An Investigation of English Articles’ Acquisition by Chinese Learners of English

Li Haiyan
Wendeng Normal College
YANG Lianrui
Ocean University of China

Abstract
The process of article acquisition is one of the foci of SLA studies. The purpose of this study is to investigate the articles’ accuracy order and usage pattern by Chinese learners of English and identify possible patterns for the development of the ability to use English articles. It was found that the ability to use the develops earlier than the ability to use a/an, and the ability to use a/an develops earlier than the ability to use “zero article” Ø. These findings help us understand why lower-intermediate and intermediate learners tend to overuse the and a/an and underuse Ø, while advanced learners often overuse Ø. For intermediate Chinese learners of English, tasks that require them to choose articles before generic nouns ([−SR, +HK]) contexts and idioms are marked. For advanced learners, article choices in [−HK] contexts can be problematic. The findings presented in this study may be valuable for those involved with English language instruction in China.

Key words: articles’ acquisition; accuracy order; usage pattern; difficulty

1. Introduction
Articles such as a/an, the, and the zero article (i.e., any instance in which a noun does not require an article) are among the most commonly used words in the English language. English articles are usually introduced to ESL/EFL learners within the first few lessons. While more challenging structures can be acquired relatively early, it may take an ESL/EFL learner many years to become proficient in the proper use of articles.
There has been a considerable amount of research published in international journals on English article proficiency for learners of English in ESL contexts (see, for example, Bulter, 2002; Chaudron, & Parker, 1990; Hakuta, 1976; Huebner, 1983; Lang, 1998; Lee et al., 1994; Lu, 2001; Master, 1987; Parrish, 1987; Thomas, 1989; Zegarac, 2004). Such research suggests that language learners develop the ability to use the definite article the before a/an (Thomas, 1989) and that the accuracy order for the article system follows the sequence the, a/an and zero article for language learners whose native languages have an article system. For learners whose native languages have no article system, the accuracy order for the article system follows the sequence zero article, the and a/an (Master, 1987).

Chinese is generally acknowledged as a language with no functional equivalent of the English article system. In the Chinese language, definiteness and indefiniteness are made apparent by word order or by the use of determiners such as zhe (this), na (that), and yi (one) (Robertson, 2000). For Chinese speakers, these seemingly simple morphemes, the, a/an, and zero article, are usually not mastered until a very late stage of English learning.

The acquisition of the English article system by Chinese learners has been the focus of various studies in China in recent years (see, for example, Cai, 2007; Cai, & Wu, 2006; Li et al., 2001; Wang, 2005; Yan, 2003; Zhou, 2007; Zhu, 2003, 2008a, 2008b). Few of these studies, however, directly address the issue of acquisition patterns. Considering this fact, this study was designed to investigate the accuracy and usage patterns of the three articles by Chinese English language learners, and the specific barriers that they encounter during the learning process. At last, the pedagogical implications of our findings are to be identified.

### 2. Literature Review

Huebner (1979, 1983) used Bickerton’s (1975) notion of the dynamic paradigm to argue that early morpheme studies (Bailey et al., 1974; Brown, 1973; Dulay & Burt, 1973; Larsen-Freeman, 1975), which considered only obligatory contexts, were unable to discern variation in the use of morphemes in an evolving interlanguage, and thus failed to delineate a complete picture of acquisition processes.

![Figure 2-1. Bickerton’s semantic wheel for noun phrase reference (Huebner, 1983)](image)

Huebner (1983) believed that the obligatory/optional/ungrammatical trichotomy used in early studies was not refined enough for article analysis. Therefore, Huebner (1983)
adopted Bickerton’s semantic wheel model \(^1\) (Figure 2-1), which suggested that pre-noun contexts be classified in terms of four semantic categories, \([\pm\text{Specific Referent (SR)}]\) and \([\pm\text{Assumed Known to the Hearer (HK)}]\) (see Table 2-1 for an example drawn from the test items in Master, 1994).

In his longitudinal naturalistic study, Huebner (1983) investigated the use of the definite article *da* by observing his subject, Ge, an adult Hmong speaker with basic-level English proficiency. After a year of observation, Huebner identified six stages in Ge’s learning trajectory of marking *da* for noun phrases (NPs). His findings provided research evidence for systematic variability, rather than random choice for article acquisition by second language learners.

### Table 2-1. Environments and examples for the semantic categories \([\pm\text{SR}, \pm\text{HK}]\) (Master, 1994)

<table>
<thead>
<tr>
<th>Category</th>
<th>Article</th>
<th>Environment</th>
<th>Example (drawn from the test items in Master, 1994)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. [−SR,+HK]</td>
<td>the, a, Ø</td>
<td>Generics</td>
<td>The favorite food of the jaguar is the wild pig. Ø Wild pigs move in bands of fifteen to twenty.</td>
</tr>
<tr>
<td>2. [+SR,+HK]</td>
<td>the</td>
<td>Unique, previously mentioned, or physically present referents</td>
<td>What is the diameter of the moon? Once there were many trees here. Now, the trees are gone. The air in this city is not very clean.</td>
</tr>
<tr>
<td>3. [+SR,−HK]</td>
<td>a, Ø</td>
<td>First-mention NPs, or NPs following existential &quot;has/have&quot; or &quot;there is/there are&quot;</td>
<td>I would like a cup of coffee, please. I always drink Ø water with my meals. There is an orange in that bowl.</td>
</tr>
<tr>
<td>4. [−SR,−HK]</td>
<td>a, Ø</td>
<td>Equative NPs, or NPs in negation, question, or irrealis mode</td>
<td>What is the sex of your baby? It’s a boy! Einstein was a man of great intelligence.</td>
</tr>
</tbody>
</table>

Huebner opened up a new avenue of research on L2 article acquisition by employing Bickerton’s (1981) noun classification system. He considered not only the presence or absence of articles in obligatory contexts, but also analyzed various types of NPs, and the articles used with each semantic type, as well as the development of ESL learners’ article system. Most developmental studies of article acquisition since then have used this approach.

Some research scholars (Master, 1987; Parrish, 1987) indicated that Ø is often overused in the early stages of L2 acquisition. They hypothesized that the ability to use zero article was acquired earlier than the other two articles. Liu & Gleason (2002) reexamined Master’s (1987) data and offered a new explanation for the overuse of the zero article and underuse of *a* and *the*. They suggested that the learners got early and accurate control of the indefinite article, which is also inconsistent with the result of Thomas (1989), i.e., the definite article *the* was possibly the first to be acquired.

According to Lu (2001) there are three reasons for this kind of inconsistency. First, different methodological approaches were used by different investigators. Some studies were longitudinal (Huebner, 1983; Parrish, 1987), while others were pseudo-longitudinal (Master, 1987), or cross-sectional (Thomas, 1989; Yamada & Matsuura, 1982) studies. Second, there is no universal standard for participants’ English proficiency level. Third,
An Investigation of English Articles’ Acquisition by Chinese Learners of English

participants’ mother language was not consistent across studies; languages represented include Chinese, Japanese, Korean, Finnish, and Russian. Hence, general development patterns that have been identified should be verified for learners with different linguistic backgrounds, in different situations.

3. Methodology

3.1 Research questions
1. What are the accuracy and usage patterns of the English articles in Chinese EFL learners’ interlanguage?
2. In what semantic contexts are there areas of difficulty underlying article choice for Chinese learners of English?

3.2 Instruments
Two methods, tests and interviews, were used to collect data for both quantitative and qualitative analysis. Both a fill-in-the-blank test and a cloze test were employed to determine participants’ accuracy rate for article usage. These two tests were adopted from Master (1994), which were designed to assess non-native English speakers’ ability to use articles. The researchers revised some items for the following reasons: First, only common nouns were bounded in Master’s tests, which were based on Bickerton’s semantic categories, suggesting that pre-noun contexts should be classified in terms of four semantic categories, \([±\text{Specific Referent (SR)}] \) and \([±\text{Assumed Known to the Hearer (HK)}]\) ( \([−\text{SR, +HK}]\), \([+\text{SR, +HK}]\), \([+\text{SR, −HK}]\), and \([−\text{SR, −HK}]\). Bickerton’s semantic categories only considered literal use of language. In order to get a more complete picture of Chinese learners’ article use, a fifth item (idioms) was added as another category by the researcher as in Thomas (1989) and Butler (2002), since the articles’ usage in such context is conventional. Second, some items which allow two possible options were excluded. For example, “If you want to read, why don’t you turn on—light?” Here both the and a were right depending on whether the light is assumed known to the hearer. Third, to avoid a double penalty, some items were not included in the test. For example, “the” in the title of the cloze “The jaguar and the wild pig” were certainly repeated in the paragraph. Three groups of subjects were asked to finish the two tests in class. The interviews were conducted in order to gather information to support the data gathered from the quantitative research.

3.3 Participants and procedures
Pseudo-longitudinal data were collected from 121 Chinese learners of English (41 males and 80 females, aged 17-24). The participants were divided into three groups each of which represented a different English proficiency level. This allowed for inter-group comparisons for article acquisition. Prior to the test, the participants were asked to respond to the test items as spontaneously as possible. They were not allowed to use dictionaries, and discussion was not permitted. On average, it took them approximately 10-15 minutes to
finish the test (there was no time limit). Out of the data analysis convenience, we randomly sampled 30 test papers from the papers of 42 second-year undergraduates studying at a Normal School (Teachers’ College), who were of the same level as second-year high school students. We also sampled 30 papers from a group of 52 second-year undergraduates (English minors) and 20 from a group of 27 fourth-year undergraduates (English majors). Hence, a total of 80 papers were selected for analysis.

Immediately after the test, four students from each group were randomly selected to be interviewed individually. One question was proposed: Which one do you think is most difficult? Please elaborate the possible reasons. Notes taken during the interviews were analyzed.

3.4 Analyses
Data analyses of the test results included both TLU (Target-like Use) and UOC (Used in Obligatory Contexts) measures. Various morpheme studies (Andersen, 1977; Bailey, Madden & Krashen, 1974; Dulay & Burt, 1973; Hakuta, 1976; Larsen-Freeman, 1975; Master, 1987; Parrish, 1987; Thomas, 1989) have used these two measures. SOC (Supplied in Obligatory Contexts) (Brown, 1973) has also been used widely. It simply indicates how often an item is accurately used when its use is required. For example, in the sentence, “Please answer ____ telephone,” use of the article the would result in a SOC score of 100%, whereas the use of a or Ø would result in a score of 0%. The total of all the scores for a given passage would be the SOC score for accuracy for the use of the.

\[
SOC = \frac{\text{Number of correct compliances in obligatory contexts}}{\text{Number of obligatory contexts}} \times 100
\]

SOC has, however, been criticized for its failure to consider over-suppliance of a morpheme in non-obligatory contexts (Anderson, 1977; Hakuta, 1976; Pica, 1983). If a morpheme is overused or overgeneralized, the SOC score will overestimate the learner’s accuracy. The TLU (Target-Like Use) measure was designed to redress this potential inflation of SOC scores. It was formulated as follows (Pica, 1983):

\[
TLU = \frac{\text{number of correct suppliances in obligatory contexts}}{(\text{number of obligatory contexts}) + (\text{number of suppliances in non-obligatory contexts})} \times 100
\]

This formula shows how the TLU measure reduces the SOC score by increasing the value for the denominator in all cases when an item is used when it should not have been. For example, if a learner encounters ten contexts that require use of the article the, five contexts that require a/an, and five contexts requiring Ø, he/she responds with the five times in the ten obligatory contexts, the three times in contexts which require a/an, and the two times in Ø contexts. His/her SOC score for the would be 50%. This score does not reflect his/her misuse of the in a/an and Ø contexts. The TLU score for his/her use of the article the would account for the erroneous use of the in non-obligatory contexts which require a/an and Ø. Adjustments made to the denominator would result in a score of 33%. (Master, personal communication, September 23, 2005) TLU measures were used to determine results for the present study.

While TLU was applied to measure article accuracy, UOC was used to analyze article
usage. Master (1987) devised UOC as a complementary measure that can be used to assess a learner’s overuse or underuse of an article:

\[
\text{UOC} = \frac{\text{the total number of suppliances in both obligatory and non-obligatory contexts}}{\text{number of obligatory contexts}} \times 100
\]

Like TLU, suppliances in non-obligatory contexts is also taken into consideration in UOC, so the learner’s overall use of the articles can be inspected. When performance on article tasks matches that of a native speaker, TLU and UOC scores will have a value of one. Using these two measures in tandem leads to a better understanding of articles acquisition: TLU scores offer uninflated estimates of participants’ accuracy levels, and UOC scores reveal information regarding learners’ actual use or overuse of articles.

During the data analyses, TLU scores and UOC scores for the, a/an, and Ø were calculated for each participant, and each of the three proficiency level groups. The means of TLU scores for each group were compared to identify the accuracy order. The means of UOC scores for each group were also compared to examine overuse or underuse of articles. Percentage scores for correct responses were calculated for each type of task, and for each proficiency level.

In order to determine accuracy order and usage pattern and to identify learning difficulties for participants, a one-way analysis of variance (ANOVA and Scheffe Tests) was used to determine whether the mean scores for the three groups were significantly different. The two independent variables were Article (with three levels: the, a/an, and Ø) and Group (also with three levels: Advanced, Intermediate, and Lower-Intermediate). A dependent repeated-measures design was necessary because the same groups of participants were used to examine usage of each article type. All inferential statistics were performed using the Statistical Package for the Social Sciences (SPSS) 11.5. P values of .05 or lower were considered to be statistically significant.

4. Results and Discussion

Data collected from the tests were analyzed in an attempt to investigate the article accuracy order and usage pattern for Chinese English language learners. First, comparative analyses were conducted among and within the three groups to analyze the two different kinds of data elicited by the two measures: TLU and UOC. Second, percentage scores for correct responses were calculated for each semantic article type, and for each proficiency level. Comparative analysis was conducted in order to compare the three groups’ ability to use the articles the, a/an, and Ø in the [±SR, ±HK] contexts and in idioms. Third, data collected from the interviews were summarized for analysis.

4.1 Results and analysis of the test data

4.1.1 Comparative analysis of the mean TLU scores among the three groups

Comparative analysis of the mean TLU scores among the three groups was conducted first.
The second-year undergraduates from a normal school were identified as Group 1, and the second-year English minors and fourth-year English majors from a university were identified as Group 2 and Group 3 respectively.

Table 4-1. Results of an ANOVA analysis of the three groups’ TLU scores

<table>
<thead>
<tr>
<th>GROUP</th>
<th>A/AN</th>
<th>THE</th>
<th>ZERO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>50.8313</td>
<td>64.1637</td>
<td>24.5103</td>
</tr>
<tr>
<td>N</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>9.98593</td>
<td>6.75992</td>
<td>10.49368</td>
</tr>
<tr>
<td>Mean</td>
<td>62.4987</td>
<td>70.3237</td>
<td>34.0270</td>
</tr>
<tr>
<td>N</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>15.17994</td>
<td>9.61837</td>
<td>12.48923</td>
</tr>
<tr>
<td>Mean</td>
<td>82.2170</td>
<td>83.3340</td>
<td>67.4685</td>
</tr>
<tr>
<td>N</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>8.38448</td>
<td>7.97723</td>
<td>14.84246</td>
</tr>
<tr>
<td>Mean</td>
<td>63.0530</td>
<td>71.2663</td>
<td>38.8186</td>
</tr>
<tr>
<td>N</td>
<td>80</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>16.96814</td>
<td>11.06358</td>
<td>21.09335</td>
</tr>
</tbody>
</table>

Table 4-1 presents each group’s mean TLU score for each of the three articles. There is a positive relationship between English proficiency level and TLU score, which suggests that there is a positive relationship between article acquisition and learners’ English proficiency level. The first and the second group showed similar patterns for article usage. For Group 3, the TLU score for the indefinite article was higher than it was for the other two groups, and much like Group 1 and Group 2, Group 3’s TLU score for the zero article was lower than its TLU score for the other two articles. This suggests that the ability to use this article is acquired later than the ability to use other articles such as the and a/an.

It shows that articles follow the accuracy order the>a/an>Ø across the groups, which suggests that the use of the is more targetlike than a/an, which is more targetlike than Ø.

In order to test this hypothesis, a one-way ANOVA and a Scheffe follow-up test were conducted. Results from the ANOVA indicated that the interaction effects for article and group were all statistically significant at the .01 level (p=.000). A Scheffe post hoc test was also conducted. Results indicated that the mean difference between all possible pairs of groups were significant, p<.05. These findings suggest an accuracy order of the articles, the>a/an>Ø.

4.1.2 Comparative analysis of the mean UOC scores among the three groups

While TLU is used to measure article accuracy, UOC is used to measure article usage. Table 4-2 presents each group’s mean UOC score for each of the three articles. UOC scores indicate that a/an was overused at a rate of 117.78% for Group 1, and 120.83% for Group 2. For Group 3 a/an was a little bit underused (98.75%). The ANOVA analysis showed there
was significant difference between all possible pairs (p<.05).

**Table 4-2.** Results of an ANOVA analysis of the three groups’ UOC scores

<table>
<thead>
<tr>
<th>GROUP</th>
<th>A/AN Mean</th>
<th>THE Mean</th>
<th>ZERO Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>117.7783</td>
<td>117.7280</td>
<td>56.9043</td>
</tr>
<tr>
<td>N</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>22.71499</td>
<td>14.29233</td>
<td>21.67155</td>
</tr>
<tr>
<td>Mean</td>
<td>120.8340</td>
<td>110.9097</td>
<td>67.3810</td>
</tr>
<tr>
<td>2</td>
<td>22.61035</td>
<td>14.10383</td>
<td>21.19570</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>13.04495</td>
<td>10.26827</td>
<td>18.88478</td>
</tr>
<tr>
<td>Mean</td>
<td>98.7510</td>
<td>97.9550</td>
<td>102.8575</td>
</tr>
<tr>
<td>3</td>
<td>13.04495</td>
<td>10.26827</td>
<td>18.88478</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>114.1674</td>
<td>110.2279</td>
<td>72.3214</td>
</tr>
<tr>
<td>Mean</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>N</td>
<td>80</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>22.36062</td>
<td>15.26548</td>
<td>27.54431</td>
</tr>
</tbody>
</table>

Moreover, results presented in Table 4-2 indicate that overuse of the article the decreased as English proficiency level increased (117.72%; 110.91%; 97.96%). Conversely, overuse of Ø increased with English proficiency level (56.90%; 67.38%; 102.86%). These results suggest that Group 1 and Group 2 learners tended to overuse the and underuse Ø, whereas Group 3 was less likely to overuse the and more likely to overuse Ø. Use of a/an increased with the participants’ English proficiency level, and then decreased once English proficiency extended beyond a certain level. We argue that this finding is directly related to the type of task participants were required to complete. We left spaces for the missing articles which may have given participants the impression that it was necessary to use a/an or the, especially participants with lower-ability levels.

### 4.1.3 Comparative analysis among three groups for article usage in [±SR, ±HK] contexts

In this section, the percentage scores for correct responses for each semantic article type in each group are presented. The performance results for the test are shown in Table 4-3. We use Type 1, 2, 3, 4, 5 respectively to refer to the following types of tasks:

1. [+SR, −HK], (a, Ø): First mention nouns, whose referent is identifiable to the speaker but not to the listener.
2. [+SR, +HK], (the): referential definites.
3. [−SR, −HK], (a, Ø): nonreferentials.
4. [−SR, +HK], (the, a, Ø): generics.
5. Idioms, (the, a, Ø): idiomatic expressions and conventional uses.

Percentage of accurate responses were averaged for each setting and presented by proficiency level and by type of task. The data in the following table illustrate the differences in performance for the three proficiency level groups.
Table 4-3. Accuracy rates for article usage (Percentage Correct)

<table>
<thead>
<tr>
<th>Type 1</th>
<th>Type 2</th>
<th>Type 3</th>
<th>Type 4</th>
<th>Type 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>62.89</td>
<td>90.32</td>
<td>55.33</td>
<td>27.78</td>
</tr>
<tr>
<td>Group 2</td>
<td>75.33</td>
<td>90.95</td>
<td>62.67</td>
<td>41.11</td>
</tr>
<tr>
<td>Group 3</td>
<td>70.00</td>
<td>94.52</td>
<td>73.00</td>
<td>91.67</td>
</tr>
</tbody>
</table>

The data presented in Table 4-3 allow us to draw a number of conclusions. For all groups, Type 2 tasks yielded the highest accuracy rates. Even Group 1 produced correct response rate of 90.32%. The rate of accuracy reached 90.95% and 94.52% for Group 2 and Group 3 respectively. ANOVA analyses indicated that, for Type 2 tasks, the difference among the three groups’ accuracy rates was not significant (p=.089).

Nonreferential _a/an_ and _zero_ articles (Type 1 task) showed a zigzag pattern (scores first increased, then decreased) across the three groups. Students in Group 2 (intermediate-ability students) scored slightly higher than both Group 1 and Group 3 on Type 1 tasks. For Group 3, Type 1 tasks yielded the lowest scores. The differences observed among groups for performance on Type 1 tasks were statistically significant at the .01 level (p=.001).

Type 3 tasks, which required the use of both _a/an_ and the _zero_ article, were difficult for Group 1. Nonetheless, correct article usage in [−SR, −HK] environments increased as the participants’ English proficiency level increased. The positive relationship between English proficiency level and performance on Type 3 tasks was significant at the .05 level (p=.017).

Results for Type 4 tasks (generics) and Type 5 tasks (idioms) presented a very different picture. Scores for these tasks rose sharply with increases in the participants’ English proficiency level. These two tasks required skillful use of the articles _a/an_ or _the_ or _zero_. These tasks were very challenging for Group 1 and Group 2 participants. Differences among the three groups for performance on these tasks were significant at the .01 level (p=.000).

In order to verify the results from the one-way ANOVA, a Scheffe post hoc test was conducted to make multiple comparisons between the groups. Results from the Scheffe post hoc test were in line with results from the one-way ANOVA.

Task order for Group 1 (from the lowest level of accuracy to the highest level of accuracy) is as follows: idioms (Type 5), generics (Type 4), nonreferentials (Type 3), referential indefinites (Type 1), and referential definites (Type 2). For Group 2, the order is: generics (Type 4), idioms (Type 5), nonreferentials (Type 3), referential indefinites (Type 1), and referential definites (Type 2). These two groups exhibit the same difficulties for them: generics and idioms. For high-ability learners (Group 3) the order is: referential indefinites (Type 1), nonreferentials (Type 3), idioms (Type 5), generics (Type 4), and referential definites (Type 2). Acquisition curves were obtained by plotting the rate of accuracy from the article matrix.

4.2 Results of the interviews

To supplement the data gathered from the test, an interview was conducted immediately after the cloze test to obtain information regarding the rationale for participants’
responses for the test. Four students were randomly selected from each group to answer the interview questions.

The students from Group 1 indicated that they were not sure about the appropriate response for many of the blanks. For example, one blank in the first part of the cloze test appeared in front of the following phrase: ____ enormous house. Two of the participants asked the researcher what “enormous” meant. Once they knew the meaning, they realized which article they should use. Article choices for idioms were also difficult for them. Most of them did not know the meaning of idiomatic phrases such as “all of a sudden” or “from hand to mouth”. They assumed that an article should be used, so they chose either a or the.

In choosing articles for the phrase “in ____ 1960s”, some participants selected the zero article because, in Chinese, the phrase, zai liushi niandai (在六十年代) is used (this phrase does not contain an article). Here, we can see how language transfer influenced these participants’ response choices. Test items 13 (smoke ____ cigarettes) and 20 (clever with ____ money) were also difficult for them. For these items, the participants wondered whether or not the should be used when referring to a certain category. Results indicate that they performed poorly on these three items.

The participants from Group 2 indicated that they had the greatest difficulty with items 28 (All of ____ sudden), 29 (in ____ 1960s), 30 and 31 (____ hand to ____ mouth). They were not familiar with the usage rules for these items. Interestingly, one participant responded incorrectly to an item (Item 42) that he did not consider to be difficult. During his interview, he acknowledged that this was due to carelessness. He inserted the article the, without realizing that the phrase used (most animals have) was not intended to be a superlative.

The Group 3 students that were interviewed felt that the test was relatively easy. Results confirm that the test was not very difficult for them. Yet two of the interviewees from Group 3 chose to use the article a for item 49 (____ great courage and strength) ([+SR, −HK] contexts). Participants indicated that perhaps this was due to the fact that in advanced English learning, they assumed that many uncountable words might become countable under certain conditions (e.g., a pleasure, a surprise, etc.). These types of errors seemed to be a product of intralingual transfer. And the result also shows that the participants have difficulty judging the countability of noun phrases.

All three groups agreed that item 25 (____ copies of rare books) was difficult to answer. This item, a Type 3 task ([+SR, −HK] contexts), requires the use of the zero article, but most of the participants used the article the instead of the zero article. The learners overgeneralized the in of-phrase, regardless of [+SR, −HK] contexts or [+SR, +HK] contexts. Only two of the twelve participants who were interviewed answered this item correctly. Neither of these two participants were able to explain why they chose the zero article for item 25. When given more details regarding this kind of usage the other participants concluded that they answered this item incorrectly because they never paid attention to this kind of rule before. Some indicated that their teachers made little mention of these rules in class. Several of the interviewees from Group 3 indicated that these rules had been explained by teachers in the past but because of insufficient practice, they had been completely forgotten.
Data collected during interviews suggest that participants from Group 1 and 2 had the most difficulty with generics and idioms. Article choice in [+SR, −HK] contexts were difficult for all the three groups. These finding are congruent with the data analysis results that are presented above.

4.3 Discussion

4.3.1 What are accuracy and usage patterns of the English articles?
The TLU results reported above suggest that the accuracy order for all groups follows the pattern: the > a/an > Ø. While TLU scores help us identify accuracy order, UOC scores help us understand underlying acquisition processes. UOC scores help us identify overuse of articles in certain specific contexts. For Group 1 and Group 2, the and a/an were both overused. Both of these groups underused the zero article, which suggests that use of the zero article may be more difficult for them.

The UOC scores of Group 1 and 2 for a/an and the were more than 100%, showing that the intermediate learners overused these two articles. As shown in Table 4-2, overuse of the article a/an decreased as English proficiency level increased, whereas the accuracy rate increased gradually. TLU scores for a/an and the support these findings.

Low-level learners in Chaudron & Parker (1990) first overgeneralized Ø in a/an contexts. With increased proficiency, the use of Ø regressed and the use of a/an increased. In the present study, there was a positive relationship between English proficiency level and use of Ø (the UOC score for Ø was 56.91%, 67.38%, and 102.86% for Group 1, Group 2, and Group 3 respectively). Findings from these two studies do not contradict each other, but rather confirm a pattern of Ø usage, that is in line with Huebner’s (1983) notions of flooding and trickling: Lower-level learners with a native language that does not include an article system initially overgeneralize Ø, and then reduce the use of Ø when testing hypotheses by trying the other articles. As they continue to become more proficient, they gradually increase the frequency with which they use Ø (this trend was apparent in the present study). The most common source of errors made by the study participants in Ø contexts was overgeneralization of a/an and the. This was true for participants at the lower-intermediate and intermediate levels of proficiency.

Master (1987) also identified a flooding-then-trickling pattern of Ø usage: The [-ART] group (English learners with a native language that does not include articles) started with Ø-flooding, and then shifted to a Ø-trickling stage when they realized that English phrases must be preceded by a specifier, such as the. So, they tested hypotheses by flooding NPs with the, while simultaneously trickling the use of Ø. A dramatic change from Ø-flooding to the-flooding occurred at the Low proficiency level. When the higher level learners recognized that Ø could also be a specifier, they began to increase the frequency with which they used Ø.

The patterns for the Ø acquisition processes discussed above are consistent with the patterns identified in the present study. The lower-intermediate and intermediate students (Group 1 and Group 2) may hypothesize that most of the blanks need to be filled with certain specifiers (such as a/an and the), whereas the advanced learners (Group
3) may have grasped the fact that Ø is also commonly used in the English language. Hence, Group 1 and Group 2 may underuse Ø while Group 3 begins to increase usage of Ø.

To sum up, the two measures, TLU and UOC, reveal useful information regarding the acquisition of English articles by Chinese learners: TLU scores allow us to identify accuracy order, the > alan > Ø for intermediate and advanced level Chinese learners; and UOC scores reveals the usage pattern: lower-intermediate and the intermediate learners overgeneralize the and alan, and underuse Ø, while advanced learners overuse Ø.

4.3.2 In what semantic contexts are there areas of difficulty underlying article choice for Chinese learners?

Analyses of the data indicated that the five different tasks present different levels of difficulty for L2 learners. Language learners appear to develop the ability to do well on each task one by one, not all at the same time. Results presented in Table 4-3 indicate that all three groups received high scores in Type 2 [+SR, +HK] contexts. This suggests that language learners’ solid command of this type of use. Important to note, however, is that the [+SR, +HK] task only required use of the article the. For Group 1 and Group 2, the was overgeneralized in other contexts. These participants were likely to use the for the following items: item 24 (a length of 12 meters), item 25 (Ø copies of rare books), and item 48 (Ø bands of fifteen to twenty). They tended to substitute the for alan or Ø in of-phrase structures. According to Lu (2001), Chinese learners are likely to overgeneralize the in of-phrase structures, even in [+SR, −HK] contexts. Participants used the instead of alan or Ø in item 42 (Ø most animals), item 51 (I once read a story), and item 54 (attack Ø human hunters). This suggests that they have difficulty distinguishing [+SR, +HK] from [+SR, −HK] contexts.

For Type 1 tasks ([+SR, −HK] contexts), which required the use of alan and Ø, the scores first increased and then decreased (62.89%, 75.33%, 70%). This trend was seen on items such as item 20 (clever with Ø money), item 36 (Ø large, catlike tracks), item 43 (have a favorite food), etc. For Type 3 tasks (−SR, −HK] contexts), which also required the use of alan and Ø, the accuracy rate increased with English proficiency level. The scores of the advanced learners on Task 1 and Task 3 were much lower than the other three tasks (70% and 73%), which suggests these two items were more difficult for Group 3. We conclude that the Chinese learners of English, particularly the advanced learners, have difficulty choosing appropriate articles in [−HK] contexts. These findings are consistent with those of Brown (1973), Parrish (1987), Thomas (1989), Yoon (1993), and Butler (2002).

The data also suggest that Type 4 tasks [−SR, +HK] (generics) and Type 5 tasks (idioms) required the most sophisticated article usage. Both tasks required skillful placement of alan, the, or Ø. The generics tasks included items such as the jaguar, the wild pig, and Ø wild pigs. The idioms tasks included items such as all of a sudden, in the 1960s, and from Ø hand to Ø mouth. These tasks were especially difficult for the intermediate learners. These findings provide empirical support for Liu and Gleason’s (2002) conclusions about the difficulties associated with acquiring the
ability to become proficient with what they termed “the cultural use” of articles. These conclusions seem reasonable: language learners are rarely exposed to generics, and distinctions between the literal meaning and actual meaning of idioms must be learned as a whole (Butler, 2002).

Results from the subject interviews suggest that the intermediate Chinese learners, had the most difficulty with article usage for generics ( [−SR, +HK] ) and idioms. They also seem to have difficulty distinguishing [+SR, +HK] contexts from [+SR, −HK] contexts. Article choices in [−HK] contexts (referential indefinites and nonreferentials), were more problematic for the advanced learners. This kind of usage was based on fixed grammar rules and was usually introduced at the very beginning of English learning. Yet due to the fact that article errors rarely lead to miscomprehension, learners devote little effort to learning the system (Master, 1995). The learners have not really acquired the article system, even in the advanced period of learning.

4.4 Pedagogical implications
This article has offered some insight into the general development patterns of the English article system for Chinese English language learners at intermediate and advanced levels. For these learners, accuracy rates for article usage increase with English language proficiency level. This implies that, for intermediate and advanced English language learners, there is a positive relationship between these two factors.

For the intermediate and advanced learners, the accuracy order is the> a/an> Ø. This shows that the mastery of Ø may be most difficult for non-native learners. The overuse of Ø persists at the highest level more than overuse of the other articles (Master, 1997). Intermediate learners’ overuse of a/an and the suggests that it is difficult for them to choose appropriate articles in [+SR] contexts (specific referent). Only after grasping the articles’ regular accuracy and usage patterns, can the instructor teach with more efficiency.

There are many factors that have the potential to make article choice difficult for Chinese English language learners. “Cultural use” of articles (in generics and idioms) is one of the difficulties for intermediate learners to master. For advanced learners, mastering the ability to use articles in [−HK] contexts (not known to the hearer) is very difficult. It may be worthwhile to develop strategies that can be used to overcome these obstacles for Chinese English language learners.

Our findings lead us to believe that grammar instruction should always be intertwined with instruction that allows students to gain insight into the cultural aspects of rules for language usage. In our study, intermediate level participants received very low article usage scores for generics and idioms. Performance on these tasks is likely related to participants’ knowledge of the cultural aspects of the English language. As Master (1995) points out, “shared knowledge” is the most difficult aspect of language that causes the greatest number of errors in language usage. It requires a constant assessment on the part of the writer of what he or she presumes the reader to know. Instruction for the cultural aspects of language can be delivered in many forms, including various kinds of after-class activities.
5. Conclusion

The findings presented in this study may contribute to our understanding of the nature of interlanguage representations, and provide some useful information that is relevant to article pedagogy. Findings such as these help us understand accuracy and usage patterns in the article acquisition process, and identify difficulties that Chinese English language learners have when choosing articles. Much more needs to be accomplished in this research area. For example, it is still unclear whether Chinese English language learners at different proficiency levels exhibit the same accuracy order. Future investigations should consider using oral production tasks in order to compare results with those that are obtained when written cloze tests are used. According to Cai & Wu (2006), besides referent specificity and hearer’s knowledge of noun phrases, countability of nouns, influence of L1 and different task types are also features which have significant effects on the usage of the English articles by Chinese learners of English. Only after we have gained a thorough understanding of the various aspects of L2 article acquisition can we hope to build a complete profile of article acquisition for Chinese English language learners.

Note

1. Bickerton (1981) claimed that articles have different interpretations that are captured in terms of two binary features, namely, whether the article and associated NP refer to a specific entity (±SR), and whether the article and associated NP are already known from the previous discourse or from context, to the person who is listening to or reading the sentence (±HK).

References


An Investigation of English Articles’ Acquisition by Chinese Learners of English


Appendix A  Test

I. Choose the correct article (a, an, the, Ø) in the following sentences.

1. There is  1  orange in that bowl.
2. Carlos is  2  student at our university.
3. What is  3  sex of your baby? It’s  4  boy!
4. I always drink  5  water with my meals.
5. What is  6  diameter of  7  moon?
6. Once there were many trees here. Now,  8  trees are gone.
7. I would like  9  cup of coffee, please.
8. My father earns $25,000  11  year.
9.  12  evening sky was really beautiful.
10. A man knocked on my door.  13  man was bleeding.
11. People who smoke  14  cigarettes often get lung cancer.
12.  15  air in this city is not very clean.
13. Einstein was  16  man of great intelligence.
14. In this family,  17  first child inherits everything.
15. This book did not sell well even though  18  author was a famous writer.
16. She owns  19  enormous house in Pasadena.
17.  20  fool though he was, he was clever with  20  money.
18. That was  21  worst storm of 1985.
19. I keep sending  22  messages to him.
20.  23  restaurant in which we ate was quite expensive.
21. This room has  24  length of 12 meters.
22.  25  copies of rare books should always be preserved.
23. I ordered a bottle of wine, but  26  bottle of wine was too cold.
24 Dr. Engelberg, _27_ physician to Marilyn Monroe, would not comment on her death.
25 All of _28_ sudden, he woke up from his nightmare.
26 In _29_ 1960s, there were lots of protests against the Vietnam War.
27 He has been thrown out of work, and his family is now living _30_ hand to _31_ mouth.

II. Choose the correct article (a, an, the, Ø) in the following paragraph.

_The_ Jaguar and _the_ Wild Pig

When hunters visit _32_ southwestern part of _33_ United States, they often find _34_ large, catlike tracks along _35_ ground. These tracks are made by _36_ spotted jaguar, _37_ greatest hunter of all _38_ North American animals and _39_ largest member of _40_ cat family on _41_ American continent. _42_ most animals have _42_ favorite food. _43_ favorite food of _45_ jaguar is _46_ wild pig. _47_ wild pigs move in _48_ bands of fifteen to twenty. They have _49_ great courage and strength in _50_ group. I once read _51_ story about _52_ courage and strength of these wild pigs. _53_ story pointed out that these pigs sometimes even attack _54_ human hunters.

Appendix B  The Interview

Which one do you think is most difficult?
Please elaborate the possible reasons.

(Copy editing: Devon Williams)

(…continued from p. 62)


(Copy editing: Joshua Lee)