Why do Hong Kong Cantonese ESL learners choose a certain English article for use?

Chan, Alice Yin Wa

Published in:
The Asian Journal of Applied Linguistics

Published: 01/03/2017

Document Version:
Final Published version, also known as Publisher’s PDF, Publisher’s Final version or Version of Record

License:
Unspecified

Publication record in CityU Scholars:
Go to record

Publication details:

Citing this paper
Please note that where the full-text provided on CityU Scholars is the Post-print version (also known as Accepted Author Manuscript, Peer-reviewed or Author Final version), it may differ from the Final Published version. When citing, ensure that you check and use the publisher's definitive version for pagination and other details.

General rights
Copyright for the publications made accessible via the CityU Scholars portal is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights. Users may not further distribute the material or use it for any profit-making activity or commercial gain.

Publisher permission
Permission for previously published items are in accordance with publisher’s copyright policies sourced from the SHERPA RoMEO database. Links to full text versions (either Published or Post-print) are only available if corresponding publishers allow open access.

Take down policy
Contact lbscholars@cityu.edu.hk if you believe that this document breaches copyright and provide us with details. We will remove access to the work immediately and investigate your claim.
Why do Hong Kong Cantonese ESL learners choose a certain English article for use?

Alice Yin Wa Chan  
Department of English, City University of Hong Kong

This article reports on a study which investigated the reasons underlying Cantonese ESL learners’ choice of an English article and the hypotheses they made in their selection. 33 Cantonese ESL learners majoring in English participated in the study, which included two cloze passages with all the articles deleted. Participants were required to provide a suitable article for each blank based on the context and give the reason behind each selection. Results showed that although English majors performed well when selecting an article for use, they did not have enough knowledge about the English article system. Certain inappropriate hypotheses were made about the effects of linguistic contexts on reference interpretation. Misconceptions about generic reference and the concept of definiteness were apparent. ESL teachers are suggested to design suitable formal and explicit teaching programmes to eradicate learners’ misconceptions about the English article system.

Keywords: Second language acquisition; English article system; learning of English articles; Chinese learners; Hong Kong

Introduction

English uses different articles (a/an, the, and ZERO) for generic, definite and indefinite interpretations. There is no one-to-one fixed relation between an article choice and the reference type expressed. A/an with singular count nouns can be used to express indefinite or generic references (e.g. “A boy is coming. A dinosaur is a prehistoric animal”). The with singular count nouns can be used to express definite or generic references (e.g. “The book is expensive. The dinosaur is extinct.”), but with plural count nouns and mass nouns the is only used for definite reference (e.g. “The tables are expensive.”). ZERO with plural count nouns and mass nouns can be used to express indefinite or generic references (e.g. “There is water everywhere. Water is essential to our lives.”).

English articles are often found to cause problems for second language learners (Huebner, 1983), especially those whose native languages lack articles, like Chinese (Austin, Pongpairoj, & Trenkic, 2015; Ionin, Zubizarreta, & Maldonado, 2008; Luk & Shirai, 2009). Various explanations have been given to account for L2 article errors, particularly the complex ways in which meaning is mapped onto form (Young, 1996), learners’ judgment of noun countability (Yoon, 1993), and L1 transfer (Luk & Shirai, 2009).

In probing the reasons underlying learners’ article choices, Butler (2002) found that learners made incorrect hypotheses showing sensitivity to the wrong contexts, postulating inappropriate word-article collocational rules that restricted certain words to collocate with a certain article; and relying excessively on structural cues. Chan (2016)
also observes that even advanced learners may not realize the importance of identifiability for definiteness and often confuse specificity, where the speaker “intends to refer to exactly one individual” (Ko, Ionin, & Wexler, 2010, p. 219), with definiteness, which presupposes existence and uniqueness and is related to the identifiability of referents in discourse (Ko et al., 2010; Trenkic, 2008).

Although previous research provides insightful findings about learner hypotheses and misconceptions, the focus has been exclusively on incorrect article choices. There seems to be no systematic analysis of reasons for correct article choices. It is unknown, therefore, whether learners’ incorrect generalizations or misconceptions about the English articles system will lead to correct article choices. This is an important area of inquiry, because such learner problems escape the notice of teachers (and learners) and therefore are not remedied. The present study aimed to fill this gap.

**Research questions**
The present study examined the reasons behind Hong Kong Cantonese ESL learners’ article selection to understand:
1. why ESL learners choose a certain English article for use;
2. what hypothesis they make in their selection of an English article for use in a certain context;
3. whether correct article choices will result from inappropriate learner hypotheses and how.

**Participants**
Participants consisted of 33 Cantonese native speakers (4 males and 29 females) aged 19 to 23 who were majoring in English at a local university, Seven had learnt English for 14-16 years, nineteen for 17-19 years, and the rest for 20 years or more. 25 had received a C or above in the Hong Kong Advanced Level Use of English (HKALE UE) exam¹, 7.5 or above in the International English Language Testing System (IELTS) test, or 5 or above in the Hong Kong Diploma of Secondary education (HKDSE)², and the rest scored D in HKALE or 6.5-7 in IELTS.

**Data collection**
Following Butler (2002), two short passages (of 204 and 229 words) were chosen from two university English workbooks. All the articles in the texts were deleted, creating 50 blanks for each of which participants were asked to insert the article they thought most suitable given the context provided by the passage. Immediately after completing each passage, they were asked to provide, verbally, the reason for each article choice. In their verbal reporting many participants used a mixed-code of Cantonese and English which, for the purposes of this paper is reported in English. A research assistant implemented the cloze passage task, recorded the reasons using a smart-phone, and transcribed the recordings. Care was taken to ensure that the items were all clear, appropriate and idiomatic, and all the three articles of *a/an* (6 blanks), *the* (23 blanks), and *ZERO* (21 blanks), and all the three types of definite (22 blanks), indefinite (12 blanks) and generic (16 blanks) references were included. Some unevenness of distribution is inevitable due to the use of authentic texts.

The passages were piloted with two native speakers and three Cantonese ESL learners with similar linguistic backgrounds. They all found the passages suitable for the target group and the task requirements clear and reasonable.
Data analysis

Performance on the cloze passage task
The participants’ performance was analysed by counting the targetlike selection of each participant on each article and each reference type. Targetlike selections are those corresponding to the original text and non-targetlike are deviations even if they were viable alternatives to the originals. Friedman procedure tests (which allow for a population which may not be normally distributed) using SPSS 16 for Windows 7 were run to compare participants’ performance on different articles and references in both texts.

Classification of reasons for article choices
Butler (2002) classified her participants’ reasons into specific and non-specific, but because specificity is a major focus of this paper, the corresponding terms of precise and imprecise are used. Butler’s (2002) specific (precise) classification of: referentiality, countability, non-generalizable hypothesis and others, are reasons based on rules of grammar and discourse or previous experience. A preliminary analysis of the findings reported here showed the use of phonological concerns and previous encounters (seen or used before) as the basis of article selection, so these two sub-categories have been added. A further level of sub-classification has been added to referentiality (dividing it into correct and incorrect reference identified) and countability (subdividing into correct detection and misdetection of countability). Butler’s (2002) sub-classifications for non-specific reasons (here called imprecise) consist of plausible choice, elimination and no clue, which are responses where participants cannot identify any reasons and these have been retained. The full set of classifications is shown in Figure 1.

Unlike Butler (2002), who only sub-classified precise reasons for non-targetlike article choices, precise reasons for both targetlike and non-targetlike article choices were sub-classified in the current study. During classification, attention was given to certain patterns of article selection, referentiality judgment, or other concerns noted from the reasons, including the correlation between the choice of an article and the interpretation of reference types. Generalizations to account for learners’ errors were then drawn with regard to their consideration of different linguistic aspects.

Results

Performance on language task
The participants’ performance in the cloze passages was reasonably good, with high targetlike selection rates for all articles. Their performance on noun phrases with the as the target article was the best, with a targetlike selection rate of 82.5% (mean= 18.97), followed by ZERO at 77.9% (mean= 16.36). Participants’ performance on noun phrases with a/an as the target article was the worst, with a targetlike selection rate of only 63.2% (mean= 3.79). A Friedman procedure test showed that there was a significant difference between a/an vs. the vs. ZERO ($X^2= 24.382, p= 0.000$).
Figure 1. Reason classifications
As for the three types of reference, as can be seen in Table 1, the targetlike selection rate for noun phrases with target generic reference was 69.1% (mean = 11.06), that for definite reference was 87.2% (mean = 19.18), and that for indefinite reference was 74% (mean = 8.88). A Friedman procedure test showed that the differences on generic vs. definite vs. indefinite references were significant ($X^2 = 28.962$, $p = 0.000$).

Table 1. Participants’ performance on different articles and different references

<table>
<thead>
<tr>
<th></th>
<th>a/an (n=6)</th>
<th>the (n=23)</th>
<th>ZERO (n=21)</th>
<th>Generic (n=16)</th>
<th>Definite (n=22)</th>
<th>Indefinite (n=12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean of Participants’ Targetlike Selection (SD= 1.24)</td>
<td>3.79 (SD= 1.64)</td>
<td>18.97 (SD= 1.93)</td>
<td>16.36 (SD= 1.47)</td>
<td>11.06 (SD= 1.47)</td>
<td>19.18 (SD= 1.51)</td>
<td>8.88 (SD= 1.59)</td>
</tr>
<tr>
<td>Percentage of Targetlike Section (Mean/n)</td>
<td>63.2%</td>
<td>82.5%</td>
<td>77.9%</td>
<td>69.1%</td>
<td>87.2%</td>
<td>74%</td>
</tr>
</tbody>
</table>

Reasons for article selection

Reason distribution
For target-like article use, 91.8% of the reasons given were precise and 8.2% were imprecise. Of the imprecise reasons, 60.4% were plausible choices, 17.9% elimination and 21.7% no clue. Of the precise reasons, 61.7% were about correct referentiality, 14.4% about incorrect referentiality, 8.3% about correct detection of countability, 0% about incorrect detection of countability, 4% about structures seen or heard before, 0.6% phonological concerns, and 11.1% idiosyncratic or non-generalizable hypothesis.

For non-targetlike article use, 91.4% of the reasons given were precise and 8.6% were imprecise. Of the imprecise reasons, 74.2% were plausible choices, 6.5% elimination and 19.4% no clue. Of the precise reasons, 5.8% were about correct referentiality, 66.5% about incorrect referentiality. 7.6% about correct detection of countability, 4% about structures seen or heard before, 0.9% about phonological concerns, and 12.2% based on idiosyncratic or non-generalizable hypothesis. Unlike Yoon’s (1993) observation, misdetection of noun countability was not serious, accounting for only 3% of article misuse.

Targetlike article choices led by incorrect judgments
All the targetlike article choices led by incorrect judgements were associated with incorrect referentiality judgments. Of these (171 cases), 36.8% were associated with correct selection of ZERO for target generic noun phrases (see examples 1b-2b in Table 2) but incorrect judgment of those phrases as having indefinite, non-specific or even specific/definite reference (reasons 1a-2a), showing that the distinction between indefinite and generic references was often undetected. The use of the for generic reference escaped many learners, as a noun phrase with the was almost always taken as definite: 11.1% of targetlike article choices were associated with correct selection of the for target generic singular noun phrases (example 3b) but incorrect judgment of those phrases as definite (reason 3a). Incorrect judgment of noun phrases as having generic reference instead of the target indefinite reference (reason 4a) was very common,
leading to 33.9% of targetlike article choices. There were many targetlike choices of definite the (45 cases), especially in noun phrases with a post-modifier (example 5b), in which the information for identifying the referent was wrongly detected (reason 5a) despite correct referentiality judgment.

Table 2. Targetlike article choices led by misidentification of reference types/misdetection of information to identify referents

<table>
<thead>
<tr>
<th>Target Reference</th>
<th>Reference Identified by Participants</th>
<th>Reason</th>
<th>Example Sentence/Phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td>generic</td>
<td>indefinite</td>
<td>(1a)</td>
<td>Domestic life is domestic life, it is not a specific domestic life, so we don’t need ‘the’ (participant 25).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1b)</td>
<td>the importance of consumer discrimination in / domestic life</td>
</tr>
<tr>
<td>generic</td>
<td>specific/definite</td>
<td>(2a)</td>
<td>I didn’t put any articles, because, it is specified. It is that domestic life. (participant 6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2b)</td>
<td>the importance of consumer discrimination in / domestic life</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3a)</td>
<td>I added ‘the’, because it is referring to this average, the average. (participant 28).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3b)</td>
<td>anxieties that beset the average human being.</td>
</tr>
<tr>
<td>indefinite</td>
<td>generic</td>
<td>(4a)</td>
<td>I didn’t put any articles because this is a general term (participant 9).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4b)</td>
<td>Shake off excess water and handle …. with / damp hands</td>
</tr>
<tr>
<td>definite</td>
<td>definite, wrong detection of information to identify referent</td>
<td>(5a)</td>
<td>For this blank, I used ‘the’, because type and quality are talking about products (participant 28).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(5b)</td>
<td>Judgment is based on the type and quality of materials</td>
</tr>
</tbody>
</table>

Key: Underlining indicates the article selected. ZERO articles are represented by /.

Relationships between reference and article choices
Although the same article can be used for showing different references, participants’ choice of a certain article, targetlike or non-targetlike, was often predominantly the result of their (mis)identification of only one reference. Definite reference, for example, was the only reference type which triggered the (mis)use of the. No participants chose the because the associated noun phrase was thought to be generic. For a/an, only 7 instances were caused by participants’ (mis)interpretation of a noun phrase being generic, whereas all the rest (164 cases) were related to indefinite reference. Once generic reference was ruled out as the interpretation of a noun phrase, participants would think that an article, the or a/an should be used (see reasons 6a-7a in Table 3), because generic reference was thought to be only used with ZERO (reason 8a).
Table 3. Relationships between reference and article choices

<table>
<thead>
<tr>
<th>Misconception</th>
<th>Reason</th>
<th>Example Sentence/Phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-generic used with an article</td>
<td>(6a) It’s referring to actual production something like not really general (participant 11).</td>
<td>(6b) the actual production of what they buy</td>
</tr>
<tr>
<td></td>
<td>(7a) I chose ‘an’ for this, because … it is talking about an example, not a general situation (participant 28).</td>
<td>(7b) an inescapable part of life</td>
</tr>
<tr>
<td>generic (misidentified) always used with ZERO</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(8a) There are places where we don’t need an article, for example, a general concept (participant 3).</td>
<td>(8b) in order to persuade / people to ….</td>
</tr>
</tbody>
</table>

Key: Underlining indicates the article selected. ZERO articles are represented by /

Misconceptions/confusion about generic reference
Because generic reference was often thought to be associated with ZERO, when generic reference was (mis)identified, ZERO was used as a default choice irrespective of the number of the head noun. Even with singular count nouns, ZERO was selected, resulting in ungrammatical expressions (see Table 4, example 9b). Participants’ choice of lexis in explaining their article selection for (thought to be) generic noun phrases is also worth noticing. The term generic was scarcely used (only 6 times by one single participant). Instead, general or all (or the equivalent in Chinese) were used 200 times for the former (reason 10a) and 96 times for the latter (reason 11a). As a result, generic noun phrases with all were interpreted (correctly or incorrectly) as generic simply because of the pre-determiner (reason 12a).

Table 4. Misconceptions/confusion about generic reference

<table>
<thead>
<tr>
<th>Misconception</th>
<th>Reason</th>
<th>Example Sentence/Phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td>default ZERO for generic</td>
<td>(9a) This has no article because … the word average is already very general (participant 5).</td>
<td>(9b) anxieties that beset / (the) average human being</td>
</tr>
<tr>
<td>“general” referring to generic reference</td>
<td>(10a) This does not need an article because it is talking about general advertisers (participant 10).</td>
<td>(10b) / (The) advertiser’s concern</td>
</tr>
<tr>
<td>“all” referring to generic</td>
<td>(11a) This should have no articles, because it is talking about … all contact lenses (participant 15).</td>
<td>(11b) You will find using / soft contact lenses to be ….</td>
</tr>
<tr>
<td>generic induced by “all”</td>
<td>(12a) Because ‘all’ is already a generic term, we don’t need the definite article (participant 5).</td>
<td>(12b) All / lenses have individual markings</td>
</tr>
</tbody>
</table>

Key: Underlining indicates the article selected. Brackets indicate the target article where different. ZERO articles are represented by /
Possible confusion between (in)definiteness and (non-)specificity

When describing their article choices for (thought-to-be) definite noun phrases, many participants were inclined to use the word *specific* rather than *definite*, as if definite reference was equivalent to specific reference (Table 5, reasons 13a-14a). By the same token, when describing their article choices for (thought-to-be) indefinite noun phrases they were inclined to say that the noun phrases were not specific rather than indefinite (reasons 15a-16a). A simple word count showed that the terms *specific* and *not specific* were used 134 and 71 times respectively in participants’ reasons, whereas *definite* and *not definite* were used only 11 times and once, respectively.

Table 5: Use of specific and not-specific for definite and indefinite references

<table>
<thead>
<tr>
<th>Lexis Choice</th>
<th>Reason</th>
<th>Example Sentence/Phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td>use of specific for definite</td>
<td>(13a) I chose ‘the’, because this passage is talking about something specific (participant 18).</td>
<td>(13b) <em>The</em> importance of consumer discrimination in domestic life</td>
</tr>
<tr>
<td>use of specific for thought-to-be definite</td>
<td>(14a) Because at the back it is talking about a particular product. It is a very specific thing (participant 16).</td>
<td>(14b) to persuade people to buy the <em>a</em> particular product</td>
</tr>
<tr>
<td>use of not specific for indefinite</td>
<td>(15a) I think it's not really specific like the concept of …… it's introducing something but it's not really specific (participant 26).</td>
<td>(15b) <em>an</em> inescapable part of life in today’s society</td>
</tr>
<tr>
<td>use of not specific for thought-to-be indefinite</td>
<td>(16a) It is not a specific domestic life, so we don’t need ‘the’</td>
<td>(16b) <em>The</em> importance of consumer discrimination in <em>domestic life</em></td>
</tr>
</tbody>
</table>

Key: Underlining indicates the article selected. Brackets indicate the target article where different. ZERO articles are represented by `/`

Effects of contexts on article selection or referentiality judgment

The immediate linguistic context of a noun phrase often had some effects on article selection. Some participants thought that the presence of a post-modifier would specify or limit noun reference, so *the* should be used with a noun phrase with a post-modifier (Table 6, reasons 17a-18a). While the hypothesis that the presence of a post-modifier was indication of definite reference was correct in most cases, as many post-modifiers were indeed combined with cataphoric reference to uniquely identify a specific subset (example 17b), there were also post-modifiers which did not help to uniquely identify a specific subset, yet the hypothesis was still applied, leading to non-targetlike article choices (example 18b).

The presence of a pre-modifier also led participants to using *the* as the default article. Two significant examples were *actual* in “the actual production of what they buy” (example 19b) and *particular* in “a particular product” (example 20b). Many participants chose *the* for both because the presence of the pre-modifiers *actual* and *particular* was thought to be specifying production and product, respectively (reasons 19a-20a).
Table 6: Article selection or referentiality judgment led by the presence of modifiers

<table>
<thead>
<tr>
<th>Effect of Contexts</th>
<th>Reason</th>
<th>Example Sentence/Phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td>targetlike article selection led by post-modifier</td>
<td>(17a) I used ‘the’, because there is ‘of’ at the back. Normally when we see the word ‘of’, we will use ‘the’ (participant 11).</td>
<td>(17b) the process of discrimination.</td>
</tr>
<tr>
<td>non-targetlike article selection misled by post-modifier</td>
<td>(18a) I put ‘the’, because it is talking about life, about the life in today’s society (participant 1).</td>
<td>(18b) an inescapable part of the (life in today’s society)</td>
</tr>
<tr>
<td>targetlike article selection led by pre-modifier</td>
<td>(19a) The actual production, because it is specifically talking about actual, not any production… It is actual, so very specific (participant 7).</td>
<td>(19b) Most people have little knowledge of the actual production of what they buy.</td>
</tr>
<tr>
<td>non-targetlike article selection misled by pre-modifier</td>
<td>(20a) Particular product should be ‘the’ particular, because particular is specifying something (participant 21).</td>
<td>(20b) … in order to persuade people to buy the (a) particular product</td>
</tr>
</tbody>
</table>

Key: Underlining indicates the article selected. Brackets indicate the target article where different. ZERO articles are represented by /

Other linguistic contexts which were used as the basis of referentiality judgment and/or article selection included the words more and all, which were sometimes thought to disallow a preceding article but sometimes thought to necessitate a default the. Words ending in -ness were regarded as a signal for the. The default article choice for the first word of a sentence was deemed to be the, that for an instruction-like sentence was the, and the article which should be used with the word first should also be the. Subject complements were regarded by many participants as information to identify noun referents, thus necessitating the use of the.

Discussion

Reasons for article selection
Cantonese ESL learners base their article decisions on different factors, including phonological environments, noun countability, and previous linguistic encounters, but referentiality (correctly or incorrectly interpreted) is the most prevalent reason. This shows that learners are, to a certain extent, aware of the core functions of English articles for showing reference and the importance of reference interpretation in article selection, yet in situations where other factors such as phonological environments and noun countability are salient enough, for example when a noun begins with a vowel or when it is uncountable or plural, such factors will be called upon for a quick decision and probably for confirmation. Their use of reference interpretation for article selection is not consistent and robust, resulting in occasional selections based on ungrounded hypotheses.
Incorrect hypotheses in article selection

Learners have some misconceptions about the functions and semantics of English articles and form incorrect hypotheses about article use. Some incorrect hypotheses show learners’ sensitivity to the wrong linguistic contexts, postulating certain co-occurrence restrictions for the choice or use/non-use of articles with other noun phrase elements and justifying their article use by such collocational rules. A prevailing misconception is the assumption that pre-modifiers function to uniquely identify a specific subset and indicate definite reference like many post-modifiers usually do. This shows learners’ confusion between identifiability and modification as a result of their possible non-distinction between specificity and definiteness as demonstrated in their use of the term specific for definite. As discussed earlier, specificity is distinct from definiteness. A noun phrase is specific if it intends to refer to exactly one individual and there exists a property which the speaker considers noteworthy in the context (Ionin, 2003, 2006). Therefore, noun phrases with definite reference may be specific (e.g. “The teacher who taught me last year left”) or non-specific (e.g. “I am looking for the hat that will suit me best”, Greenbaum & Nelson, 2009, p. 62), and those with indefinite reference may also be specific (e.g. “I met an American”) or non-specific (e.g. “I want a car”). However, many ESL learners often equate specificity with definiteness because of their failure to appreciate the importance of identifiability for definiteness and tend to regard specific noun phrases such as a boy in “I saw a boy” as definite (Chan, 2016).

For noun phrases with a classifying pre-modifier (e.g. a demographic transition), such a tendency is aggravated by the presence of the pre-modifier (i.e. demographic), the classificatory modification of which is presumably regarded as information rendering the noun phrases identifiable.

Other inaccurate hypotheses are manifested in learners’ co-occurrence restrictions disallowing/specifying certain articles (e.g. the) to co-occur with certain other noun phrase elements (e.g. all, first) based on their previous linguistic encounters, and/or their (partial) knowledge about article use. One significant exemplar is their confusion about generic reference. As is well-known, generic reference refers to a whole class of things. However, students often interpret it as referring to all entities of a group irrespective of whether the group is a sub-group or a class. Therefore, when encountering the pre-determiner all, learners tend to assume a default use with ZERO, on the incorrect assumption that the pre-determiner already serves to delimit the reference interpretation of the noun phrase (i.e. generic) and thus precludes an article.

The findings also reveal incorrect hypotheses induced by sentence elements other than those in the noun phrase itself, such as a subject complement appearing in the same sentence as a subject noun phrase whose reference interpretation is at issue. Many learners assign definite reference to a subject noun phrase (e.g. The essence of advertising in “The essence of advertising is persuasion”) on the basis of the equivalence relation between the subject and an identifying attribute (e.g. persuasion in “The essence of advertising is persuasion”). This shows that while learners are aware that reference interpretation is discoursal and may call for information outside of the immediate noun phrase, they are unaware of the distinction between identifiability (i.e. definiteness) and identity (i.e. equivalence) and that the equivalence relation between an attribute and its carrier does not contribute to the identifiability of the referent of the carrier or delimit an article choice.

Incorrect hypotheses manifested in learners’ one-to-one mapping between an article and a reference type are also prevalent, showing their partial knowledge about the English article system. ZERO is often regarded as the default (or even only) article choice for generic reference, and the and a/an are assumed to be used only for definite
and indefinite reference, respectively. Therefore, whenever the is used, a noun phrase is considered definite, whereas whenever a/an is used, it is considered non-specific. Such a simplified mapping between meaning and form may be the result of the technical names of the articles (the being called the definite article and a/an the indefinite article), which, according to Trenkic (2008), are “bona fide lexical meanings” that second language learners “attribute to article forms” (p. 10). However, it may also be the results of the more widespread use of ZERO but more restricted use of the or a/an for generic reference (Chan, 2016).

**Effects of incorrect hypotheses**
The effects of learners’ incorrect hypotheses/inaccurate referential judgments can be detrimental, leading to not just incorrect article uses but also occasional correct uses. Although it is not clear how serious the adverse consequences of correct article uses led by incorrect hypotheses/referential judgments are, such incorrect hypotheses may be unnoticed and even reinforced by the correct uses and ingrained in learners’ minds. The associated correct article uses are often unstable and vulnerable, as a different context or a (slight) change to the context may necessitate the use of a different article, but learners who hold incorrect hypotheses/referentiality judgments may not understand or be aware of such subtle differences. These explain why learners’ article use is sometimes targetlike (e.g. “Most people have little knowledge of the actual production of what they buy”, cf. 19b) but sometimes non-targetlike (e.g. “in order to persuade people to buy the particular product …”, cf. 20b)): Learners who have made incorrect hypotheses about the function of a pre-modifier may be correctly led to use the for the former but incorrectly led to use the same article for the latter.

**Implications**
The results have important implications. Theoretically, they add to our understanding of ESL learners’ explicit and implicit knowledge in the acquisition of English articles and of the phenomenon of noticing. It has been asserted that second language learning starts with explicit knowledge (DeKeyser, 2001, 2007), which “exists as declarative facts that can be stated” (Ellis, 2009, p. 13). Through practice, explicit knowledge will become implicit knowledge, which forms the basis of automatized performance (DeKeyser, 2001, 2007). However, our findings show that in the use of English articles, fluent performance may not be grounded on implicit knowledge, and given the very small number of members in the English article system, targetlike use does not necessarily reflect implicit knowledge or the result of continued practice after the acquisition of explicit knowledge. As is well-known, correctly used language forms (e.g. irregular past forms) do not categorically manifest learners’ mastery of those items but can be a reflection of their developmental stages (e.g. irregular past forms and base forms treated as separate vocabulary items). Although it is premature to claim that certain interlanguage stages exist in ESL learners’ article acquisition, it is evident that in this area of English learning, error analysis alone is not adequate, as learners’ correct article choices in certain contexts may be more indicative of their interlanguages. More in-depth analyses probing learners’ thinking processes like reason elicitation are needed to ascertain such interlanguages.

As for noticing, the results show that certain features in English article use are less readily noticed than others. One significant example is the importance of identifiability for definiteness and the related subtle linguistic difference between specificity and
definiteness. It may be argued that learners’ use of specific for definite is simply a random choice of lexis with the term specific used in a non-linguistic, everyday sense meaning more or less like unique (Chan, 2016), but there is reason to believe that the distinction between specificity and definiteness is less transparent and more difficult to notice, thus resulting in the observed fluctuating patterns.

Pedagogically, our results highlight the importance of explicit teaching in this area. ESL teachers should design explicit teaching programmes targeting learners’ explicit knowledge of the English article system, so that learners will be given a chance to verbalize the relevant concepts. Their use of English articles are then more likely to be led by their linguistic knowledge instead of ungrounded intuitions (Chan, 2016). When tackling learners’ problems, teachers should be conscious of not just incorrect uses but also the shadowy problems underlying correct uses. They need to probe into learners’ thinking processes and understand the root of their problems instead of merely focusing on learner performance, which may be deceptive. Learners’ misconceptions should be eradicated and relevant linguistic knowledge should be explicitly taught. It is advisable to set aside time for learners to voice the reasons underlying their article choices by making reference to reference representations and interpretations, or to analyse the reference of each article in an authentic text, substitute another article for a target article and discuss the possible changes to the interpretation of the sentence/text after substitution.

Limitations
Reporting on the reasons behind each article choice was tiring. Although only two passages were used, many participants spent over an hour explaining their choices. Some encountered difficulties in giving explanations due to their lack of confidence or insufficient metalanguage. The fatigue caused by the tediousness of the task and the frustration arising from their difficulties in verbalizing their concepts may have hindered their reporting process and resulted in unclear/unjustified answers.

Restricting possible correct answers to three alternatives (as is inevitable with English articles) may have encouraged guessing. The controlled task of completing blanks may also have masked learners’ ability to use English articles in a free and authentic context. Other learning tasks, such as grammaticality judgment and/or free conversations, may be more authentic in obtaining a more comprehensive picture.

Finally, no native English speakers were invited to provide baseline data. It is possible that native speakers would also show variability in their article selection. Although this would not impinge on the reliability of the results and the insights obtained it would have made an interesting additional comparison.

Conclusions
The results of this study investigating Cantonese ESL learners’ article use and their metalinguistic knowledge have important implications for these and similar learners. Referentiality is their most dominant concern for choosing an article and outweighs all other concerns. Although many learners can use English articles correctly, their choices may not necessarily be led by reasonable linguistic generalizations. Instead, many choices may be based on their ungrounded intuitions or misconceptions about the English article system, so even when an article is correctly chosen, learners may not be able to correctly interpret the intended reference of a noun phrase they are reading.
The learning of English articles is, thus, not just about the correct use of an article in a certain linguistic context, but also about the representation (as a speaker/writer) and interpretation (as a listener/reader) of intended reference. Targetlike article selection might hinder learners’ awareness of their errors and misconceptions, so awareness-raising activities targeting explicit knowledge about the functions and semantics of English articles will be useful in enhancing advanced learners’ acquisition of English articles and eradicating their misconceptions.

Acknowledgments
The work described in this article was supported by the Hong Kong Research Grants Council (Project Number: CityU 11400614). The support of the council is acknowledged. I would also like to thank all the respondents of the survey for their participation, my research assistants for their administrative help, and the consultants in the City University Statistical Consulting Unit for their expert advice on statistical analyses.

Notes
1. The Hong Kong Advanced Level Use of English (UE) examination, now no longer in use, tested students’ ability to understand and use English at a level that was required for tertiary education and/or for future employment (Hong Kong Examinations and Assessment Authority, 2013). UE Grade E or above is regarded as equivalent to a pass in the GCE A level examinations (Hong Kong Examinations and Assessment Authority, n.d.).
2. The Hong Kong Diploma of Secondary Education examination measures the attainment of students after six years of secondary education and is now the only public examination in Hong Kong. 5** is the highest grade that students can attain for a certain subject, followed by 5* and 5.

About the author
Alice Y.W. Chan is an associate professor in the Department of English, City University of Hong Kong. Her research interests include second language acquisition, English grammar, English phonetics and phonology, and lexicography.

References


