



Use and Misuse of the English “the”: A Case of Hong Kong Cantonese ESL Learners

Alice Yin Wa Chan

City University of Hong Kong, Hong Kong, China

This paper discusses Hong Kong Cantonese ESL learners’ judgment of the use of the English definite article. 33 Cantonese ESL learners majoring in English participated in the study, which consisted of a grammaticality judgment task requiring participants to judge the grammaticality of 60 sentences with regard to the use (or non-use) of the definite article. Two post-task interviews of four interviewees each were conducted to collect in-depth reports about learners’ thinking processes during grammaticality judgment and the strategies they employed. The results of the study show that the generic use of the definite article is largely unknown to Cantonese ESL learners, and learners make certain incorrect hypotheses about the use of *the*. Reliance on L1 translations is also observed, showing learners’ unawareness of the functional and semantic differences between similar items in the two languages. It is suggested that ESL teachers alert learners to their misconceptions and help learners eradicate their faulty assumptions. The generic use of *the*, alongside other structures used to represent generic reference, should be incorporated into the curriculum of an advanced ESL classroom. Future research about English article use should not only focus on analyses of learner errors but also tap into learners’ thinking processes.

Keywords: Cantonese ESL learners, use of the English definite article, grammaticality judgment, the definite generic

Introduction

Articles, being some of the most common English function words, are used in nearly all English written and spoken discourse. English articles can be used to show the scope of reference of the head noun, including definite reference (e.g., *The boy I saw yesterday is called John*), indefinite reference (e.g., *I met a boy*), and generic reference (e.g., *Boys have a lower voice than girls*). The use of English articles has often been found to be problematic for second language learners (Huebner, 1983; Master, 1987; Parrish, 1987; Robertson, 2000; Thomas, 1989). For learners whose native languages lack articles, the problems are even more acute (Ionin, Zubizarreta, & Maldonado, 2008).

Learner Problems in Article Use

Errors associated with different articles have been found. The definite article, which is described as having a “wide variety of usage” and a “higher frequency of use than the indefinite article” (Liu & Gleason, 2002, p. 2), has been the focus of much research. As its name suggests, the definite article is most commonly used for showing definite reference, including anaphoric use (e.g., *Guangzhou is a city in*

China. *The city is prosperous.*), immediate situation use (e.g., “Don’t go in there, chum. *The dog will bite you.*” (Hawkins, 1978, p. 112)), visible situation use (e.g., *Please give me the book.*), etc. (Liu & Gleason, 2002). Many errors with the use of *the* involve overuse (Ionin, Ko, & Wexler, 2004; Ko, Ionin, & Wexler, 2010; Mayo, 2008; Parrish, 1987; Thomas, 1989; van Hout, Harrigan, & de Villiers, 2010), especially in contexts where *a/an* is required, such as in “**I lost the health tooth, and I have realized after some time how it was valuable for me. It happened unexpectedly – I bit off the solid sweet and that’s all*” (quoted in Ionin et al., 2004, p. 4). Thomas (1989) also found that learners overgeneralize *the* in referential indefinite contexts (e.g., *I saw a boy* written as *I saw the boy*). On the other hand, overuse of *a* for *the* in definite contexts is also observed, as in “**He is staying with a mother of his best friend*” (Ionin et al, 2004, p. 35).

Apart from its use for showing definite reference, *the* (+ singular noun) is also used for showing generic reference (e.g., *The dinosaur is extinct*). Known as the definite generic (Snape, 2018) and is associated with the [+species] feature (Ionin, Montrul, Kim & Philippov, 2011), this use of the definite article is rather rare (Parrish, 1987; Tarone & Parrish, 1988; Whitman, 1974), as generic reference can also be represented by *a/an* when one member of a class is picked out to represent the whole class (e.g., *A computer is a machine which can help us store and organize information*), or by *ZERO* when the head noun is plural (e.g., *Dinosaurs are extinct*) or mass (e.g., *Water is useful*). Although EFL learners have been found to be able to make a distinction between the use of *the* for showing generic and definite references (Snape, 2018), misuse of *the* for target *a/an* or *ZERO* for generic reference is very common (e.g., “**These things destroy the society*” (Crompton, 2011, p. 20)).

Explanations for Article Errors

Various explanations have been put forward to account for learner problems in English article use. One contributing factor is L1 transfer. Empirical evidence for L1 transfer and recovery from transfer has been obtained in various studies, including Austin, Pongpairoj, and Trenkic (2015), Ionin and Montrul (2010), Luk and Shirai (2009) and Mede and Gurel (2010). Serratrice, Sorace, Filiaci, and Baldo (2009) also found evidence for L1 transfer for learners whose native languages have articles, yet for learners whose native languages do not have articles, L1 transfer may not be that apparent. For L2 learners whose L1 lacks articles, variability in article productions is seen as a result of the competition between the article + noun forms licensed by the L2 and the article-less forms licensed by the L1 (Trenkic & Pongpairoj, 2013).

Arguments countering the effects of L1 transfer on the L2 acquisition of English articles, however, also exist. Examples include Zdorenko and Paradis (2008), and Zdorenko and Paradis (2012). It is found that similar learner problems (e.g., using *the* for *a* in indefinite specific contexts; showing a fluctuating pattern in article choice) are encountered by children irrespective of whether their L1 has articles or not, so children’s acquisition of English articles is developmental rather than transfer-based.

Learners’ association of *the* with the [+ specific] feature¹ has also been argued as the cause of the overuse of *the* with indefinites (Butler, 2002; Ionin et al., 2004), where learners use the definite form to signal a specific salient referent rather than an identifiable referent (Trenkic, 2004). In a similar fashion, the absence of the [+ specific] feature has been argued to be associated with L2 learners’ tendency to overuse *a* for *the* in definite contexts (Ionin et al., 2004), suggesting that learners associate *the* with the specific reference feature irrespective of whether the target reference is definite or indefinite. Learners’ establishment of non-target form-meaning connections for English articles has also been put forward as a plausible explanation for article errors, in such a way that *a/an* is often used to individuate non-discrete referents whereas *the* is often associated with discrete referents (Trenkic, 2002).

Regarding the overuse of *the* for target *a/an* or *ZERO* for showing generic reference, it has been argued that learners tend to associate *the* with [+Hearer’s Knowledge] ([+HK]) contexts (Geng, 2010): Since both generic contexts and definite contexts share [+ HK] (see Endnote 1), and since learners have an incorrect assumption that generic nouns refer to all entities of a group instead of a whole class, learners often mistakenly think that nouns in generic contexts are identifiable and are thus definite, resulting in

overuse of *the* in such contexts (Chan, in press). How learners understand “what articles signal, and what representations are assigned to them in interlanguage grammars” (Trenkic, 2008, p. 1) is, thus, a main reason for article errors.

Rationale for Study

As mentioned earlier, problems with the use of English articles are more acute for learners whose native languages lack articles (Ionin et al., 2008). Cantonese is a language which does not have structural equivalents of English articles, and it has been found that Cantonese ESL learners of all proficiency levels (advanced, upper-intermediate, lower-intermediate) encounter problems with the use of English articles, including overuse, omission and substitution (Chan, in press). In line with previous research, confusion between specificity and definiteness has also been observed and recognized as a reason for learners’ overuse of *the* for *a* in indefinite contexts (Chan, 2016, 2017a).

Previous studies on English article use, however, mostly focused on learner performance, whereas the strategies that learners adopt in their use of English articles have relatively been less studied². Given the wide variety and high frequency of usage and the ubiquity of errors in the use of the definite article for showing both generic and definite references, it is worth implementing a study which probes into the strategies that Cantonese ESL learners adopt in their use of the definite article.

Although different kinds of errors with *the* have been found and various explanations have been put forward, much less focus has been put on the definite generic. Liu and Gleason (2002), for example, only focused on the non-generic uses of *the*. Chan (in press), on the other hand, did recognize Cantonese ESL learners’ lack of awareness of the definite generic, but it did not examine learners’ actual problems with the use of the definite generic and the corresponding reasons. The present study aimed to bridge this research gap.

Objectives of Study

By means of a grammaticality judgment task and two post-task focus group interviews, the present study aimed to investigate Cantonese ESL learners’ judgment of the use of *the*, including its generic and non-generic uses. The following research question underlay the study:

1. What strategies do Cantonese ESL learners adopt or what hypotheses do they make when deciding whether to use *the* in a certain context?

Given that one of the main focuses of the present study was on the definite generic, research question 1, in effect, subsumed the following:

2. What strategies do Cantonese ESL learners adopt or what hypotheses do they make when encountering the definite generic?

It may be argued that the inclusion of only a grammaticality judgment task and two focus group interviews could not yield results extensive enough for arriving at a reliable conclusion about ESL learners’ use of the definite article, as students often use English articles in different channels in their learning, including essay writing, oral presentations and classroom discussions. However, a grammaticality judgment task, which often requires some conscious language analysis and can measure learners’ metalinguistic knowledge (Kwon, 2018), was deemed suitable for the aim of the present study, as such a task would engage participants in deliberate judgments of the use/non-use of *the* in a certain

context and yield insightful answers to the research questions³. The two post-task focus group interviews could provide additional data for understanding the strategies participants adopted.

Participants

A homogenous group of 33 Cantonese university English majors from a local university participated in the study. Participant selection was based on convenience sampling, including 24 second year and 9 final year students, 4 males and 29 females aged from 19 years to 23 years at the time of the study. All of them were taking English as their majors. Seven of them had learnt English for 14-16 years, nineteen for 17-19 years, and the rest for 20 years or more. 25 of them had received a C or above in the Hong Kong Advanced Level Use of English (UE)⁴ exam, 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)⁵, and the rest had received D in UE or 6.5-7 in IELTS. In view of their linguistic backgrounds, the participants could all be regarded as advanced ESL learners.

Methodology

Grammaticality Judgment Task

60 sentences (or sentence groups)⁶ with or without errors in the use of *the* were included in the study. Participants were required to judge the grammaticality of the sentences and correct the ungrammatical ones accordingly. For each of a total of 43 sentences, there was one noun phrase which was erroneous. Of these, 20 sentences had a missing *the*, whereas 23 sentences had a superfluous *the*. The remaining 17 sentences were grammatical sentences used as distractors⁷.

The sentences included in the task were adapted from documented instruments such as the task sheet used by Liu & Gleason (2002), authentic examples from other sources such as children's fiction and linguistic books, as well as some contrived sentences (see Appendix A). Care was taken to ensure that the items were all clear, appropriate and idiomatic, and that the target corrections for the ungrammatical sentences included different structures and represented different reference types. Table 1 shows the distribution of target structures and reference types for the erroneous noun phrases in the ungrammatical sentences.

TABLE 1

Distribution of Target Structures and Target References for the Erroneous Noun Phrases in the Ungrammatical Sentences

Target Structures	No. of erroneous noun phrases/ungrammatical sentences ⁸ N= 43
Singular (head noun) with <i>the</i>	11
Mass (head noun) with <i>the</i>	6
Plural (head noun) with <i>the</i>	3
Mass (head noun) without <i>the</i>	16
Plural (head noun) without <i>the</i>	7
Target References	No. of erroneous noun phrases/ungrammatical sentences N= 43
Generic	23
Definite	13
Indefinite ⁹	7

Before the implementation of the actual study, 100 sentences (including the 60 chosen) were piloted with two native English speakers and three Cantonese ESL learners also majoring in English at other local universities to ensure that the sentences were in idiomatic English, that the errors in the ungrammatical ones were indeed unacceptable, that the corresponding corrections targeted were idiomatic, and that the level of difficulty of the sentences was suitable for the target group. One of the native speakers was Australian (48 years of age) and the other was British (67 years). Both of them had received tertiary education beyond the undergraduate level. The local students were aged between 18 years and 21 years, with similar linguistic backgrounds as to those of the participants. 40 sentences were deleted after the pilot study (because they did not meet one or more of the requirements specified above), and the remaining 60 were used for the actual task.

Focus Group Interviews

Two focus group interviews with four interviewees each were conducted to probe into learners' thinking processes. Participation was on a voluntary basis. The interviews focused on interviewees' general knowledge of article use (Q.1-Q.2, Appendix B), the problems they encountered in the general use of English articles as well as those they encountered in the grammaticality judgment task (Q.3-Q.4, Appendix B), the strategies they used in their grammaticality judgment (Q.5, Appendix B), and their interpretation of the sentences (Q.6-Q.7, Appendix B). The interviewees were requested to illustrate their strategies using examples from the grammaticality judgment task. Photocopies of the task sheets completed by the interviewees were brought to the interviews to refresh their memories. Relevant technical terms, including *generic reference*, *definite reference*, or *indefinite reference*, were not used by the interviewer to minimize possible hints on or biases towards a response. The interviews were conducted in a mixed code of Cantonese and English (mainly in Cantonese, with some English words or expressions, such as *reference*, *articles*). The proceedings were recorded using a smart phone.

Data Analysis

Grammaticality judgment task

An ungrammatical sentence was considered to have been accurately judged only if the correction given was relevant and grammatical. For example, participants judging the erroneous sentence **Although his father is a pianist, he doesn't like playing ^ piano at all* as incorrect but giving irrelevant "corrections" such as **Although his father is **the** pianist, he doesn't like playing ^ piano at all* were regarded as having given an incorrect judgment, because the target error *the piano* was not corrected.

Similarly, participants judging the grammatical sentence *I heard that **the elephant** never forgets* as ungrammatical with a wrong "correction" **I heard that **elephant** never forgets* were also regarded as having given an incorrect judgment.

Irrelevant grammatical mistakes or corrections which did not affect the reference interpretation and/or article use of the sentences were not taken into account, so a correction such as *I resisted **in giving** rewards until I experienced the power of giving rewards* for the grammatical sentence *I resisted **giving** rewards until I experienced the power of giving rewards* (Gray, 1999, p. 123) was ignored and the corresponding judgment made was regarded as correct.

The percentages of correct judgments (and incorrect judgments) were calculated by dividing the number of sentences correctly judged (and those incorrectly judged) by the total number of sentences judged. Both the correct and incorrect judgments were also classified according to the target reference types and with respect to the target structures of the erroneous noun phrases (e.g. *the* with a plural head noun, *ZERO* with a mass head noun).

Focus group interviews

The post-task interviews recordings were transcribed verbatim by the research assistant who conducted the interviews and corresponding responses were analyzed by the researcher. No coding system was specifically designed to code the interview data. Instead, when explicit reference was made by an interviewee(s) for a certain (grammatical or ungrammatical) sentence(s) and its associated corrections, the strategies employed/hypotheses made in the decision of that particular sentence, as well as the strategies/hypotheses which reflected a similar phenomenon for other sentences, were identified. These strategies/hypotheses will be discussed in the Results and Discussion section below, with actual responses from the interviewees paired up with the sentences under discussion. All the interviewees' Cantonese utterances were translated into idiomatic English by the researcher for inclusion in this paper.

Results and Discussion

Grammaticality Judgment

For the ungrammatical sentences in the grammaticality judgment task, 65.8% were correctly judged as incorrect with correct corrections given. The rest were incorrectly regarded as correct (34.2%). When the target structures and references were taken into account, 81% of the erroneous sentences which should have had *the* + singular head nouns for generic reference (e.g., **Bicycle is a cheap form of private transport*) were incorrectly judged as correct.

By the same token, 24.9% of the erroneous sentences which should have had *the* (with different kinds of head nouns) for definite reference (e.g., **I will never forget ^happiness I had when I dated him*) were incorrectly judged as correct.

On the other hand, 35.4% and 17% of those which should have had *ZERO* + plural or mass head nouns, respectively, for generic reference (e.g., **The dinosaurs are extinct*; **The medicine is a widely respected profession*) and 33.3% of those which should have had *ZERO* + mass head nouns for indefinite reference (e.g., **We all need to do ~~the~~ exercise to keep us healthy no matter how old we are*) were incorrectly judged as correct (see Table 2).

TABLE 2
Judgments of Incorrect Noun Phrases

Target Structures and References	Correct Judgments (Correctly judged to be incorrect and correct corrections given)	Incorrect Judgments (Incorrectly judged to be correct)
Generic Reference		
<i>ZERO</i> + plural (n= 6)*	128/198 (64.6%)	70/198 (35.4%)
<i>ZERO</i> + mass (n= 10)*	274/330 (83%)	56/330 (17%)
<i>the</i> + singular (n= 7)#	44/231 (19%)	187/231 (81%)
Sub-Total	446/759 (58.8%)	313/759 (41.2%)
Definite Reference		
<i>the</i> + singular (n= 4)#	106/132 (80.3%)	26/132 (19.7%)
<i>the</i> + plural (n= 3)#	73/99 (73.7%)	26/99 (26.3%)
<i>the</i> + mass (n= 6)#	143/198 (72.2%)	55/198 (27.8%)
Sub-Total	322/429 (75.1%)	107/429 (24.9%)
Indefinite Reference		
<i>ZERO</i> + plural (n= 1)*	33/33 (100%)	0/33 (0%)
<i>ZERO</i> + mass (n= 6)*	132/198 (66.7%)	66/198 (33.3%)
Sub-Total	165/231 (71.4%)	66/231 (28.6%)
Total	933/1419 (65.8%)	486/1419 (34.2%)

Note. *the erroneous noun phrases all have an unnecessary *the*
the erroneous noun phrases all have *the* omitted

As for the grammatical sentences, *the* + singular noun phrases for generic reference (e.g., *I heard that the elephant never forgets*) were often mis-judged and “corrected” with *the* omitted¹⁰. There were also cases of correct non-use of *the* (e.g., *You are so dirty. You’ve got egg on your shirt*) being mis-judged and “corrected” with a redundant *the*.

Tables 3 and 4 show some examples of correct sentences with incorrect corrections and some examples of incorrect sentences judged as correct, respectively.

TABLE 3

Examples of Correct Sentences Judged as Incorrect and Corresponding Incorrect Corrections

Correct Sentences	Incorrect Corrections	Percentages
At <i>Ocean park</i> I saw some penguins.	*At <i>the Ocean Park</i>	7 participants; 21.2%
Animals are talented. I heard that <i>the elephant</i> never forgets. *I heard that <i>elephant/ the elephants</i> never forgets/forget.	24 participants; 72.7%
A special kind of shortening characterized by the omission of <i>the initial unstressed syllable of a word</i> is the aphetic form (adapted from Jackson & Amvela, 2000, p. 89).	*.... the omission of <i>initial unstressed syllable of a word</i> is the aphetic form.	6 participants; 18.2%
<i>Tests conducted on babies</i> reveal that baby girls respond almost twice as much to a loud sound than boys (Pease A. & Pease B., 2001, p. 30).	* <i>The tests conducted on babies</i> reveal	7 participants; 21.2%
Our understanding of <i>the human brain</i> is increasing dramatically every day (Pease A. & Pease B., 2001, p. 48).	*Our understanding of <i>human brain</i> is	11 participants; 33.3%

TABLE 4

Examples of Incorrect Sentences Judged as Correct

Sentences with redundant <i>the</i>	Percentages
* The <i>dinosaurs</i> are extinct.	9 participants; 27.3%
*I like to watch the <i>movies that are black and white</i> , because they give me a nostalgic feeling (adapted from Liu & Gleason, 2002).	12 participants; 36.4%
*I like reading very much, especially the <i>books about physics</i> .	15 participants; 45.5%
*If we have a mental illness we will see a psychiatrist, but I never understand how the <i>psychiatrists</i> can help us.	15 participants; 45.5%
*A lot of women want to end the <i>male superiority</i> .	21 participants; 63.6%
Sentences Missing <i>the</i>	Percentages
*^ <i>Professor who taught biology last year</i> explained things very clearly.	4 participants; 12.1%
*^ <i>Furniture in this house</i> does not suit me.	9 participants; 27.3%
*I saw two women across the street. I later realized that ^ <i>woman wearing a blue jacket</i> was my teacher.	8 participants; 24.2%
*We have been told that ^ <i>beaver</i> can build dams.	26 participants; 78.8%
*If we look at ^ <i>clause</i> as a unit of communication, we can say that the first constituent is the point of departure (adapted from Downing & Locke, 2006, p. 223)	25 participants; 75.8%

It can be seen from the above that both overuse of *the* and omission of *the* were observed, and even grammatical sentences were “corrected”, showing that learners’ grammaticality judgments should have been led by some (correct or) incorrect hypotheses about the use and/or non-use of *the*. The following sections will address these strategies/hypotheses.

Learner Strategies and Hypotheses (Research Question 1)

Reliance on previous linguistic repertoire

In determining the use/non-use of *the* in a certain context, learners tended to use different kinds of strategies, one of which was reliance on their intuitions, habitual uses, and/or previous linguistic repertoires. In the post-task interviews, Student J commented on his/her use of intuitions in making a

decision, and in deciding whether to use *the* or not for *Ocean Park* and *baleen whales*, Student A and Student C expressed their reliance on their previous linguistic repertoires.

Comment	Sentences under Discussion
(a) <i>You can have the. You can have no the. I use my own intuition.</i> (Student J)	<i>General strategies employed</i>
(b) <i>Actually I wrote at the Ocean Park. I think it reads well. Ever since I was young I have used the Ocean Park.</i> (Student A)	<i>At Ocean Park I saw some penguins. I think penguins are beautiful animals.</i>
(c) <i>Very often when you are talking about the characteristics of a group, then we don't need the at the beginning.</i> (Student C)	<i>Baleen whales are more often found singly instead of in groups.</i> (Mathews & Read, 1982, p. 55)

Reliance on previous linguistic repertoires was occasionally manifested in learners' modelling the target syntactic structure on the structure of similar noun phrases not given in the sentences. For example, in deciding whether to use *the* or not for *Ocean Park*, Student E relied on similar noun phrases such as *Hong Kong*, *Lantau Island*, etc. as signposts for guiding his/her decision.

Comment	Sentence under Discussion
(d) <i>We don't need the for Hong Kong. We don't need the for Landmark Island, so according to the same system, to be consistent, I don't think we need the for Ocean Park.</i> (Student E)	<i>At Ocean Park I saw some penguins. I think penguins are beautiful animals.</i>

Effects of immediate linguistic environments

Other strategies that learners adopted in their judgment of the sentences included their making reference to the immediate syntactic environments in which an article appeared. A predominant misconception was that the presence of a post-modifier in a noun phrase should always necessitate *the* for the head noun, as shown by Student Cr's comment below:

Comment	Sentence under Discussion
(e) <i>I think probably the books they are talking about are a specific kind, about physics. The books are about physics. Probably just a group, so I think with the.</i> (Student Cr)	<i>*I like reading very much, especially the books about physics.</i>

Learners' assumption that the presence of a post-modifier in a noun phrase necessitates the use of *the* has already been observed in Chan (2017a). However, instead of resulting in correct use of *the* for definite reference as attested in Chan (2017a), the present study revealed that the same hypothesis could result in unnecessary use of *the* in non-definite contexts. This shows that while learners are aware of the significance of the syntactic environments on the interpretation of a noun, they often mistakenly associate noun phrase post-modifiers as a default indication of definite reference, probably as a result of the widespread use of noun phrase post-modifiers for cataphoric, definite reference (e.g., *The boy with a yellow shirt has left*).

Noun countability

Although noun countability has mainly been found to be important for learners' determination of the need for *a/an* (Chan, 2017b), it was also regarded as an important criterion for the use/non-use of *the* in the present study. Learners made their decisions about *the* based on their judgments of the countability of a noun. However, in line with the findings of Chan (2016) that Cantonese ESL learners often have difficulties judging the countability of an English noun, noun countability was sometimes misjudged in the present study despite correct grammaticality judgments and corrections:

Comment	Sentence under Discussion
(f) <i>I don't know whether furniture is countable or uncountable. I assume that it is countable, but then I thought: why is there no s? Therefore, I add the.</i> (Student E)	<i>*^ Furniture in this house does not suit me.</i>

On the other hand, some learners mistakenly thought that if a noun was uncountable, then the definite article would not be needed by default:

Comment	Sentence under Discussion
(g) <i>I remember furniture is uncountable, then uncountable does not need the, so I don't have the.</i> (Student H)	<i>As above</i>

Problems with the Definite Generic (Research Question 2)

The definite generic (e.g., *The elephant never forgets*) was often rejected, because learners had a misconception that only ZERO should be used with generic reference and *the* + singular should have definite reference by default.

Comment	Sentence under Discussion
(h) <i>Normally we should say Elephants never forget. I don't understand why there is the elephant here.</i> (Student A)	<i>Animals are talented. I heard that the elephant never forgets.</i>

Learners' rejection of the definite generic also resulted in their inaccurate use of *the* + plural for generic reference:

Comment	Sentence under Discussion
(i) <i>I think they are talking about all elephants. I think the elephant never forgets seems to be very odd. I have not learnt that the is compulsory, I only think that if it is expressing all elephants, then we should say elephants. For generic, there shouldn't be the with singular. I just added "s" to the elephant and delete "s" from forgets.</i> (Student K)	<i>As above</i>

As argued in Chan (2017a), many ESL learners regard generic reference as referring to all entities of a group instead of a class of entities, so sentences such as *All the teachers are well-paid*, which have definite reference referring to all entities of a subgroup, are often interpreted as having generic reference. This "referring to all" confusion may have given rise to default pluralization of the head noun, resulting in incorrect use of *the* + plural for generic reference.

Bare singular noun phrases were also found as a result of learners' rejection of the definite generic:

Comment	Sentence under Discussion
(j) <i>I think in the same way as Student A. I think it is very odd to say the elephant, so I just deleted the.</i> (Student E)	<i>As above</i>

The use of a bare singular noun phrase for generic reference was likely the result of the widespread use of ZERO + plural for generic reference: As learners were familiar with this structure for generic reference, in seeing the sentence *The elephant never forgets*, they opted for the ZERO article without corresponding addition of a plural morpheme for the head noun in order to retain a seemingly similar structure with correct subject verb agreement. As Trenkic, Mirkovic, and Altmann (2014) claim, learners of a native language without structural equivalents of English articles often have the tendency to favor bare noun phrases. Our findings have provided evidence for such a tendency.

Possible L1 Influence

Although the study was not designed to investigate L1 influence, the results showed that learners occasionally resorted to their native language in grammaticality judgment despite the absence of structural equivalents of English articles in Cantonese, albeit with some misconceptions about the two language systems. In line with the findings of Ionin, Baek, Kim, Ko, and Wexler (2012) that L2 learners may be influenced by transfer from related categories such as demonstratives, our findings witnessed a “translation and comparison” strategy, in that an article choice (whether to use *the* or not) was validated by comparing a given English sentence or phrase with its seemingly equivalent Chinese version, notably a phrase or sentence with a Cantonese determiner showing definite reference (e.g., demonstratives 嗰個 (*that*) or 嗰啲 (*those*)). When a Cantonese demonstrative was needed in the corresponding Chinese sentence, then learners would be inclined to use *the* in the English sentence. When the corresponding Chinese sentence did not require a demonstrative, then learners would be inclined to use *ZERO* in the English sentence. Student E’s comment demonstrates this strategy:

Comment	Sentences under Discussion
(k) Like number 18, during the first two years . The first two or three years. They are just like 嗰個 (<i>that</i>).... Sometimes I translate it into Chinese... I translate that into Chinese and see if it reads smooth. (Student E)	*During ^ first two or three years of development , we need to interact with other language-users in order to learn to speak. (Yule, 2006: 149)

This strategy of learners’ seeking confirmation for their English article choices from the need for a corresponding Cantonese demonstrative may be explained by the similarity between the functions of English *the* and Cantonese demonstratives. As discussed in Chan (2004), Cantonese demonstratives are used for deictic functions pointing at or referring back to noun phrases mentioned in the same context. This anaphoric function is similar to the most common function of English *the*. With the results of the present study, it is not clear whether learners’ reliance on the L1 will have adverse effects on their acquisition of English *the*, but this is an area worth looking into, as it not only demonstrates learners’ unawareness of the functional and semantic differences between English articles and Cantonese demonstratives and/or other determiners, but it also reveals the fact that learners’ correct article choices may be the result of ungrounded comparisons between the L1 and the L2¹¹.

Implications

The results of the present study are illuminating, revealing learners’ unawareness of the use of the definite generic, their inappropriate reference made to the immediate linguistic environments of a noun, their problems with noun countability and bare singular noun phrases, as well as possible L1 influence. Despite the fact that Cantonese is an article-less language and there are no structural equivalents of English articles, there is reason to believe that the use of a contrastive analysis approach to the teaching of English articles would help eradicate learners’ inappropriate reliance on the L1. ESL learners can be engaged in explicit analyses discussing the functional and semantic similarities and differences between such similar but distinct items in the two languages as English articles and Cantonese demonstratives, possessives and even numerals.

For advanced ESL learners, another important area of teaching focus is the use of the definite generic. Although the definite generic is very rare (Parrish, 1987; Tarone & Parrish, 1988), there has been evidence showing that *the* is the most likely article used for generic reference in subject position, in the first sentence of a paragraph, in introductions and conclusions, and in marking the topic of a scientific essay (Master, 1987). As suggested by the Noticing Hypothesis (Schmidt, 1990), one has to notice a language feature before one can learn it. Advanced ESL learners should be alerted to the definite generic, such as when there are contexts which allow generalizations to be made about a class with clear defining

characteristics. Direct correction and metalinguistic explanations, which direct learners to pay attention to specific linguistic forms with grammatical explanations (Domakani, 2008; Oyama, 2017), are needed to enhance long-term learning. ESL teachers can introduce explicit analyses targeting the different uses of *the* to consolidate learners' understanding of the use of the definite article for representing different references, as well as comparing and contrasting the definite generic with the use of other articles for showing generic reference, such as *ZERO* + plural/mass and *a/an* + singular. A truth value determination and discussion task using some short stories and a few accompanying test statements, like the following¹², may serve this purpose.

There are two female kangaroos in our zoo. They are very special. They carry their babies on their backs!!! But we all know that a female kangaroo normally carries its babies in its pouch.

<i>The female kangaroo carries its babies in its pouch.</i>	T/F
<i>The female kangaroo carries its babies on its back.</i>	T/F
<i>The female kangaroos carry their babies on their backs.</i>	T/F
<i>Female kangaroos carry their babies on their backs.</i>	T/F
<i>Female kangaroos carry their babies in their pouches.</i>	T/F

In addition to showing the stories in written form, ESL teachers can prepare pictures depicting the situations described in the stories (e.g., two female kangaroos in the zoo carrying their babies on their backs) so as to provide both visual and written stimuli for motivating interest. After familiarizing students with the meanings of the stories, teachers can engage learners in a discussion about the truth values of the accompanying test statements by focusing on the differences and similarities between the sentence structures of the different statements as well as the references represented.

Conclusion

In this paper, the results of a study investigating Cantonese ESL learners' judgment of the use of *the* are reported. It is found that learners sometimes rely on their previous linguistic repertoire as the basis for deciding whether *the* is needed in a certain context. Certain misconceptions about *the* are evident, including the use of *the* + plural for generic reference. Inappropriate use of bare singular noun phrases and confusion about similar items in the target and native languages are also witnessed. It is suggested that contrastive analyses between similar but distinct items in the native and target languages, as well as grammatical analyses focusing on the definite generic, be incorporated into the curriculum of an ESL classroom. If learners have a heightened awareness of their problems, their use of *the* will then be led by enhanced explicit knowledge instead of ungrounded assumptions. Future research into the performance and knowledge of students of other proficiency levels (e.g., intermediate and elementary levels), of other disciplines, (e.g., social sciences or engineering), or with other mother tongues (e.g., Korean or French), will also be useful in yielding insightful findings about other learner hypotheses as well as the extent of L1 influence.

Acknowledgements

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 participants of the study for their participation and my research assistants for their administrative help.

The Author

Alice Yin Wa Chan is Associate Professor at the Department of English, City University of Hong Kong. Her research interests include second language acquisition, English grammar, English phonetics and phonology, and lexicography.

Department of English
City University of Hong Kong
83 Tat Chee Avenue
Kowloon, Hong Kong
Tel: 852 3442 9752
Email: enalice@cityu.edu.hk

References

- Austin, G., Pongpairoj, N., & Trenkic, D. (2015). Structural competition in second language production: Towards a constraint-satisfaction model. *Language Learning*, 65(3), 689–722.
- Butler, Y. G. (2002). Second language learners' theories on the use of English articles: An analysis of the metalinguistic knowledge used by Japanese students in acquiring the English article system. *Studies in Second Language Acquisition*, 2, 451–480.
- Chan, A. Y. W. (2004). Noun phrases in Chinese and English: A study of English structural problems encountered by Chinese ESL students in Hong Kong. *Language, Culture and Curriculum*, 17(1), 33–47.
- Chan, A. Y. W. (2016). How much do Cantonese ESL learners know about the English article system? *System*, 56, 66–77.
- Chan, A. Y. W. (2017a). Why do Hong Kong Cantonese ESL learners choose a certain English article for use? *Asian Journal of Applied Linguistics*, 4(1), 16–29.
- Chan, A. Y. W. (2017b). The effectiveness of using a bilingualized dictionary for determining noun countability and article selection. *Lexikos*, 27, 183–213.
- Chan, A. Y. W. (2019). L1 influence on the acquisition of English articles by Cantonese ESL learners in Hong Kong: Presence or absence? *Journal of Language and Communication*, 6(1), 103–113.
- Chan, A. Y. W. (in press). Typology and contexts of article errors: Investigation into the use of English articles by Hong Kong Cantonese ESL learners. *International Review of Applied Linguistics in Language Teaching*.
- Crompton, P. (2011). Article errors in the English writing of advanced L1 Arabic learners: The role of transfer. *Asian EFL Journal*, 50, 4–34.
- Domakani, M. R. (2008). Formal instruction and the acquisition of linguistic features. *The Journal of Asia TEFL*, 5(2), 35–58.
- Downing, A., & Locke, P. (2006). *English grammar: A university course* (2nd ed.). London: Routledge.
- Geng, J. (2010). The semantic analysis of the definite article' misuse by Chinese learners of English. *Asian Social Science*, 6(7), 180–184.
- Gray, J. (1999). *Children are from heaven: Positive parenting skills for raising cooperative, confident, and compassionate children*. New York: HarperCollins Publishers.
- Hawkins, J. (1978). *Definiteness and indefiniteness*, San Diego, CA: Academic Press.
- Huebner, T. (1983). *A longitudinal analysis of the acquisition of English*. Karoma, Ann Arbor.
- Huebner, T. (1985). System and variability in interlanguage syntax. *Language Learning*, 35, 141–163.
- Ionin, T., Baek, S., Kim, E., Ko, H., & Wexler, K. (2012). *That's not so different from the*: Definite and demonstrative descriptions in second language acquisition. *Second Language Research*, 28, 69–101.

- Ionin, T., Ko, H., & Wexler, K. (2004). Article semantics in L2-acquisition: The role of specificity. *Language Acquisition*, 12, 3–69.
- Ionin, T., & Montrul, S. (2010). The role of L1 transfer in the interpretation of articles with definite plurals in L2 English. *Language Learning*, 60(4), 877–925.
- Ionin, T., Montrul, S., Kim, J.-H., & Philippov, V. (2011). Genericity distinctions and the interpretation of determiners in second language acquisition. *Language Acquisition*, 18(4), 242–280.
- Ionin, T., Zubizarreta, M. L., & Maldonado, S. B. (2008). Sources of linguistic knowledge in the second language acquisition of English articles. *Lingua*, 118, 554–576.
- Jackson, H., & Amvela, E. Z. (2000). *Words, meaning and vocabulary: An introduction to modern English lexicology*. London; New York: Continuum.
- Ko, H., Ionin, T., & Wexler, K. (2010). The role of presuppositionality in the second language acquisition of English articles. *Linguistic Inquiry*, 41(2), 213–254.
- Kwon, S. H. (2018). A study on NNS teachers' language awareness. *The Journal of Asia TEFL*, 15(1), 51–65.
- Liu, D., & Gleason, J. L. (2002). Acquisition of the article *the* by nonnative speakers of English: An analysis of four nongeneric uses. *Studies in Second Language Acquisition*, 24, 1–26.
- Luk, Z. P., & Shirai, Y. (2009). Review article: Is the acquisition order of grammatical morphemes impervious to L1 knowledge? Evidence from the acquisition of plural -s, articles, and possessive 's. *Language Learning*, 59(4), 721–754.
- Master, P. (1987). *A cross-linguistic interlanguage analysis of the acquisition of the English article system* (Doctoral dissertation). University of California, Los Angeles.
- Matthews, A., & Read, C. (1982). *Themes: An integrated skills course for late intermediate and advanced students*. Glasgow: Collins.
- Mayo, M. del P. G. (2008). The acquisition of four nongeneric uses of the article 'the' by Spanish EFL learners. *System*, 36(4), 550–565.
- Mede, E., & Gurel, A. (2010). Acquisition of English articles in early bilingualism. *EUROSLA Yearbook*, 10(1), 193–219.
- Oyama, R. (2017). Effects of cognitive comparison during focus-on-form instruction on Japanese EFL learners' acquisition of grammar. *The Journal of Asia TEFL*, 14(1), 1–15.
- Parrish, B. (1987). A new look at methodologies in the study of article acquisition for learners of ESL. *Language Learning*, 37, 361–383.
- Pease, A., & Pease, B. (2001). *Why men don't listen & women can't read maps*. London: Orion Publishing Group.
- Robertson, D. (2000). Variability in the use of the English article system by Chinese learners of English. *Second Language Research*, 16(2), 135–172.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied linguistics*, 11(1), 17–46.
- Serratrice, L., Sorace, A., Filiaci, F., & Baldo, M. (2009). Bilingual children's sensitivity to specificity and genericity: Evidence from metalinguistic knowledge. *Bilingualism: Language and Cognition*, 12(2), 239–257.
- Snape, N. (2018). Definite generic vs. definite unique in L2 acquisition. *Journal of the European Second Language Association*, 2(1), 83–95.
- Tarone, E., & Parrish, B. (1988) Task-related variation in interlanguage: The case of articles. *Language Learning*, 38, 21–44.
- Thomas, M. (1989). The acquisition of English articles by first- and second-language learners. *Applied Psycholinguistics*, 10, 335–355.
- Trenkic, D. (2002). Form-meaning connections in the acquisition of English articles. In S. Foster-Cohen, T. Ruthenberg, & M. Poschen (Eds.), *EUROSLA Yearbook 2* (pp. 115–133). Amsterdam: John Benjamins.

- Trenkic, D. (2004). Definiteness in Serbian/Croatian/Bosnian and some implications for the general structure of the nominal phrase. *Lingua*, 114(11), 1401–1427.
- Trenkic, D. (2008). The representation of English articles in second language grammars: Determiners or adjectives? *Bilingualism: Language and Cognition*, 11, 1–18.
- Trenkic, D., Mirkovic, J., & Altmann, G. (2014). Real-time grammar processing by native and non-native speakers: Constructions unique to the second language. *Bilingualism: Language and Cognition*, 17(2), 237–257.
- Trenkic, D., & Pongpaioj, N. (2013). Referent salience affects second language article use. *Bilingualism: Language and Cognition*, 16, 152–166.
- van Hout, A., Harrigan, K., & de Villiers, J. (2010). Asymmetries in the acquisition of definite and indefinite NPs. *Lingua*, 120(8), 1973–1990.
- Whitman, R. L. (1974). Teaching the article in English. *TESOL Quarterly*, 8, 253–262.
- Zdorenko, T., & Paradis, J. (2008). The acquisition of articles in child L2 English: Fluctuation, transfer, or both? *Second Language Research*, 24, 227–250.
- Zdorenko, T., & Paradis, J. (2012). Articles in child L2 English: When L1 and L2 acquisition meet at the interface. *First Language*, 32(1-2), 38–62.
- http://www.hkeaa.edu.hk/DocLibrary/HKALE/Subject_and_Syllabuses/2013/2013as-e-ue.pdf
- http://www.hkeaa.edu.hk/en/recognition/ce_al_recognition/

Appendix A

Some Sample Sentences Included in the Grammatical Judgment Task

CORRECT SENTENCES

- At Ocean Park I saw some penguins. I think penguins are beautiful animals.
- I read many detective stories last week. I like reading detective stories.
- I bought two new computers yesterday. I think computers will override humans someday.
- Animals are talented. I heard that the elephant never forgets.
- Coffee is more tasty than water, but it is less healthy.

UNGRAMMATICAL SENTENCES WITH OMISSION OF *THE*:

- *[^] *Professor* who taught biology last year explained things very clearly.
- *Before the test, we were told to write all our answers in [^] *answer books* provided.
- *Although his father is a pianist, he doesn't like playing [^] *piano* at all.
- *[^] *Bicycle* is a cheap form of private transport.
- *I saw two women across the street. I later realized that [^] *woman* wearing a blue jacket was my English teacher.

UNGRAMMATICAL SENTENCES WITH UNNECESSARY *THE*:

- *We all need to do ~~the~~ *exercise* to keep us healthy no matter how old we are.
- *The garden is covered with ~~the~~ *colorful flowers*.
- *~~The~~ *dinosaurs* are extinct.
- *~~The~~ *beer* is a drink.
- *I like reading very much, especially ~~the~~ *books about physics*.

Appendix B

Interview Questions

1. Do you think your knowledge of the English article system is enough for your learning needs?
2. Do you think your knowledge of the English article system helps you in your selection of articles? If so, how? If not, why?
3. What difficulties did you encounter in the use and/or understanding of English articles?
4. Did you encounter any difficulties in the grammaticality judgment task? If yes, what?
5. What strategies did you use to overcome the difficulties? Have you been taught these strategies?
6. For the sentences which you found inaccurate, was your original understanding of the sentences affected by the mistake(s) identified?
7. For the sentences which you found inaccurate, did you think that more than one answer could be acceptable? Which sentences? Why did you think so?

¹ Under Huebner's (1983, 1985) semantic wheel analysis, noun phrases are classified by two features of referentiality: Specific Reference (SR) and Hearer's Knowledge (HK). Nouns are classified as +/- *specific reference* and +/- *hearer's knowledge*. Four basic contexts determine article use: Type 1: [-SR, +HK]: generic (e.g. *The dinosaur* is extinct); Type 2: [+SR, +HK]: referential definites, including nouns previously mentioned (e.g. Mary bought a purse. *The purse* was expensive), specified by entailment (e.g. I arrived at his place and rang *the bell*), unique in all contexts (e.g. *The moon* is round), etc.; Type 3: [+SR, -HK]: referential indefinites, including first mention nouns, whose referent is identifiable to the speaker but not to the listener (e.g. I saw *a boy*); and Type 4: [-SR, -HK]: non-referentials, including attributive indefinites (e.g. John is *a teacher*), nonspecific indefinites (e.g. I want *a new pen*), etc.

² Chan (2017a) and Butler (2002) are notable exceptions, but the focuses of these studies were on the selection of an English article in a certain context instead of deliberate judgments of whether to use *the*.

³ A grammaticality judgment task has been used in the literature for understanding ESL/EFL learners' acquisition of *the*, such as Liu & Gleason (2002).

⁴ The Hong Kong Advanced Level Use of English (UE) examination aimed to test students' ability to understand and use English at a level that was required for tertiary education and/or for future employment. (http://www.hkeaa.edu.hk/DocLibrary/HKALE/Subject_and_Syllabuses/2013/2013as-e-ue.pdf). It was normally taken by F.7 students in Hong Kong who had completed their two-year matriculation studies. UE Grade E was regarded as equivalent to Grade E in the GCE A level examinations (http://www.hkeaa.edu.hk/en/recognition/ce_al_recognition/).

⁵ The Hong Kong Diploma of Secondary Education examination aims to measure the attainment of students upon their completion of six years of secondary education and has been the only public examination in the new 3-3-4 education system in Hong Kong since 2012. 5** is the highest grade that students can attain for a certain subject, followed by 5* and 5.

⁶ For the sake of convenience, in the rest of the paper, the word *sentence(s)* will be used to refer to the items included in the grammaticality judgment task, although some of the items were actually sentence groups instead of isolated sentences.

⁷ It may be argued that a balanced distribution of sentences with missing *the*, superfluous *the*, and distractors may yield more reliable results, and so may the same number of sentences with different structures (e.g. with mass head nouns, with plural head nouns, etc.) or different references (e.g. definite vs. generic). This conviction was not taken in the study, as the same number of questions for different types of structures /references might give hints to participants as to what kinds of structures to expect. As the objective of the study was not to compare participants' performance on different kinds of structures, comparable but not identical numbers of sentences with different structures/references were deemed more appropriate.

⁸ There is only one erroneous noun phrase in each ungrammatical sentence, so the number of target noun phrases is the same as the number of ungrammatical sentences.

⁹ Although *the* is not used for showing indefinite reference, the target structures for some erroneous noun phrases with an unnecessary *the* showed indefinite reference.

¹⁰ In a grammatical sentence, there were often more than one correct noun phrase. Different participants could make incorrect judgments on different correct noun phrases, so a numerical analysis of the percentages of correct sentences being misjudged (like those given in Table 2 for incorrect sentences) is not meaningful.

¹¹ A similar kind of reliance on the L1 has also been observed in Chan (2019), which reports the results of two other sub-studies of the present study.

¹² The truth value judgment task described here was inspired by Ionin & Montrul's (2010) study, which introduced a similar task to investigate the interpretation of plural noun phrases by Spanish learners of English.