Analysis of the Process of Article Selection in English by Japanese EFL Learners: Using Path and Decision-Tree Analysis

Toshiaki TAKAHASHI  
Yamaguchi University

Abstract

It is well known that Japanese EFL learners have great difficulty in the accurate use of English articles. The present study investigates this difficulty by looking at the process of article selection in English by Japanese native speakers. We use path analysis to compare the hypothesized processes of choosing English articles and the actual processes employed by Japanese EFL learners. The results of path analysis indicate that the process used by Japanese EFL learners is close to that proposed by Master (1990). These results, as well as those of decision-tree analysis, indicate that the criteria General and SR_HK do not have a direct link with each other and each independently affects the choice of definite article. When the SR_HK criterion is met, the definite article is used. In addition, when the General criterion is not met, the definite article is used. As regards the choice of indefinite articles (the indefinite article and the zero article), both the General criterion and the Countable criterion need to be satisfied, while the SR_HK is not relevant to their choice. Finally, the implications of the present study are discussed.

1. Introduction

Japanese learners have great difficulty in the accurate use of English articles. In order to best teach the correct use of English articles, it is critical to understand how Japanese EFL learners determine which article is to be used; that is, how they select English articles and what steps they take in making such decisions.

If we knew the criteria used to select English articles, and through what kind of process articles are selected, this knowledge would aid teachers in giving direction on the proper use of English articles. It would also enable students to alter their decision-making processes for choosing English articles (i.e., what criteria Japanese EFL students use when making the distinction between definite and indefinite, what decision is to be made first, etc.).

As Swan (1994) pointed out, pedagogical grammar needs to satisfy the criterion of simplicity. If rules are too complicated or there are too many of them, they are not appropriate from the viewpoint of pedagogical grammar. In the same way, it would seem practical and
efficient to teach the selection of English articles through a few simple criteria and simple steps to apply these criteria.

There are studies, although not many, that provide simple steps for the selection of English articles and some criteria for the proper selection of English articles (for more details on these models, see section 2.2). Since such models outline a process for the selection of English articles, we term them “process models” for the selection of English articles. Because all these process models present steps for making choices in the form of flowcharts, they would be helpful when teaching English articles.

Although there are some overlaps between these process models, there are still significant differences between them. For example, one process model begins with the step of determining the countability of a noun phrase (NP), while other models begin with the step of determining whether the reference is general. Similarly, one model uses a criterion of whether the reference is unique in the context, while other models use a criterion of whether the reference is specifically known.

We do not know whether Japanese EFL learners follow one of the existing process models, nor do we know whether they follow a different model. Since none of the process models have been empirically tested, it is important to analyze processes by which Japanese learners select English articles based on these models in order to determine what criteria are used and what steps are taken when they select English articles.

In the present study, analysis was performed using elements of the use of English articles known to be difficult for Japanese ESL learners, so that we could understand which aspects of English article selection cause the most difficulty for Japanese EFL learners, and thus determine pedagogical implications for the instruction of English articles.

2. Background

2.1 Criteria likely to be used for selecting English articles

When selecting English articles, learners must make various decisions. For example, learners cannot determine whether the zero article can be used if they do not know whether the NP is countable.

Previous research has indicated that there are several factors (or criteria) that are considered to be in play when the learner selects English articles. (See below; due to space limitations, please refer to Takahashi (2006) and Takahashi (2013) for details.) However, all factors (or criteria) may not necessarily be linked with the choice of English article. For example, all of the criteria from (f) to (h) below are considered relevant to the use of the definite article, and although criterion (h) seems to be the most relevant (Takahashi, 2006), this does not mean that Japanese learners actually use this criterion. In addition, as we discuss in the next section, different process models use different criteria for making decisions concerning the use of English articles. Thus, it is
important to examine which below-mentioned criteria the Japanese EFL learners will use for their decisions:

(a) NP is countable (Yoon, 1993).
(b) The referent is definite (Komiya, 1998).
(c) The reference is general or specific (Master, 1994).
(d) The noun refers to an individuated entity (Wierzbicka, 1988).
(e) The noun refers to a concrete (or an abstract) entity (Young, 1996).
(f) The referent is specifically known to the hearer (Heubner, 1979).
(g) The referent is specifically understood (identifiable) (cf. Yule, 1995).
(h) The referent is context unique (Zehler & Brewer, 1982).

In the above criteria, criterion (f) refers to whether the referent is both specific (=+SR) (Specific Referent)) and known to the hearer ( [+HK] (Heubner, 1979; Parrish, 1987; Thomas, 1989)). In the following text, the abbreviation SR_HK is used for this criterion in order to save space in tables and figures. Criterion (g) refers to whether the referent is specific or whether the hearer can identify which entity the speaker is referring to. In the following text, the abbreviation Specific may be used. In the same way, criterion (h) is abbreviated as Unique hereafter. As regards criterion (b), definiteness is very difficult to define. Therefore, the present study defines definiteness as either being specifically known (criterion (f)), specifically understood (criterion (g)), or context unique (criterion (h)).

2.2 Process models

In the present study, a process model of English articles is defined as a model of the process of selecting English articles. Although there is relatively little research focusing on process models for English articles, they have significant potential for pedagogical application. Teachers may find it useful to know the criteria that their students use and the steps that they follow to select English articles.

The criterion of being a process model in this study refers to whether a model describes the criteria to use and what steps to follow to select English articles in the form of a flowchart or a decision tree. In addition, to increase the validity of the present study, process models for English article selection have been taken from peer-reviewed international journal articles, or books or chapters of books published by highly regarded international publishers.
The process models found were those proposed by (1) Swan (1984), (2) Lindstromberg (1986), (3) Master (1990), and (4) Robinson (2010), which are described in more detail below. Since the present study deals neither with plural nouns nor the generic use of articles, some details not relevant to the present study are omitted from the schematic representations of these models. Although some models (i.e., Lindstromberg (1986) and Robinson (2010)) incorporate a separate step to deal with the use of the definite article for anaphoric reference, other models do not. For the convenience of comparison, these steps were omitted from schematic representations and were excluded from analysis.

Swan (1984) proposed the following process of selection of English articles.

Step 1: Determine whether the reference is general (whether the speaker is generally referring to objects or things). If yes, the zero article is used. If no, go to Step 2.

Step 2: Determine whether the referent is definite. If yes, the definite article is used. If no, go to Step 3.

Step 3: Determine whether the NP is a countable noun. If yes, the indefinite article is used. If no, the zero article is used.

For the sake of simplicity, the present study focuses only on the use of articles with singular nouns. (Therefore, the step relating to the use of articles with plural nouns is omitted from the description above.) It should be noted that Swan’s model is not specific about which of the two criteria (i.e., Specific, SR_HK) is supposed to exist under this model when determining whether the definite article is used.

Next is a model proposed by Lindstromberg (1986). This is a very complex process model (the most complex model of the four). Lindstromberg’s model, called “The Decision Sheet,” is two pages long. Its description has been greatly simplified for our purposes, whilst still retaining the essential features of the model.
The process models found were those proposed by (1) Swan (1984), (2) Lindstromberg (1986), (3) Master (1990), and (4) Robinson (2010), which are described in more detail below. Since the present study deals neither with plural nouns nor the generic use of articles, some details not relevant to the present study are omitted from the schematic representations of these models. Although some models (i.e., Lindstromberg (1986) and Robinson (2010)) incorporate a separate step to deal with the use of the definite article for anaphoric reference, other models do not. For the convenience of comparison, these steps were omitted from schematic representations and were excluded from analysis.

Swan (1984) proposed the following process of selection of English articles.

**Step 1:** Determine whether the reference is general (whether the speaker is generally referring to objects or things). If yes, the zero article is used. If no, go to Step 2.

**Step 2:** Determine whether the referent is definite. If yes, the definite article is used. If no, go to Step 3.

**Step 3:** Determine whether the NP is a countable noun. If yes, the indefinite article is used. If no, the zero article is used.

For the sake of simplicity, the present study focuses only on the use of articles with singular nouns. (Therefore, the step relating to the use of articles with plural nouns is omitted from the description above.) It should be noted that Swan’s model is not specific about which of the two criteria (i.e., Specific, SR_HK) is supposed to exist under this model when determining whether the definite article is used.

Next is a model proposed by Lindstromberg (1986). This is a very complex process model (the most complex model of the four). Lindstromberg’s model, called “The Decision Sheet,” is two pages long. Its description has been greatly simplified for our purposes, whilst still retaining the essential features of the model.

**Step 1:** Determine whether the NP is a countable noun. If no, the zero article is used. If yes, go to Step 2.

**Step 2:** Determine whether the referent is unique in the context. If yes, the definite article is used. If no, go to Step 3.

**Step 3:** Determine whether the reference is general (whether the speaker is generally referring to objects or things). If yes, the indefinite article (non-specific reference) is used. If no, the indefinite article (specific reference) is used.

The next model was proposed by Master (1990). This is a very simple process model (the simplest of the four).

**Step 1:** Determine whether the referent is “identified” or “classified.” If “identified,” the definite article is used. If not, go to Step 2. What Master (1990) meant by “identified” is not clear, but it seems very close in meaning to either Specific or SR_HK, while “classified”...
seems to mean General (the opposite of “identified”).

Step 2: Determine whether the NP is a countable noun. If it is, the indefinite article is used. If not, the zero article is used.

For the sake of simplicity, the present study focuses only on the use of articles with singular nouns. (Thus, a step relating to the use of articles with plural nouns is omitted from the steps described above.)

Finally, Robinson (2010) proposed the following model, which is very similar to the model proposed by Swan (1984).

Step 1: Determine whether the reference is general (whether the NP is the subject of a generalization). If it is, the indefinite article is used. If not, go to Step 2.

Step 2: Determine whether the referent is definite. If it is, the definite article is used. If not, go to Step 3.

Step 3: Determine whether the NP is a countable noun. If yes, the zero article is used. If no, the indefinite article is used.

3. Method

3.1 Participants

Twenty university students (including five graduate students) (8 male, 12 female) participated in the present study. The average TOEIC score of the participants was 617.75, with a standard deviation of 72.39. The participants were pursuing English-related majors, in their third or fourth year of the undergraduate program or in the first year of the graduate program. All of the participants were Japanese.

As described in the procedure section below, the participants were individually tested in a quiet room. They were asked to answer 21 questions on English articles and answer a questionnaire relating to the 8 criteria affecting the choice of English articles (as discussed in 2.1).
The present study recruited a relatively higher level of EFL students as participants. The reasons for this were as follows: (1) responding to the questionnaire requires metalinguistic awareness of the choice of English articles; (2) all participants must be able to understand the meaning of the criteria.

Due to the limited availability of qualified participants, the number of participants is not large. However, as described later in section 3.5, the sample size is considered to be sufficient for data analysis (e.g., path analysis).

3.2 Materials
3.2.1 Multiple-choice questions on English articles

As described in the procedure section, each participant was individually tested in order to ensure that the meaning of the criteria was clear. For practical reasons, the maximum test time for each individual was limited to less than an hour. Therefore, it was necessary to limit the number of questions (i.e., the number of multiple-choice questions on English articles, and the number of questionnaire questions).

In addition, one of the aims of the present study was to investigate difficulties that EFL learners have with English articles. The present study used the 21 questions on English articles that were answered with less than 70% accuracy in Takahashi (2008), such as the following example (See Appendix for the 21 questions used in the present study). Please note that three questions on idiomatic article usage and three questions concerning the use of the definite article for anaphoric reference were excluded (cf. 2.2)).

(Q 1) 実生活で、これまでパンダを見たことがありますか？
Have you seen (a/the/x) panda before in real life? (x indicates the zero article)

Since the questions used in Takahashi (2008) contained a wide range of English article usage (i.e., eight types of use of the definite article, such as immediate situation, etc.) with a balanced number of countable/uncountable nouns and abstract/concrete nouns, and they were conducted among a total of 123 high-school students, it seems reasonable to assume that the 21 questions that were answered with less than 70% accuracy in Takahashi (2008) reflect difficulties encountered by Japanese EFL learners.

A Japanese translation was used to eliminate the possibility of student inability to answer correctly owing to a lack of vocabulary. Since the Japanese language does not have an equivalent of English articles, it was assumed that the addition of a Japanese translation would not affect the choice of English articles.

Contextualization of the article insertion test would have been ideal but this requires longer texts and longer processing times. In addition, longer texts without Japanese translation would hinder participants’ understanding of the text, and would affect the reliability of the test itself.
3.2.2 Questionnaire survey

The questionnaire survey contained questions concerning the eight factors considered to
influence the choice of English articles (i.e., the eight criteria described above in section 2.1).

After answering each of the 21 multiple-choice questions on English articles, the
participants were asked to judge the countability of an underlined noun (e.g., panda).

The participants were then asked to make the following judgments about the underlined
noun in the question on English articles (e.g., panda). They were instructed to mark with a circle
as many choices as they considered appropriate.

(a) The noun is used in a general meaning.
(b) The noun is used in an abstract sense.
(c) The noun refers to an individuated entity.
(d) The noun refers to a concrete entity.

Finally, the participants were asked to make the following judgments about the underlined
referring expression (e.g., panda). They were instructed to mark with a circle as many choices as
appropriate.

(a) What the speaker is referring to by the referring expression is specifically known to
the hearer.
(b) What the speaker is referring to by the referring expression is understood by the
hearer.
(c) What the speaker is referring to by the referring expression is unique in the context.

3.3 Procedure

Participants were individually tested in a quiet room. At the beginning of the experiment, all
the criteria were briefly explained in Japanese by the researcher in order to make sure all the
criteria and instructions were clear to them. Each participant was instructed to answer
multiple-choice questions on English articles and a questionnaire about them as described in
sections 3.2.1 and 3.2.2. Only one question was answered at a time. After answering each question,
participants were asked to complete a questionnaire about the eight factors described above.

They were instructed to answer the English-article test and questionnaire at their own pace.
On average, participants took approximately 40–50 minutes to finish both the test and
questionnaire.

3.4 Method of statistical analysis

Path analysis was used to determine what criteria Japanese EFL learners use and what
sequence they follow to select English articles. IBM SPSS AMOS 23 was used for model building
and path analysis.
The hypotheses were created based on the four process models of English articles discussed in section 2.2. Since the model by Swan (1984) is not specific on the criteria used to determine the definiteness of the NP, making it difficult to determine whether the model used criterion (f) (=SR_HK) or criterion (g) (=Specific), two types of hypotheses were tested (Swan 1 and Swan 2). In the same way, the model by Master (1990) is not specific about the criteria used to determine the definiteness of the NP, and it was difficult to determine whether the model used criterion (f) (=SR_HK) or criterion (g) (=Specific). Therefore, two types of hypotheses were tested (Master 1 and Master 2).

3.5 Assessment of model fit

The goodness of fit for each model was determined using a Goodness of Fit Index (GFI) and an Adjusted Goodness of Fit Index (AGFI), Normed Fit Index (NFI), Comparative Fit Index (CFI), Root Mean Square Error Approximation (RMSEA), and Akaike’s Information Criterion (AIC).

Values for GFI, AGFI, NFI, and CFI range from 0 to 1 and are derived from the comparison of a hypothesized model with the independent model. A value equal to or greater than 0.95 indicates an acceptable fit to the data for GFI, and a value equal to or greater than 0.90 indicates

<table>
<thead>
<tr>
<th>Models</th>
<th>Criteria for Definiteness</th>
<th>GFI</th>
<th>AGFI</th>
<th>NFI</th>
<th>CFI</th>
<th>RMSEA</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>The</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swan 1</td>
<td>SR_HK</td>
<td>0.997</td>
<td>0.973</td>
<td>0.932</td>
<td>1.000</td>
<td>0.000</td>
<td>18.969</td>
</tr>
<tr>
<td>Swan 2</td>
<td>Specific</td>
<td>0.998</td>
<td>0.982</td>
<td>0.989</td>
<td>1.000</td>
<td>0.000</td>
<td>18.638</td>
</tr>
<tr>
<td>Master 1</td>
<td>Unique</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.000</td>
<td>18.013</td>
</tr>
<tr>
<td>Master 2</td>
<td>Specific</td>
<td>0.996</td>
<td>0.980</td>
<td>0.900</td>
<td>1.000</td>
<td>0.000</td>
<td>17.416</td>
</tr>
<tr>
<td>Robinson</td>
<td>SR_HK and Unique</td>
<td>0.999</td>
<td>0.990</td>
<td>0.991</td>
<td>1.000</td>
<td>0.000</td>
<td>16.695</td>
</tr>
<tr>
<td>An</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swan 1</td>
<td>SR_HK</td>
<td>1.000</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
<td>0.000</td>
<td>18.019</td>
</tr>
<tr>
<td>Swan 2</td>
<td>Specific</td>
<td>0.997</td>
<td>0.975</td>
<td>0.993</td>
<td>1.000</td>
<td>0.000</td>
<td>18.497</td>
</tr>
<tr>
<td>Master 1</td>
<td>Unique</td>
<td>0.999</td>
<td>0.986</td>
<td>0.996</td>
<td>1.000</td>
<td>0.000</td>
<td>18.276</td>
</tr>
<tr>
<td>Master 2</td>
<td>Specific</td>
<td>0.999</td>
<td>0.950</td>
<td>0.965</td>
<td>0.999</td>
<td>0.013</td>
<td>18.033</td>
</tr>
<tr>
<td>Robinson</td>
<td>SR_HK and Unique</td>
<td>0.997</td>
<td>0.974</td>
<td>0.989</td>
<td>1.000</td>
<td>0.000</td>
<td>16.857</td>
</tr>
<tr>
<td>Zero</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swan 1</td>
<td>SR_HK</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.000</td>
<td>18.002</td>
</tr>
<tr>
<td>Swan 2</td>
<td>Specific</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.000</td>
<td>18.002</td>
</tr>
<tr>
<td>Master 1</td>
<td>Unique</td>
<td>0.996</td>
<td>0.982</td>
<td>0.984</td>
<td>1.000</td>
<td>0.000</td>
<td>17.023</td>
</tr>
<tr>
<td>Master 2</td>
<td>Specific</td>
<td>0.979</td>
<td>0.904</td>
<td>0.930</td>
<td>0.978</td>
<td>0.023</td>
<td>21.982</td>
</tr>
<tr>
<td>Robinson</td>
<td>SR_HK and Unique</td>
<td>1.000</td>
<td>0.998</td>
<td>0.999</td>
<td>1.000</td>
<td>0.000</td>
<td>20.105</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Criterion</th>
<th>≥0.950</th>
<th>≥0.900</th>
<th>≥0.900</th>
<th>≥0.900</th>
<th>≤0.050</th>
<th>The smaller the better</th>
</tr>
</thead>
</table>

Table 1 Results of Assessment of Model Fit in Path Analysis
an acceptable fit for AGFI, NFI, and CFI. By convention, there is a good fit if RMSEA is less than or equal to 0.05.

AIC is regarded as an information theory goodness of fit measure and is used to compare different models. The models that generate the lowest values are regarded as optimal. The absolute AIC value is irrelevant and only the AIC value of one model relative to the AIC value of another model is meaningful. When other indices indicate that a model is acceptable, the model with the lowest AIC value is considered to be the model with the best fit.

Although the number of students was not large ($N = 20$), the number of sets of the questionnaire, which consisted of eight questions about the eight criteria (see section 2.1), was five for questions employing the indefinite article (an Questions), nine for questions employing the definite article (the Questions), and seven for the questions employing the zero article (zero Questions) (see Appendix). Thus, the sample size for path analysis was 100 ($= 20 \times 5$) for an Questions, 180 ($= 20 \times 9$) for the Questions, and 140 ($= 20 \times 7$) for zero Questions, thus meeting the minimum recommended sample size for path analysis (i.e., 100).

4. Results and Discussion

4.1 Results of path analysis

As described above, when other indices indicate that a model is acceptable, the model with the lowest AIC value is considered to be the model with the best fit. Therefore, Master’s (1990) model using SR_HK as the criterion for definiteness (Master 1) attained an adequate fit as indicated by GFI = 0.996, AGFI = 0.980, NFI = 0.900, CFI = 1.000, RMSEA = 0.000, and AIC = 17.416 for questions employing the definite article; GFI = 0.990, AGFI = 0.950, NFI = 0.965, CFI = 0.999, RMSEA = 0.013, and AIC = 18.033 for questions employing the indefinite article; and GFI = 0.996, AGFI = 0.982, NFI = 0.984, CFI = 1.000, RMSEA = 0.000, and AIC = 17.023 for questions employing the zero article.

Figure 5. Path Solution for Questions Employing the Definite Article
Presented in Figure 5 is the path model (Master 1) obtained for questions employing the definite article. As Figure 5 indicates, the path solution under this model does not suppose no interaction between the two variables (General and SR_HK). This supports the model by Master (1990) in that the process for selecting the definite article and that for selecting indefinite articles (i.e., the indefinite article and the zero article) are independent from each other. Thus, learners seem to select the definite article when they think that the referent is specifically known to the hearer (SR_HK) irrespective of whether the referent is general or not. In the same way, learners appear to use the definite article when they think that the reference is not general, irrespective of whether the referent is specifically known to the hearer (SR_HK).

The path solution also indicates that participants in the present study used SR_HK (not Unique) as the criterion for the definiteness of an NP. This may show that SR_HK does not contribute to the accuracy in the use of the definite article (=Score). If learners used Unique as the criterion for definiteness, the criterion for definiteness may have had a stronger link to the accuracy. Or it may be that since the questions on English articles contained many uncountable or abstract nouns, known to be difficult for Japanese EFL learners, learners may have had difficulty applying SR_HK (or any other criteria) when determining the definiteness of the NP. Thus, the General criterion may have been the only option left to determine whether the definite article should be used.
The obtained path solution also supports the prediction that countability is not linked to the choice of the definite article as indicated by a correlational value of 0.000 ($p = 0.997$) (between the variable Countable (the countability of an NP) and the variable Score (the accuracy in the use of the definite article)), as predicted by Master (1990).

Presented in Figure 6 is a path model (Master 1) obtained for questions employing the indefinite article. The path solution indicates that the countability of an NP plays a major role in the choice of the indefinite article. As Master (1990) predicted, the criterion for the definiteness of an NP is not strongly related to the selection of the indefinite article.

Finally, Figure 7 shows the path model (Master 1) obtained for the questions employing the zero article. The path solution indicates that the criterion of whether the reference is general (=General) was the most important factor. As in the case of the indefinite article, the criterion for the definiteness of an NP is not strongly related to the selection of the zero article.

4.2 Insignificant paths

Only a limited number of paths were statistically significant. In the case of Figure 5, the only significant path was between General and Score ($−0.238, p < 0.001$). In the case of Figure 6, significant paths were found between Countable and Score ($0.633, p < 0.001$) and between General and Score ($0.181, p < 0.05$). In the case of Figure 7, significant paths were found between General and Score ($0.409, p < 0.001$), between Countable and Score ($0.394, p < 0.001$), and between SR_HK and Score ($0.158, p < 0.05$). Although no other paths were significant, the present study did not remove the insignificant paths, as the final model (Master 1) had already obtained a good fit. Toyoda (2003) said that it is not always necessary to remove insignificant paths if a model has obtained an acceptable fit. Besides, if some modification is made (such as the deletion of a statistically insignificant path from a model), it would be necessary to verify the modified model by using a different group of students (Tabachnick & Fidell, 2007), which is obviously not practical. Furthermore, the lack of a significant relationship between variables can also provide invaluable information.

However, given that there is no significant link between General and Countable, there is a possibility that all three criteria are independent of each other (at least for the Japanese EFL learners participating in the present study). This remains to be tested in future research.

4.3 Ad hoc analysis using decision-tree analysis

Decision-tree analysis was used to identify major factors affecting the choice of English articles. Decision-tree analysis is a popular classification method that uses a tree-like graph or model to show how a previous decision or step affects the decisions or consequences that follow it. R software (R version 3.2.0; rattle package, http://www.R-project.org/) was used to create a decision tree.
Since the results of path analysis had identified the three criteria (SR_HK, General, and Countable) as the main factors affecting the choice of English articles, decision-tree analysis used these three criteria to create a decision tree.

As Figure 8 shows, the primary factor affecting the selection of the definite article is whether the criterion General is accurate (i.e., General = 1). In other words, those students who correctly judged that the reference was not general tended to accurately select the definite article. The classification result performed with decision-tree analysis shows that among the 180 data, there are 72 data points (40%) on the left node and 108 data points (60%) on the right. As Figure 8 shows, the average accuracy rate was 70% (0.7) when this criterion was judged accurately.

As illustrated by Figure 9, the ability to judge the countability of an NP is vital for the accurate selection of the indefinite article. If students can judge the countability of an NP accurately, and if they can correctly judge whether the reference is general, they are likely to appropriately select the indefinite article (with an accuracy rate of 81%). Thus, the result of decision-tree analysis seems compatible with the result of path analysis.
As Figure 10 illustrates, similar patterns of results were obtained for questions employing the zero article (zero Questions): Both the generality criterion (General) and the noun countability criterion (Countable) were critical for accurate use of the zero article. The accuracy rate was 83% when both criteria were judged accurately.

5. Conclusions

The results of the present study indicate that Master’s (1990) model can best account for the observed data compared with the other models. As Master’s (1990) model predicted, there is no direct link between SR_HK and General and they seem to be independent from each other. When the SR_HK criterion was satisfied, the definite article was used, irrespective of whether the reference was general. In addition, when the reference was not general, the definite article was used irrespective of whether the SR_HK criterion was met. Thus, not only the SR_HK criterion but also the General criterion seemed to play an important role when determining whether the definite article should be used. The results also indicate that the accurate use of the definite article is not related to the Countable criterion. The results of path analysis and decision-tree analysis show that both the general criterion and the countable criterion need to be met for the use of the indefinite articles (i.e., the indefinite article and the zero article). Although both the General criterion and the Countable criterion must be met, the former played a major role in the selection of the zero article, whereas the latter played a major role in the selection of the indefinite article. One important pedagogical implication of the current study is that much more emphasis should be placed on the role of the generality criterion when teaching which article should be used. This would be especially true when the noun is abstract or uncountable and criteria for the definiteness, such as SR_HK, are difficult to apply.

In most cases, participants in the present study seemed to determine whether the definite article is used based on whether the referent is specifically known to the hearer (SR_HK), rather than whether the referent is context unique. It remains to be seen whether this is the reason for a lower causal link between criteria (SR_HK) and accuracy, or whether all criteria for definiteness, including SR_HK and Unique, are difficult to apply when the noun is abstract or uncountable.
As Figure 10 illustrates, similar patterns of results were obtained for questions employing the zero article (zero Questions): Both the generality criterion (General) and the noun countability criterion (Countable) were critical for accurate use of the zero article. The accuracy rate was 83% when both criteria were judged accurately.

5. Conclusions
The results of the present study indicate that Master’s (1990) model can best account for the observed data compared with the other models. As Master’s (1990) model predicted, there is no direct link between SR_HK and General and they seem to be independent from each other. When the SR_HK criterion was satisfied, the definite article was used, irrespective of whether the reference was general. In addition, when the reference was not general, the definite article was used irrespective of whether the SR_HK criterion was met. Thus, not only the SR_HK criterion but also the General criterion seemed to play an important role when determining whether the definite article should be used. The results also indicate that the accurate use of the definite article is not related to the Countable criterion. The results of path analysis and decision-tree analysis show that both the general criterion and the countable criterion need to be met for the use of the indefinite articles (i.e., the indefinite article and the zero article). Although both the General criterion and the Countable criterion must be met, the former played a major role in the selection of the zero article, whereas the latter played a major role in the selection of the indefinite article. One important pedagogical implication of the current study is that much more emphasis should be placed on the role of the generality criterion when teaching which article should be used. This would be especially true when the noun is abstract or uncountable and criteria for the definiteness, such as SR_HK, are difficult to apply.

In most cases, participants in the present study seemed to determine whether the definite article is used based on whether the referent is specifically known to the hearer (SR_HK), rather than whether the referent is context unique. It remains to be seen whether this is the reason for a lower causal link between criteria (SR_HK) and accuracy, or whether all criteria for definiteness, including SR_HK and Unique, are difficult to apply when the noun is abstract or uncountable.

Aknowledgement

This work was supported by JSPS KAKENHI Grant Number 15K02680. The author would like to thank Enago (www.enago.jp) for the English language review.

References


Appendix

For the sake of convenience, the questions are grouped according to 3 question types and have been renumbered sequentially. The Japanese translation has been omitted from the Appendix due to space restriction.

An Questions
(Q 1) Have you seen (a / the / x ) panda before in real life?
(Q 2) These days no one is worried about (an / the / x ) attack from outer space.
(Q 3) I have (a / the / x ) poor ability at expressing my ideas.
(Q 4) You’re (a / the / x ) real help to someone like me.
(Q 5) English is (a / the / x ) subject that he is not at all interested in.

The Questions
(Q 6) Ichiro was busy taking photos with (a / the / x ) new camera that he got for his birthday.
(Q 7) When you decorate (a / the / x ) ceiling, please use white paint.
(Q 8) The most powerful person in the United States is (a / the / x ) president.
(Q 9) A: Where’s (a / the / x ) coffee?  B: It's on the table.
(Q10) I like to study (a / the / x ) history of Egypt.
(Q11) I really have no idea. This is (an / the / x ) only information I have.
(Q12) Everybody knows that (a / the / x ) current housing shortage is mainly caused by the high price of land.
(Q13) (An / The / x ) art of paper folding has a long history in Japan
(Q14) Personally I think (a / the / x ) salary is not enough.

Zero Questions
(Q15) I don't like (a / the / x ) rice cooked in rice cookers.
(Q16) For cooking I prefer (a / the / x ) dark chocolate, which is mildly bitter.
(Q17) Please contact us on 01234 234321 for (an / the / x ) information on our products.
(Q18) One of the major problems today is (an / the / x ) air pollution.
(Q19) At present, I am doing (a / the / x ) research on Japanese culture.
(Q20) A: What are you doing at college? B: I am studying (a / the / x ) modern Italian art.
(Q21) Do you drink (a / the / x ) milk for breakfast?