was motivated by Higuchi (2007), in which many advanced learners of Chinese made incorrect responses to seemingly easy items in a two-choice cloze-type test of the Chinese numeral classifier sequence *yi ge*. In Experiment 1, we tested native Chinese speakers as subjects, using the same material but allowing three possible choices for each item: (1) *yi ge*, (2) zero, or (3) *yi ge*/zero (i.e., either acceptable). The results revealed that there were several items which elicited a *yi ge*/zero response, thereby indicating that the validity of a two-choice test is questionable. In Experiment 2, both a two-choice and a three-choice test were administered to Japanese learners of Chinese. The results showed that the subjects made more correct responses to the 'easy' items marking new information status than did Higuchi's subjects. Combined with Higuchi's findings, the present results may be interpreted as suggesting that Japanese learners in a beginning or intermediate stage first acquire the association between *yi ge* and new information status of the following noun, but unlearn that simple association at a more advanced level. The grammatical features which seem to engender this learning pattern are briefly discussed. **Keywords**: *yige*, two- and three-choice cloze test, Japanese learners of Chinese, U-shaped learning

提要 樋口（2007）通过二选一（φ 或者“一个”）的选择填空测试发现，很多具有高级水平的日本汉语学习者却不能正确填写有些似乎很容易的空栏。本研究在樋口（2007）的基础上进行了进一步的考察。在实验 1 中，我们以汉语母语者为被试，用樋口（2007）的调查材料作为材料进行了测试。不过我们的测试方法是三选一（φ、“一个”或者“φ/一个”）。结果揭示，在好几个空栏中，被试倾向于填写“φ/一个”。因此，二选一的选择填空测试有一定的问题。在实验 2 中，我们以日本的汉语学习者为被试进行了二选一和三选一这两种测试。结合樋口（2007）的发现，本研究的结果显示，在相对初期的学习阶段，学习者能够习得“一个”导入新信息名词的用法，但是到了较高级的学习阶段反而不能正确使用该用法。我们将对日本的汉语学习者为什么会产生这样的情况进行初步的探讨。 **关键词**：一个，二选一和三选一完形填空测试，日本汉语学习者，U 型习得曲线

1. Introduction

The Chinese numeral classifier sequence *yi ge* is used in intricate ways depending upon the linguistic context in which it appears, and a great amount of research has been conducted regarding its syntactic (Lü, 1989; Nakagawa & Li, 1992), semantic (Wang, 1954; Okochi, 1985; Chen, 2003), and pragmatic (Sun, 1988; Li 2000, Chen, 2003; Biq, 2004) aspects. However, while such studies certainly have enhanced our understanding of this sequence, little research has been devoted to the issue regarding to what extent or how non-native speakers such as Japanese learners of Chinese are able to use it correctly at the text level.

Higuchi (2007) is one of the few who addressed this issue. Using a cloze-type test in which all of the tokens of *yi ge* and potential locations before nouns appearing in a Chinese translation of a well-known Japanese fairy tale *Momo Taro* (*The Peach Boy*) were replaced with blanks, she asked intermediate- and advanced-level Japanese college students of Chinese to put a circle in the blanks if *yi ge* is appropriate and to put an x if inappropriate (see Table 1 below). The results showed that there were seven items out of 30 (23%) to which more than 21 subjects (50%) made incorrect responses. Higuchi (2007) thus concluded that *yi ge* is difficult even for intermediate- and advanced-level students.

We agree with Higuchi that *yi ge* is a difficult item in Chinese, but some of the error items reported in her study do not appear consistent with our intuitions or what our teaching experience suggest. Specifically, 50% or more of her subjects incorrectly responded to items 3 and 4 (__老爷爷 and __老奶奶, both of which carry new information) by choosing the zero classifier (see Table 3). In our judgments, these items seem to be easy for beginning- and intermediate-level learners. If so, where does the discrepancy between Higuchi (2007) and our impression come from?

Actually, there may be no discrepancy. One possibility is that Japanese learners of Chinese may exhibit a kind of U-shaped learning curve when they learn the use of *yi ge*. It may be the case that while beginning- and intermediate-level students first learn a simple and rough rule of *yi ge* and correctly use it most of the time in simple contexts, advanced students, faced with more complex contexts, unlearn it and repeatedly engage in trials and errors until they eventually acquire a correct set of rules. We will return to this issue in the Discussion section.

On the other hand, we note a possible problem with the Higuchi test, which had only two choices for each test item although our and other native Chinese speakers' judgments suggest that some items in the test are ambiguous, i.e., either the numeral classifier sequence *yi ge* or the zero classifier Φ are acceptable de-

pending on the interpretation of the context. If an *yi ge* response (+c response, hereafter) is appropriate for some items, a zero response (-c response) for others, and either an *yi ge* or zero response (+/-c response) for still others, a two-choice test would fail to assess learners' true knowledge of this classifier.

The present study consisting of two experiments is a replication and extension of Higuchi (2007). The first experiment, using native Chinese speakers as subjects, attempted to ascertain whether there are ambiguous cases which allow +/-c responses to occur in Higuchi's two-choice test. In the second experiment, we replicated the Higuchi (2007) study using Japanese learners of Chinese, but, in addition to a two-choice test, we gave subjects a three-choice test, where (1) only a +c response is appropriate, (2) only a -c response is appropriate, or (3) a +/-c response is acceptable. If we confirm in the first experiment that some of the cloze items are indeed ambiguous and if the second experiment shows that results are significantly different between the two-choice and the three-choice test, we would be in a better position to account for why our intuitive judgment for some apparently easy items is not consistent with Higuchi's findings.

2. EXPERIMENT 1

The aim of this experiment is to determine whether there are ambiguous items in a two-choice cloze-type test of *yi ge* as employed by Higuchi (2007).

2.1. Method

2.1.1. Subjects

Eighteen native Chinese speakers who were studying in graduate programs at Hiroshima University participated in this study. Their ages ranged from 22 to 35 years.

2.1.2. Materials

The text and test blanks used in this study were exactly the same as those used by Higuchi (2007) (the test format which was originally constructed by Okochi, 1985) except that we had three choices instead of two choices. The instructions given were that "you put a circle (○) in the blank if you think *yi ge* should appear there, an x (×) if you think *yi ge* should not appear there, and a triangle (△) if you think either *yi ge* or no classifier can be acceptable." As stated in the Introduction, we will refer to those responses as a +c, a -c, and a +/-c response, respectively. The test format is presented in Table 1.

Table 1: The Test Material

很久很久(1)以前，在(2)地方，住着(3)老爷爷和(4)老奶奶，(5)夏日的一天，(6)老爷爷要上(7)山打(8)柴。(9)老奶奶："快去快回啊！"

(10)老爷爷走后，(11)老奶奶又自言自语："哎，我也到(12)河边洗(13)衣服去吧。"说完她就端着(14)盆到(15)河边去了。哗啦哗啦，哗啦哗啦，(16)老奶奶使劲地洗。

洗了一会儿，忽然看见有(17)东西从(18)上游一起一浮地飘下来。(19)老奶奶使劲地停住(20)手，歪着(21)脑袋思忖起来。那(22)东西圆乎乎的，有(23)西瓜那么大；白里透绿，绿　泛红。说它像(24)桃子吧，却比(25)桃子大；说它像(26)瓜吧，又比(27)瓜圆。就在她想着的(28)当儿，那(29)东西已经飘过来，可以看得清清楚楚了。原来是很大很大的(30)桃子。

Translation:

Once upon a time, there lived an old man and old woman in a country. On a summer day, the old man went into the mountain to cut wood. His wife said to him, "Take care," and then said to herself, "Well, I'll go to the river to wash clothes." She went there with a washtub and clothes. She washed the clothes there. She soon found something floating down the river. She stopped washing and wondered what it was. It was a round object. It was as big as a watermelon. It was white, green, and pale-red. It was too big to be a peach, and too round to be a melon. It came closer to her and she found it to be a big peach.

2.1.3. Procedure

The subjects were tested individually or in small groups with no discussion allowed among subjects. It took approximately 15 minutes to complete the task..

2.2. Results and discussion

The distribution of responses is shown in Table 2.

Table 2: Distribution of Chinese Speakers' Responses in the Three-Choice Test

		No. of Responses (*N*=18)		
	Noun	-c (×)	+c (○)	+/-c (△)
1	以前	18	0	0
2	地方	0	18	0
3	老爷爷	0	18	0
4	老奶奶	0	16	2
5	夏日	15	0	3
6	老爷爷	18	0	0
7	山	18	0	0
8	柴	18	0	0
9	老奶奶	18	0	0
10	老爷爷	18	0	0
11	老奶奶	18	0	0
12	河边	18	0	0
13	衣服	18	0	0
14	盆	6	0	12
15	河边	18	0	0
16	老奶奶	18	0	0
17	东西	0	18	0
18	上游	18	0	0
19	老奶奶	18	0	0
20	手	18	0	0
21	脑袋	18	0	0
22	东西	0	2	16
23	西瓜	1	2	15
24	桃子	4	1	13
25	桃子	14	0	4
26	瓜	4	1	13
27	瓜	14	0	4
28	当儿	15	0	3
29	东西	4	1	13
30	桃子	0	18	0

Table 2 shows that substantial numbers of subjects took items 14, 22, 23, 24, 26, and 29 as +/-c, i.e., ambiguous. The binominal test indicated that +/-c responses were significantly greater than the other response(s) (+c and/or -c responses), for items 22 and 23 at the 1% level and for items 24, 26, and 29 at the 5% level. For item 14, no significant difference was observed, $p > .2$, suggesting that +/-c and -c responses are equally acceptable.

We also see from Table 2 that although the numbers of subjects who made +/-c responses were small, items 4, 5, 25, 27, and 28 may be regarded as more or less ambiguous. In sum, 11 items out of 30 (37%) allow a +/-c response. It is thus concluded that the validity of Higuchi's two-choice test is questionable. But the question remains why many advanced-level learners made errors on such easy items.

3. EXPERIMENT 2

The question addressed in this experiment was how Japanese learners of Chinese respond to a two-choice cloze test and a three-choice cloze test using the same text and the same blanks.

3.1. Method

3.1.1. Subjects

Twenty-two Japanese students studying Chinese as a second language at Hiroshima University participated in this experiment. Fifteen were 21- to 22-year-old undergraduates who majored in Chinese language/literature and had studied Chinese for three years. These students may be regarded as beginning- or intermediate-level students. The remaining seven subjects were graduate students aged 22 to 32 who had studied Chinese for more than four years, one majoring in Chinese linguistics, four in Chinese literature, and two in others. Their knowledge of Chinese was generally considered to be somewhat better than that of the undergraduates.

3.1.2. Materials

The two-choice test was the same as those used by Higuchi (2007) and the three-choice test was the same as that in Experiment 1.

3.1.3. Procedure

Fifteen undergraduate subjects first took the three-choice test. The instructions were the same as those given in the first experiment. The subjects were

asked to fill in each blank with a circle if *yi ge* is appropriate, an x if the zero classifier is appropriate, or a triangle if the item is ambiguous allowing both possibilities. After completing this three-choice test, they were asked what they would do if they had to choose either *yi ge* (a circle) or a zero quantifier (an x) in each item. Each subject was asked put a circle or an x next to his or her previous response for each item. For the seven graduate students, only the two-choice test was administered.

Each subject was tested individually. It took approximately 30 minutes to complete the two versions and approximately 25 minutes to complete the two-choice test.

3.2. Results and discussion

Table 3 shows error rates in the present studies' two-choice test and Higuchi's (2007) two-choice test ($N = 22$), and those in the present three-choice test ($N = 15$).

We compared the mean error rates between the subjects of this study and those of Higuchi (2007) in the two-choice test condition (see the first and second columns of Table 3). The means were 24.7% and 23.5%, respectively, indicating no significant difference, $t(29) = 0.37$. This result may appear to suggest that the knowledge of the Chinese numeral classifier sequence *yi ge* was essentially the same between the two subject groups. However, attention should be paid to items 3 and 4 which exhibited significant group differences; for these items, the error rates were 4.5% and 9.1% in this study compared to 50.0% and 64.3% in the Higuchi study, respectively. Insofar as these items are concerned, the subjects of this study outperformed those of the Higuchi study. Because these items could be regarded as "outliers," we recomputed the error rates with these two items excluded. The results then revealed that the mean was 26.0% ($SD = 20.1$) for this study and 21.1% for the Higuchi study, the latter being significantly lower than the former, $t(27) = 2.34$, $p < .05$. Aside from items 3 and 4, therefore, it is reasonable to conclude that Higuchi's subjects constituted a more advanced learner group than did the present subjects.

Table 3
Mean Error Rates in the Two-Choice Test
and Mean Response Rates in the Three-Choice Test

		Error Rate (%)		Response Rate (N=15) (%)		
	Word	This Study*	Higuchi	-c (×)	+c (○)	+/-c (△)
1	以前	4.5	0.0	93.3	6.7	0.0
2	地方	13.6	7.1	13.3	73.3	13.3
3	老爷爷	4.5	50.0	0.0	66.7	33.3
4	老奶奶	9.1	64.3	0.0	66.7	33.3
5	夏日	13.6	11.9	86.7	6.7	6.7
6	老爷爷	13.6	2.4	73.3	13.3	13.3
7	山	18.2	19.0	60.0	13.3	26.7
8	柴	18.2	11.9	86.7	0.0	13.3
9	老奶奶	18.2	2.4	80.0	6.7	13.3
10	老爷爷	13.6	2.4	80.0	6.7	13.3
11	老奶奶	9.1	4.8	86.7	6.7	6.7
12	河边	18.2	28.6	66.7	13.3	20.0
13	衣服	18.2	14.3	73.3	13.3	20.0
14	盆	9.1	26.2	33.3	46.7	6.7
15	河边	13.6	19.0	80.0	13.3	20.0
16	老奶奶	13.6	2.4	73.3	20.0	13.3
17	东西	22.7	23.8	26.7	53.3	20.0
18	上游	9.1	0.0	86.7	6.7	6.7
19	老奶奶	18.2	2.4	80.0	13.3	6.7
20	手	45.5	14.3	66.7	26.7	6.7
21	脑袋	27.3	9.5	46.7	26.7	26.7
22	东西	63.6	66.7	46.7	20.0	33.3
23	西瓜	63.6	54.3	46.7	40.0	13.3
24	桃子	50.0	66.7	26.7	20.0	53.3
25	桃子	18.2	19.0	40.0	13.3	46.7
26	瓜	45.5	23.8	26.7	13.3	60.0
27	瓜	22.7	9.5	46.7	6.7	46.7
28	当儿	9.1	0.0	80.0	20.0	0.0
29	东西	77.3	85.7	53.3	13.3	33.3
30	桃子	59.1	61.9	46.7	26.7	26.7
	Mean	24.7	23.5	56.9	22.2	20.9
	SD	20.0	24.8	26.9	19.9	15.5

In the Introduction, we stated that items 3 and 4 seem to be easier for beginning- and intermediate-level learners of Chinese. In fact, we note from the results of Table 3 that even in the three-choice test, two thirds of our subjects correctly made a +c response for these items and that most of the remaining one third who made a +/-c response made a correct +c response in the two-choice test. Why then did the good learners of Chinese in the Higuchi (2007) study produce such high error rates for these particular items? It may be possible to provide a plausible answer to this by further analyzing our subjects' responses in the three-choice test condition.

We conducted a correlation analysis for the 30 test items, counting the number of subjects who made a +c, -c, or +/-c for each item. The results are presented in Table 4.

Table 4
Correlation between +c, -c, and +/-c Response

	+c	−c	+/−c
+c	-		
−c	-.81**	-	
+/−cq	.13	-.68*	-

*$p < .01$, **$p < .001$

Of interest was a significant negative correlation of -.68, $p < .01$, between the -c and +/-c responses. This is taken to imply that there was a tendency for the subjects to make either a -c response or a +/-c response in the three-choice test although the tendency was not so strong as the one between the +c and -c responses. Thus, it would follow that the subjects were more likely to make a –c response when a +/-c response was not permitted in the two-choice test. Now if we assume that this tendency generally applies to advanced-level learners, many of Higuchi's subjects might incorrectly have taken the linguistic contexts involving test items 3 and 4 as ambiguous. If so, many of such subjects would have been more likely to make a -c response for these items in the two-choice test. On the other hand, many subjects in the present study who were less advanced than Higuchi's subjects did not see these items as ambiguous.

This line of reasoning allows the following hypothesis for a developmental course of acquisition of *yi ge* in Japanese learners. The learner may first learn a simple use of *yi ge* marking new information status of the following noun, which is most frequently observed in most contexts at the beginning and intermediate levels. The learner, however, may later learn a more complex case where the

noun carries new information but both +c and -c responses are acceptable. At this advanced stage, he/she may make an 'advanced-level' mistake concerning a simple case in which only a +c response is appropriate. Such a case would exhibit a kind of U-shaped acquisition pattern. Learning of items 3, 4, and 14 may be an example of this learning pattern.

4. GENERAL DISCUSSION

The results obtained in this study may be taken as indicating that beginning- or intermediate-level learners of Chinese first learn the simple one-to-one association between *yi ge* and new information whereas more advanced unlearn that simple association and seek for a more accurate set of rules but make trials and errors at this stage. If such is the case, Japanese learners of Chinese may exhibit a case of U-shaped language learning (See Brown, 1973, Rumelhart & McClelland, 1986, and Taatgen & Anderson, 2002, for a case of English-speaking children learning the English past tense which is taken as a paradigm example of U-shaped learning). In the case of English-speaking children learning the English past tense, the irregular verb *feel*, for example, may be acquired in the following order: *felt* (and *feel*) → *feeled* (and *felt*) → felt. (cf. Pinker, 1995). In the case of Japanese learners learning *yi ge* in the context involving test items 3 and 4, the acquisition order may be represented by (1) only *yi ge* → (2) either Φ (and *yi ge*) → (3) only *yi ge*. In the present small-scale pilot study, however, it is difficult to spell out in detail the mechanisms underlying this possible U-shaped development. We thus present our tentative speculation on a three-stage learning pattern here. At stage 1, beginning- and intermediate-level learners learn grammatical features such as [±count] (i.e., +count and -count) and [±new] (i.e., -new = old) for Chinese nouns. They also learn that *yi ge* goes with the features [+count] and [+new]. This rule applies to virtually all examples which appear in the elementary texts to which they are exposed. The linguistic fact, however, is that numeral-classifier usage is not that simple. At stage 2, advanced-level learners encounter examples to which the above simple rule fails to apply. For example, in sentences 1 and 2 below, the nouns underlined do not necessarily require *yi ge* even though they are thought of as having the features [+count] and [+new]. At this stage, if the feature [+definite] is added, *yi ge* would become optional.

(1) 一只小白鸽，把四封同样内容的叶形信，分别送给了音乐家、画家、诗人和哲学家。

"A little white dove sent respectively to the musician, the painter, the

poet, and the philosopher a leaf-shaped letter in which the same message was written."

(2) 老虎大王派出了自己的飞毛腿传令兵——羚羊，把布告贴遍整个森林。

"King Tiger ordered the antelope, the fastest runner in the forest, to put up the posters all over the forest."

Another set of examples where the rule involving the [±count] and [±new] features is not sufficient for correct use may be found in sentences 3 and 4.

(3) 现在，我走在荒凉的野地里，蝙蝠在我的头上飞来飞去，猫头鹰在树上怪叫。

"I was now walking along in a wasteland where bats were flying overhead and an owl was hooting in a tree."

(4) 我坐在喷泉边，只见喷泉的水花闪着银光，像撒着无数的珍珠。

"I sat by the fountain. The spray of water from the fountain was glittering like pearls."

These nouns in sentences 3 and 4 carry the feature [-focus] and can be used with the zero classifier. For example, the *owl* in sentence 3, having the features [+count], [+new], and [-focus], may not be preceded by *yi ge*. Note that the [±new] feature is not the same as the [±focus]; and thus, the feature [+new] is not incompatible with the feature [-focus].

The features [±definite] and [±focus] are not so easy to learn, and therefore learners may commit errors when they encounter sentences such as 1 to 4. We speculate that Higuchi's (2007) subjects, who were at this stage, might incorrectly have taken 老爷爷 in item 3 and 老奶奶 in item 4 as having the [+definite] and/or [-focus] feature in this context. Such being the case, because in the two-choice test, +/-c responses are not allowed and because of the tendency indicated by the negative correlation between +/-c and -c, many of Higuchi's subjects made –c responses to items 3 and 4.

At stage 3, after this trial and error stage, learners would eventually learn the features [±definite] and [±focus] and complete the U-shaped learning process.

5. CONCLUSION

This study demonstrated that when we assess second-language learners' knowledge of the Chinese numeral classifier sequence *yi ge*, the validity of a two-choice cloze-type test used by Higuchi (2007) and others is questionable. We proposed that a three-choice test, where all three +c, -c, and +/-c responses are allowed, assesses the learners' knowledge more accurately. Furthermore, this study suggests that Japanese learners may exhibit a U-shaped learning pattern in learning this classifier; for example, they may differentially learn grammatical features such as [±definite] and [±focus] of *yi ge* from intermediate to advanced levels.

There are several limitations in this pilot study and the generality of the present findings remain uncertain. Further research using many more test items and learners of Chinese as a second language is needed.

Acknowledgments

The authors are grateful to Peter Skaer of Hiroshima University, anonymous reviewers, Zheng-sheng Zhang, the editor of this journal, for their helpful comments and suggestions on an earlier draft.

References

Biq, Y-O. (2004). Construction, reanalysis, and stance: 'V yi ge N' and variations in Mandarin Chinese. *Journal of Pragmatics*, 36, 1655-1672.

Brown, R. (1973). *A First Language: The Early Stages*. Cambridge, MA: Harvard University Press.

Chen, P. (2003). Indefinite determiner introducing definite referent: a special use of '*yi* 'one' + classifier' in Chinese. *Lingua*, 113, 1169-1184.

Higuchi, S. (2007). Suryochi no hyogen goyo ni kansuru ichi kousatu. *Chugo kugo Kyoiku*, 5, 46-69 [A study of errors in classifiers. *Chinese Language Education*].

Li, W. (2000). The pragmatic function of numeral-classifiers in Mandarin Chinese. *Journal of Pragmatics*, 32, 1113-1133.

Lü, S. (1984). Ge zi de yingyong fanwei, fu lun danwei ci qian yi zi de tuoluo. In S. Lü (Ed.). *Hanyu yufa lunwenji* (*zengdingben*), pp. 145-175. Beijing: Shangwu Yinshuguan. [The scope of ge and the phenomenon of dropping of yi in yi ge.

In *A collection of papers on Chinese grammar*].

Nakagawa, M. and Li, (1992), Nicchu ryogengo ni okeru suryoshi hyogen. *Ni hongo to chugokugo no taisho kenkyu ronbunnshu*, pp. 95-116. Tokyo: Ku roshio Shuppansha [Classifiers in Japanese and Chinese. *Papers on Contrastive Linguistics on Japanese and Chinese*].

Okochi, Y. (1985). Ryoshi no kotaika kino. *Chugokugogaku*, 232, 1-13. [The function of individualization in classifiers. *Bulletin of the Chinese Language Society of Japan*].

Pinker, S. (1995). Why the child holded the baby rabbits: A case study in language acquisition. In L. R. Gleitman and M. Liberman (Eds.). *Language: An Introduction to Cognitive Science*, 2nd ed. Vol. 1, pp. 107-133. Cambridge, Mass.: MIT Press.

Rumelhart, . E. and McClelland, J. L. (1986). On learning the past tense of Eng lish verbs. In J. L. McCelland & D. E. Rumelhart (ds.). *Parallel Distributed Processing: Exploring in the Microstructure of Cognition*, pp. 216-271. Cambridge, MA: MIT Press.

Sun, C. (1988). The discourse function of numeral classifiers in Mandarin Chinese. *Journal of Chinese Linguistics*, 16, 298-323.

Taatgen, N. A. and Anderson, J. R. (2002). Why do children learn to say "Broke"? A model of learning the past tense without feedback. *Cognition*, 86, 123-155.

Wang, L. (1954). "Yi", "yi ge." In L. Wang (Ed.). *Zhongguo xiandai yufa (xia)* [*Modern Chinese grammar (II)*], pp. 113-135. Beijing: Zhonghua Shuju.