

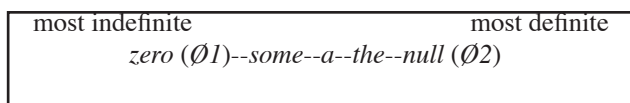
Acquisition of the Zero and Null Articles in English

Peter Master
San José State University

This paper analyzes spoken interlanguage data from 15 non-native speaker (NNS) at three English interlanguage levels representing five native language (L1) backgrounds (Chinese, Japanese, Russian, Spanish, and German) to describe the acquisition of the zero and null articles, the first of which occurs in indefinite and the second in definite noun phrases. The lack of a marked difference in the acquisition of the two forms suggests that learners are generally not aware of the distinction between the zero and null articles.

For many years, the zero article was ignored in article research because it was invisible or thought to require some kind of (potentially biased) judgment to determine whether or not it was present. For example, two well-known studies, Brown (1973) and Dulay and Burt (1974), failed to consider the acquisition of zero articles. Pedagogical materials followed suit in requiring students to determine whether or not an article (*a*, *an*, or *the*) was required in a passage. It is common practice now, however, to treat the zero article as an article in its own right, a practical solution as it allows the contrast of \emptyset and *a* in countability and number distinctions. This practice is also justified in light of the fact that the zero article is the most frequently occurring free morpheme in the English language (Master, 1997). A study of the zero article in interlanguage analyses does have a methodological limitation, however, in that there is no *a priori* means of determining whether the zero article was used deliberately, or whether an overt article was simply left out by mistake. Nevertheless, for the purposes of this paper, I refer to the subjects "use" of the zero article as if it were an overt morpheme like *a* or *the*.

The zero article has two forms, which Chesterman (1991) labeled the zero and the null article.¹ He described the following continuum between zero and null (p. 182):



The zero and null articles are thus maximally opposed in the article schema. The zero article ($\emptyset 1$) occurs most frequently with indefinite noncount (e.g., *milk*) and plural count nouns (e.g., *eggs*). It is also the preferred means of describing generic or nonspecific nouns, especially in the plural count form.² The types of nouns that occur with the zero article are summarized in Table 1 (examples from Master, 1997).

Table 1. Noun Types with the Zero Article (Ø1)

Noun Type	Example
Mass (vs. Count)	The boys ate <u>chicken</u> . The locals shot <u>boar</u> . Mice like <u>cheese</u> .
Abstract (vs. Concrete)	<u>Prison</u> dehumanizes people. My brother is in <u>bed</u> . Herbert traveled by <u>car</u> . They communicate by <u>radio</u> .
Intentional vagueness	<u>Replication of cells</u> takes place over several hours. <u>Animals in underground caves</u> are often blind and colorless.
“Adjective” (vs. Noun)	He was <u>man</u> (i.e., manly) enough to accept his fate.

[Note: The zero article occurs in front of each underlined noun.]

The null article (Ø2) is the most definite of the articles. The null article occurs most frequently with bounded singular proper nouns, that is, entities with “an exterior boundary that ...is limited *a priori*” (Chesterman, 1991, p. 86). It also occurs with certain singular count nouns. The types of nouns that occur with the null article are summarized in Table 2 (examples from Master, 1997).

Table 2. Noun Types with the Null Article (Ø2)

Noun Type	Example
Bounded Proper	<u>Italy</u> is a fascinating country.
Rank, Position, or Post	Mr. Jones was appointed <u>chairman</u> . She was hired as <u>special assistant to the president</u> .
Familiar (vs. Unfamiliar)	
Time	After <u>dinner</u> , we'll see a movie. It usually snows in <u>winter</u> . I'll call you <u>next week</u> .
Place	I left it at <u>home</u> .
And-coordinated NP	<u>Brother and sister</u> were inseparable.

[Note: The null article occurs in front of each underlined noun.]

The zero and null articles can be readily distinguished by their paraphrasability by either an indefinite or a definite article, respectively (Master, 1997). The following sample sentences based on the NP *lunch* exemplify Chesterman's continuum:

- a. **People should never go without (Ø1) lunch.** Indefinite zero article with an indefinite noncount noun. Paraphrasable by “a meal.”
- b. **Did you have some lunch?** Indefinite determiner *some* with an indefi-

nite noncount noun. *Some lunch* is less indefinite than *lunch* because it is partitive (Greenbaum, & Quirk. 1990; p. 124 call unstressed *some* a “partitive indefinite”), suggesting a limited amount. Paraphrasable by “a meal.”

c. **It was a lunch of uncommon quality.** Indefinite determiner *a* with an indefinite count noun; because of its countability, *a lunch* is more definite than *some lunch*. Paraphrasable by “a meal.”

d. **The lunch consisted of bread and cheese.** Definite determiner *the* with a definite count noun (a specific, identified lunch, e.g., the one that was served to the exhausted fire fighters). Paraphrasable by “the meal.”

e. **Thanks for (Ø2) lunch!** Definite null article with a definite count noun (a specific, identified lunch, e.g., the one we have just finished). More definite than *the lunch* in that it presumes that the party being thanked has detailed knowledge of the entire shared (familiar) experience. Paraphrasable by “the meal.”

Thus, sentence (a) *People should never go without (Ø1) lunch* may be paraphrased as “People should never go without a meal” but not as “*People should never go without the meal.” Likewise, sentence (e) *Thanks for (Ø2) lunch!* may be paraphrased as “Thanks for the meal!” but not as “*Thanks for a meal!” The distinction between zero and null is thus determined entirely by the context in which it occurs.

Master (1987) studied the spoken acquisition of the English article system, including the articles *a(n)*, *the*, and \emptyset , the zero article, although this work did not consider the null article or articles with proper nouns. The 1987 study analyzed recorded interlanguage (IL) data from 20 learners who spoke five languages at four interlanguage levels, each determined by the negation criteria described by Cazden, Cancino, Rosansky, and Schumann (1975) to approximate the rough interlanguage levels of *basilang* and low, mid, and high *mesolang* (the terms *basilang* and *mesolang* in the interlanguage continuum parallel the terms *basilect* and *mesolect* in the creole continuum, *basilect* being the farthest away from the *acrolect*, which is closest to the standard variety of a major international language). As four learners at successive levels rather than a single individual together described the full IL continuum, the data constituted five pseudolongitudinal studies, three for languages with no article system, Chinese, Japanese, and Russian, or [-ART] languages, and two for languages with article systems, Spanish and German, or [+ART] languages. The learners included nine males and eleven females with an age range of 13-93 years. The amount of time they had been in the U.S. ranged from three weeks to 32 years, and the level of education ranged from elementary to university level.

The data were taken from tape-recorded informal interviews that were conducted between 1980 and 1986. The topic of the interviews was general information about the learners, including what life was like in their home countries, how they had come to the U.S., what their feelings were about the U.S., etc., and every effort was made to make the learners feel relaxed and comfortable so as to provide an optimally natural speech sample. The majority (15/20) of the interviews were

conducted by students from John Schumann's *Contrastive Analysis* classes (English 241K) at UCLA during the years mentioned, the remaining five by the author. Each of the learner's interviews was transcribed and analyzed by the original interviewer in terms of verb phrase morphology (tense, aspect, number) and negation as well as certain aspects of noun phrase morphology (plural and possessive markers), but not for article use. A prediction was then made as to the learner's interlanguage (IL) level. In selecting the participants for this study, I considered the learner's negation characteristics, the interviewer's stated IL level, and the learner's morpheme acquisition at the 20%, 50%, and 70% accuracy levels in order to be sure that the learners had the desired native language at the appropriate IL level. The accuracy of each transcript was carefully double-checked against the original tape-recording, and a record was kept for every *uh* that may have been mistaken for the article *a* (these accounted for 2.6% of the data at the basilevel and diminished to zero by the mid-mesolang level and were thus not considered a serious problem). The final form of the data was a series of entries for each learner indicating the number of the utterance, the page and line number from whence it came in the transcript, the noun phrase being considered plus minimal attendant context to justify its classification, the article used, the article required, and the classification code.

Analysis of the data revealed continuous improvement in the accuracy of use of the system as a whole (in contrast to accuracy with specific articles or usages), suggesting that a simple test of general article usage might serve as a valid predictor of English proficiency, as was also found by Oller and Redding (1972). The validity of these results was also supported by patterns of individual article acquisition that paralleled three published studies of article acquisition (Andersen, 1977; Hakuta, 1976; and Huebner, 1983). Based on these results, it was assumed that the individual learners adequately represented the interlanguage level with which they were associated.

BACKGROUND TO THE PRESENT STUDY

The present study reanalyzes the zero article data from Master (1987) in terms of both zero article ($\emptyset 1$) and null article ($\emptyset 2$) usage to determine if there is a distinction in the way the zero and null articles are acquired in spoken discourse. Because the null article occurs primarily with proper nouns, which were ignored in Master (1987) since the rules of article usage were thought at the time to be different for proper nouns, the gathered but rejected 1987 data on proper noun usage are included in the reanalysis. The present study constitutes a functional analysis of language development (see e.g., Andersen, 1990; Bates & MacWhinney, 1982; Correman & Kilborn, 1991; Dietrich, Klein, & Noyau, 1995; Klein, 1990; Pfaff, 1987; Tomlin, 1990) concerning an aspect of syntax that is, like word order, relative clauses, and verb phrases, subject to cross-linguistic influence (Odlin, 1989).

Because it was clear from Master (1987) that real progress on the article system did not occur until the low mesolang level and beyond, data from the

five basilarang (i.e., beginning level) learners in the 1987 study are ignored in the present study as the learners at this level appeared to be grappling with the grosser distinctions between \emptyset and *the* and could not be expected to have any sense of the subtle differences between zero and null (indeed, the bulk of the \emptyset usage at this level was probably by omission). With the deletion of the basilarang group, 15 learners remain, three for each of the five language groups. Although there are only 15 learners, which is too small a number to allow for anything but a descriptive statistical analysis, the present study provides data regarding the spoken acquisition of the two zero articles that may aid future research.

ANALYSIS OF ZERO AND NULL ARTICLE USAGE

One of the potential problems of including proper nouns in the data set is that, because their use is controlled by context rather than lexical and discourse rules (e.g., if you tell a story about a trip to various countries, you will probably use more proper nouns than if you tell a story about walking through the woods, presuming your companion is not a botanist), the occurrence of \emptyset is likely to be inconsistent. However, the percentage of undifferentiated \emptyset tokens in the data set is essentially the same (42.8% accuracy for the [-ART] group and 42.9% for the [+ART] group), so their comparison is justified, even though the percentage of required \emptyset 2 (null) with proper names was 92.8% for the [-ART] group and 80.3% for the [+ART] group.

The interlanguage data were analyzed in two ways: 1) all the learners in the [-ART] group vs. all the learners in the [+ART] group to compare the performance of the two groups, and 2) the three interlanguage levels of each language group considered separately. The articles were studied in terms of accuracy (i.e., the correct use of an article when an article was required, also known as *supplied in obligatory context*, or SOC) and usage (i.e., the use of any article in any context, which Master (1987) labeled *used in obligatory context*, or UOC, and used as a replacement for the more traditional target-like usage (TLU), that has been used to account for overusage of a particular morpheme; these distinctions are further discussed on pp. 11-13).

The following research questions guided the analysis:

1. What are the differences between the [-ART] learners vs. the [+ART] learners in terms of
 - a. the accuracy of \emptyset 1 and \emptyset 2 usage?
 - b. the overuse of \emptyset ?
 - c. the use of \emptyset 1 in indefinite contexts?
 - d. the use of \emptyset 2 with proper nouns?
2. What are the tendencies of each language group within the [-ART] and [+ART] categories in the spoken acquisition of \emptyset 1 and \emptyset 2 in terms of the accuracy of \emptyset 1 and \emptyset 2 usage?

It was expected that the [-ART] and [+ART] groups would show different acquisition patterns for zero and null.

[-ART] VS. [+ART] PERFORMANCE

The first analysis examined the 3,347 noun phrases produced by the [-ART] group, which consisted of nine learners, and compared it with the 1,848 noun phrases produced by the [+ART] group, which consisted of six learners.

Ø1 vs. Ø2 Accuracy

The accuracy of Ø1 usage for the [-ART] group was 88.6%, whereas the accuracy of Ø2 usage was 90.2%. The small difference (< 2%) suggests that the [-ART] speakers may not differentiate between the Ø1 and Ø2 articles.

The accuracy of Ø1 usage for the [+ART] group was 88.4%, whereas the accuracy of Ø2 usage was 94.3%. The slightly greater difference (almost 6%) between Ø1 and Ø2 accuracy for the [+ART] group could indicate that the [+ART] speakers differentiate to some extent between the Ø1 and Ø2 articles as they do in their first languages. However, the percentages suggest there is little differentiation between the two.

Indeed, upon closer analysis, it was found that it was the Spanish speakers (82.1%) rather than the German speakers (94.7%) that accounted for the lower figure. The Spanish learners in this study appear to have had more difficulty with the multi-layered nature of the article system (which simultaneously shows definiteness, countability, and number) and could not deal with all the layers at the same time. The low-mesolang Spanish speaker, for example, uttered noun phrases such as “I live on **a* four floor” and “We are having laundry in **a* same building” (showing correct number and countability but not definiteness) in the data set, even though the definite article would be used in this case in Spanish just as it is in English. Schumann (1978) also found a native speaker of Spanish to say in his low-proficiency English “*Picture is very dark” even though in Spanish (*El cuadro es muy oscuro*) the definite article would be present, while Andersen (1977) found a surprisingly low accuracy with the article *a* for some of his Spanish learners (in areas of contrast between English and Spanish) despite the fact that an indefinite article with a similar function exists in Spanish. The German learner at this level used more German vocabulary in her speech than did the Spanish learner but did not evince a similar strategy. However, I recently encountered the following sentence in a Danish student’s essay: “I was once a substitute in a first grade.” which suggests that the strategy is not limited to Spanish speakers. This usage by [+ART] speakers may be an example of what Zobl (1982) described as his developmental complexity restraint, in which “L1 influence may modify a developmental continuum at that point at which a developmental structure is similar to a corresponding L1 structure and where further progress in the continuum amounts to an increase in complexity beyond that of the L1 structure” (p. 180), leading to an extension of the scope of the current developmental structure. In general, however, he found that [+ART] speakers achieve target-like control more quickly than [-ART] speakers, as did Master (1987).

In any event, barring this effect in the Spanish learners, there is little difference between the accuracy of Ø1 and Ø2 performance in either the [+ART] group or the [-ART] group.

Ø “Overuse”

Ø “overuse,” the use of zero or null in contexts where another article is required, is considered first in definite and then in indefinite contexts.

Ø Overuse in Definite Contexts

The [-ART] group used Ø in obligatory contexts requiring *the* 19.4% of the time, whereas the [+ART] group did the same only 9.5% of the time. Table 3 shows the breakdown in specific common noun contexts that require *the* in English. The greatest difficulty for the [-ART] group was with implied postmodification (e.g., “after swimming I go to *office”) and shared knowledge (e.g., “I read *newspaper”), whereas the least difficulty was with subsequent mention (“when I worked in *military college”), ranking adjectives (e.g., “*next state is Chicago”), and postmodified NPs (e.g., “if I use *Japanese standard of marriage”).

Table 3. Incorrect Use of Ø1 in Contexts Requiring *the* (in % errors)

Category	[-ART n = 162]	[+ART n = 46]	[±ART n = 208]
Subsequent Mention	11.6	8.7	11.0
Shared Knowledge (S.K.)	20.2	13.0	18.7
Ranking Adjectives	11.6	23.9	14.2
Postmodified NP	10.4	21.7	12.8
Implied Postmodification	27.2	15.2	24.7
S. K. of Uniqueness	19.1	17.4	18.7
Total	100.1	99.9	100.1

In other words, article rules of thumb (ignoring the exceptions) such as “Always use *the* when an NP is referred to again, when it occurs with a ranking adjective, and when a noun is postmodified” appear to be easier for [-ART] learners to apply than rules for situations in which article usage is determined by context, such as implied postmodification and shared knowledge. This finding is supported by a comment from a Japanese nonnative speaker of English cited in Master (1995):

[Y]our explanation of the articles that are not influenced by contexts (*the* next or *the* following) were very helpful and I always remember those rules when I write. However, the articles which are determined by the contexts in which they are in are still hard for me to distinguish because I am not competent enough to apply your discussion to other contexts (p. 198).

For the [+ART] group, on the other hand, the greatest difficulty was ranking adjectives (e.g., “it is *same to my country”) and postmodified NPs (e.g., “one of

*big rivers around the world”), whereas the least difficulty was subsequent mention (e.g., “Gretel is in *house”) and shared knowledge (e.g., “I am go to *opera”). The problems with ranking adjectives may have been due to the interference phenomenon mentioned earlier as ranking adjectives also take the definite article equivalent of *the* in German and Spanish. On the other hand, the relatively fewer errors with subsequent mention and shared knowledge probably result from the transference of a more general principle in the L1, which also applies in English: mark nouns that are known through prior mention or shared knowledge with *the*.

∅ Overuse in Indefinite Contexts

∅1 was used correctly 88.6% of the time by the [-ART] group and 88.4% of the time by the [+ART] group. The breakdown into categories is shown in Table 4. Table 4 shows that the [+ART] group was more accurate in the use of *a* with singular count nouns, making less than half the ∅1-for-*a* errors (15.1%) of the [-ART] group (34.2%).

Table 4. Accuracy of ∅1 in Specific Indefinite Contexts (in %)

Category	Used	Required	[-ART]	[+ART]	[+ART]
Singular Count	∅1	<i>a</i>	34.2	15.1	27.6
Plural Count	∅1	∅1	91.2	87.8	90.1
Noncount	∅1	∅1	86.5	89.4	87.5
<i>a = one</i>	∅1	<i>a</i>	7.1	0	3.1

Note: Italicized percentages represent correct usage; percentages in plain type represent % errors.

The [+ART] group had slightly more difficulty with plural count nouns, whereas the [-ART] group had slightly more difficulty with noncount nouns. The latter result is supported by Hiki (1989), who found that Japanese students had greater difficulty assigning the correct article to noncount (his term was nonbounded), particularly abstract, nouns.

∅2 with Proper Nouns

The accuracy of ∅2 + proper noun use was quite high for both groups, with 90.2% accuracy for the [-ART] group and 94.3% accuracy for the [+ART] group. The high percentage for the [-ART] group was no doubt partially due to the high *used in obligatory context* (UOC) figure for this group (see below). However, the [+ART] group used ∅2 incorrectly with proper nouns requiring *the* more frequently (27.0% of ∅-for-*the* usage) than did the [-ART] group (19.6% of ∅-for-*the* usage). This may have been due to the confusion arising from the fact that some proper nouns with ∅2 in English take the definite article in Spanish and German (e.g., Peru is *El Peru* with the definite article in Spanish; Switzerland is *Der Schweiz* with the definite article in German).

The categories of proper nouns in which ∅2 was used when *the* was required

are shown in Table 5. Both groups had the greatest number of errors with political divisions, which includes the names of countries, states, cities, etc. However, every error that the figures represent was either Ø2 + United States (20 errors) or Ø2 + Soviet Union (3 errors).

Table 5. Incorrect Use of Ø2 with Proper Nouns Requiring *the* (in %errors)

Category	[-ART] n = 44	[+ART] n = 17	[±ART] n = 61
Business Names	27.3	11.8	23.0
Political Divisions	31.8	52.9	37.8
Geographical Names	11.4	5.9	9.8
Persons	4.5	11.8	6.6
Time	6.8	0	4.8
Other	18.2	17.6	18.0
Total	100	100	100

Many learners overgeneralize the name of a political division such as Ø2 + *Panama* to the title of a political division such as *the* + *Republic of Panama*. Thus, many learners say *United States* instead of *the United States*. The problem is no doubt exacerbated by the fact that the full title (*The United States of America*) is less commonly heard than *the U.S.*, *the States*, or *America* and also that the adjective form, as in *United States* (or more likely *U.S.*) *policy*, does not take *the* unless the headnoun requires it (e.g., *the U.S. Post Office*).

The most common correct use of Ø2 with common nouns was after prepositions (e.g., *at home*, *in class*) and with sequence adjectives showing relationship to the present (e.g., *next week*, *last year*). These uses both lie within the “familiar” category of Ø2 usage. The “name” type, e.g., *He was named captain of the team*, did not occur in the data set, which is not surprising given its rather limited utility.

ACQUISITION TENDENCIES IN THE TWO GROUPS

The second analysis examined each language group within the [-ART] and [+ART] language groups for potential tendencies in the acquisition of the zero and null articles.

Undifferentiated Ø UOC and SOC for Each Specific Language Group

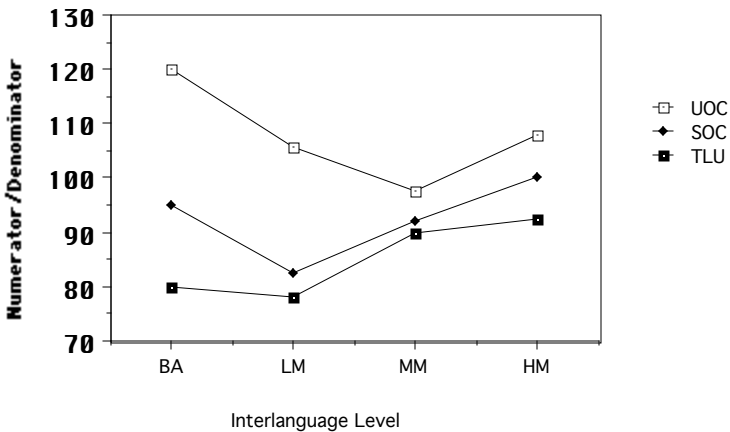
TLU has been used in interlanguage studies as a more realistic alternative to report simple accuracy of use than SOC. The TLU measure attempts to rectify any inflated indication of accuracy of use by adding the number of uses of a morpheme supplied outside an obligatory context to the **denominator**, effectively reducing the accuracy rate (e.g., $5/10 = 50\%$; $5/10+1 = 45\%$). Without this adjustment, the SOC measure only applies to performance within obligatory contexts. For example, *the*

is required in front of superlative adjectives; thus, any use of a superlative is an obligatory context for *the*. If one looked only at the use of articles with superlatives, the SOC score would accurately reflect performance within this context, but if the learner mistakenly used *the* in front of comparative structures as well (e.g., “He bought *the smaller house than I did”), SOC would ignore such errors. TLU, on the other hand, acknowledges errors such as these, though it does not differentiate between what is accurate within obligatory contexts and what occurred outside obligatory contexts.

The UOC measure was described in Master (1987) as an alternative to the TLU measure. The UOC measure adds the number of morphemes supplied outside an obligatory context to the **numerator** (hence the term *used in obligatory context* as opposed to “[correctly] supplied in obligatory context”) and is presented alongside accuracy, which may be directly compared because SOC and UOC share the same denominator (number of obligatory contexts). It is thus a little more precise than TLU, although it is still only an approximation. A native speaker has a UOC of 100%, i.e., uses a morpheme whenever it is required, as well as an SOC of 100%. A comparison of TLU, SOC, and UOC is shown in Figure 1, which shows the Spanish learners’ use of *the*.

A high UOC figure suggests that an article is frequently being used outside of obligatory contexts and is therefore not under control to the extent that a lower UOC figure would suggest.

Fig. 1 Sample Comparison of UOC, SOC, and TLU



Note: UOC=used in obligatory content; SOC=supplied in obligatory context; TLU=target-like usage; BA=basilang; LM=low mesolang; MM=mid mesolang; HM=high mesolang

Keeping in mind that the ideal SOC and UOC is 100%, Figure 1 shows that accuracy (SOC) for these learners was quite high at the BA level, diminished at the LM level, but then rose at the MM level to become almost 100% at the HM level. The UOC, on the other hand, is quite high at the BA level, suggesting that part of the reason for the high SOC number is that the learner was using *the* in many contexts where it was not required (a characteristic of BA [+ART] speakers), and thus was not the result of high control of the article. There is still some overusage (UOC) at the HM level for these learners, but the distance between UOP and SOC is much less, suggesting greater control of the article. The TLU numbers, in comparison, which have been conventionally reported in place of SOC, do not provide nearly so much information; they merely represent a “penalized” accuracy figure. This is why I made the decision to replace the TLU with the SOC/UOC figures.

UOC for the [-ART] group was considerably higher (120.3%) than for the [+ART] group (105.9%), presumably because the [+ART] group has an article system and thus some notion of the differentiation between *the* and \emptyset . The [-ART] group, on the other hand, often used \emptyset in place of any article at all (a possible transfer from L1), which accounts for the higher figure. UOC can only be supplied for the undifferentiated \emptyset article because there is no way to label the article out of context as $\emptyset 1$ or $\emptyset 2$.

Article UOC for each specific language group is shown in Table 6. Table 6 shows that for all learners but the Japanese LM³, \emptyset UOC decreases with increasing IL level (though this is not true of *the* and *a*). This consistency suggests that increasing proficiency is associated with increasing control (or perhaps “reining in”) of the zero article (See Table 6). Table 6 also shows that the [+ART] group overuses \emptyset to a substantially lesser extent (105.9%) than the [-ART] group (120.3%). Note also that, while *the* and *a* are not the focus of this study, the SOC and UOC figures are closer for *the*, and even closer for *a*. The latter finding suggests that *a* is not “experimented with” to the extent of *the* and \emptyset , perhaps because learners tend not to use *a* until they have sufficiently acquired the notion of countability, as suggested in Master (1987).

$\emptyset 1$ vs. $\emptyset 2$ Accuracy for Each Specific Language Group

Accuracy (SOC) for each L1 group in the use of the zero and null articles is shown in Table 7. The first two columns indicate $\emptyset 1$ and $\emptyset 2$ accuracy while the third ($\emptyset 2$ PNU and $\emptyset 2$ PNR) indicates $\emptyset 2$ accuracy with proper nouns, which account for over 98% of $\emptyset 2$ usage for the entire group.

Table 7 shows that $\emptyset 1$ accuracy increases continuously for the Chinese learners, but not for the rest of the learners. For $\emptyset 2$ accuracy, the picture is somewhat different. Both the [-ART] and the [+ART] groups generally show continuous increasing $\emptyset 2$ accuracy across IL levels. The exception is the Chinese HM learner,

Table 6. UOC (Used) vs. SOC (Accurately Supplied) in All Five Language Groups

Chinese	N	ØU	ØR	ØSOC%	ØUOC	ØUOC%	theU	theR	theSOC%	theUOC	theUOC%	aU	aR	aSOC%	aUOC	aUOC%
LM	388	88	104	84.6	169	162.5	132	174	75.8	155	89.1	60	110	54.5	64	58.2
MM	348	161	179	89.9	213	119.0	82	119	68.9	108	90.8	22	50	44.0	23	46.0
HM	319	99	110	90.0	116	105.5	132	141	93.6	145	103.8	48	60	80.0	50	83.3
<u>Japanese</u>																
LM	374	46	111	41.4	86	77.5	144	165	87.3	233	141.2	31	94	33.0	44	46.8
MM	327	158	163	96.9	213	130.7	72	96	75.0	81	84.4	31	68	45.6	31	45.6
HM	389	164	171	95.9	216	126.3	88	109	80.7	94	86.2	76	109	69.7	79	72.5
<u>Russian</u>																
LM	340	158	175	90.3	241	137.7	60	109	55.0	88	80.7	6	56	10.7	6	10.7
MM	424	162	177	91.5	201	113.6	129	146	88.4	146	100	75	101	74.3	77	76.2
HM	438	240	242	99.2	267	110.3	85	103	82.5	87	84.5	83	93	89.2	84	90.3
[E-ART]	3347	1276	1432	89.1	1722	120.3	924	1162	79.5	1137	97.8	432	741	58.3	458	61.8
<u>German</u>																
LM	183	63	66	95.5	84	127.3	62	74	83.8	65	87.8	33	43	76.7	33	76.7
MM	428	163	181	90.1	189	104.4	129	153	84.3	141	92.1	87	94	92.6	95	101.1
HM	315	127	128	99.2	131	102.3	103	104	99.0	104	100	79	83	95.2	79	100
<u>Spanish</u>																
LM	428	177	216	81.9	226	104.6	95	116	81.9	126	108.6	56	96	58.3	75	78.1
MM	318	137	145	94.5	151	104.1	110	125	88.0	115	92.0	46	48	95.8	51	106.3
HM	176	51	56	91.1	58	103.6	92	93	98.9	99	106.5	15	27	55.5	16	59.3
[E-ART]	1848	718	792	90.7	839	105.9	591	665	88.9	650	97.7	316	391	80.8	349	89.3
[E-ART]	1595	1503	2224	89.7	2561	117.8	1361	1656	82.9	1787	97.8	748	1132	66.1	807	71.3

Note: UOC=used in obligatory context; SOC=supplied in obligatory context; N=number of occurrences; U=used; R=required; LM=low mesolang; MM=mid mesolang; HM=high mesolang

Table 7. \emptyset 1 vs. \emptyset 2 Accuracy for Each Specific Language Group

L1	IL	\emptyset 1U	\emptyset 1R	%	\emptyset 2U	\emptyset 2R	%	\emptyset 2PNU	\emptyset PNR	%
Chinese	LM	63	75	84.0	24	28	85.7	21	25	84.0
	MM	91	105	86.7	70	74	94.6	62	66	93.9
	HM	62	63	98.4	41	51	80.4	39	49	79.6
Japanese	LM	26	79	32.9	20	32	62.5	20	32	62.5
	MM	112	116	96.6	46	47	97.9	36	36	100
	HM	133	140	95.0	31	31	100	27	27	100
Russian	LM	64	69	92.8	94	106	88.7	93	105	88.6
	MM	103	116	88.8	59	61	96.7	56	58	96.6
	HM	1209	211	99.1	31	31	100	30	30	100
SubT [-ART]	863	974	88.6	416	461	90.2	384	428	89.7	
German	LM	52	53	98.1	11	13	84.6	11	13	84.6
	MM	76	87	87.4	87	94	92.6	84	86	97.7
	HM	105	106	99.1	22	22	100	18	18	100
Spanish	LM	72	104	69.2	105	112	93.8	65	70	92.9
	MM	93	100	93.0	44	45	97.8	41	42	97.6
	HM	37	42	88.1	14	14	100	12	12	100
SubT [+ART]	435	492	88.4	283	300	94.3	231	241	95.9	
Total [\pmART]	1298	1466	88.5	699	761	91.9	615	669	91.9	

Note: IL=interlanguage level; LM=low mesolang; MM=mid mesolang; HM=high mesolang; U=used; R=required; PNU=proper nouns used; PNR=proper nouns required; Bold=percentages to be compared.

who showed decreased accuracy at the HM level.

A closer analysis of Ø2 usage by the Chinese HM learner shows an interesting phenomenon. In 49 uses of Ø2 usage with proper nouns, the HM learner got 39 correct. The remaining ten errors are as follows:

NP#	Page-Line	Context
70	5-4	during <u>the</u> Christmas
72	5-6	in <u>the</u> California
96	5-50	they lefted <u>the</u> Taiwan
102	6-5	to invasion-- <u>the</u> southeast China
122	6-42	<u>the</u> mainland China
202	10-12	in <u>the</u> Singapore
216	10-38	in <u>the</u> Indonesia
305	16-26	like <u>the</u> Monterey Park

The patterns of occurrence of the proper NPs *the Taiwan* and *the China* are shown in Tables 8 and 9.

Table 8. Pattern of Occurrence of *Taiwan* in the Chinese HM Data

NP	#Page-Line	Context
2	1-6	I from Taiwan
44	2-52	in Taiwan
91	5-46	they used Taiwan as a base
96	5-50	they lefted the Taiwan
100	6-4	they want to use Taiwan
104	6-6	in Taiwan
123	6-43	Taiwan is a part of China
221	10-43	with Taiwan

Table 8 shows that the Chinese HM learner used Ø2 + *Taiwan* correctly four times prior and four times after the erroneous usage. Table 9 shows that the Chinese HM learner used *the* + *China* incorrectly twice, once just before using it correctly twice. Both of the incorrect instances are premodified (with *southeast* and then *mainland*), which may have caused the speaker to perceive the noun as being definite (or “identified”; see Master, 1990). This pattern is almost immediately broken, however, by the correct use of Ø2 with a premodified NP in NP #128.

One explanation, which is consistent with the finding that learners appear to make little distinction between the zero and null articles, is that the Chinese HM learner may have overgeneralized the subsequent mention rule for common nouns (i.e., mark all subsequent mentions of the same referent with *the*) to proper nouns, but then failed to apply the rule consistently. In NP #122, the speaker used *the* + *mainland China* and then in NP #123 used Ø + *China* and then Ø + *mainland China* shortly thereafter in NP #128. There is then a considerable gap (approximately four pages of dialog) before *the* + *China* occurs again. As if introducing the NP after such a long hiatus, two NPs are marked with *the*, followed immediately by three uses in fairly close succession that are correctly

Table 9. Pattern of Occurrence of *China* in the Chinese HM Data

NP	#Page-Line	Context
102	6-5	to invasion--the southeast China
122	6-42	the mainland China
123	6-43	Taiwan is a part of China
128	6-51	are from mainland China
205	10-22	don't want to offend the China
211	10-32	they will offend the China
212	10-33	China will cause some agitation
214	10-37	China launched a coup d'etat
222	10-44	that will offend China

marked with \emptyset . In Table 8, we can see a similar pattern in that NP #96 with *the* is followed by the use of \emptyset with the same proper noun four NPs later. In this case, however, there are earlier uses of the same NP with \emptyset . Perhaps the speaker has forgotten to apply the subsequent mention rule, thus appearing to use \emptyset to mark subsequent mentions. Master (1995) found the following in this regard:

One example of what many advanced students with no article system in their L1s appear to do is to apply the correct shared knowledge article rules early in a paragraph but then fail to maintain the marking of the shared knowledge NP throughout the paragraph as English requires, perhaps reasoning that the reader must know by now what the writer is talking about so why should it require constant remarking? (p. 198)

On the other hand, there may be an interlanguage rule operating here that deliberately marks subsequent mentions with zero. Indeed, Mellow, Reeder, and Forster (1996) found that the zero article is often used to mark known NPs in the discourse of [-ART] speakers. Furthermore, the pattern of marking a subsequent mention of a definite NP with \emptyset is not limited to English/[-ART] interlanguages; it also occurs in standard English when two or more definite NPs (marked with *the* or a possessive determiner) are followed by a unified subsequent mention as \emptyset + NP. This unified structure has been called a zero *and*-coordinated noun phrase (Master, 1988) and is shown in the following example:

The cowboy and his horse galloped off into the sunset. (\emptyset 2) Horse and rider were later found dead by a poisoned well (Master, 1988, p. 174).

The pattern occurs relatively rarely and primarily in literary texts, but it does exist.

It is difficult to determine for sure what the speaker's rules for \emptyset 2 usage are, and it may be that the alternation is simply a product of IL variation that eventually gives way to the correct form, much as Huebner (1983) showed the use of *she* as an

object pronoun to gradually diminish over time in the interlanguage of his subject, Ge. However, it does show an interesting confusion of the two definite markers (*the* and null). It also suggests that the article rules for common nouns are not kept neatly separated from those for proper nouns and that the entire system must be considered in the developing interlanguage. Incorporation of the null article would allow article researchers to do this.

CONCLUSION

A few tentative conclusions may be drawn from the present study. For the most part, while [+ART] speakers have an advantage over [-ART] speakers in using the article system as a whole, the tendency of the latter to overuse \emptyset appears to serve them well in getting zero and null usage correct, with the caveat that there is no *a priori* mechanism for determining whether \emptyset was used deliberately or whether the article was simply deleted by mistake.

With \emptyset 1 usage, the [-ART] group had more errors with implied postmodification and shared knowledge, as was also found in (Master, 1995). The [+ART] group, on the other hand, had greater problems with ranking adjectives and postmodified nouns. This may be due to the layering of functions that article usage requires, causing difficulties when the same structure occurs in the L1. \emptyset 1 accuracy alone does not increase regularly by IL level in this study. \emptyset 2 accuracy, on the other hand, does show continuous increase by IL level (with the exception of the Chinese HM learner as explained).

A further conclusion is that articles with proper nouns, which are usually ignored in interlanguage article analyses because they are thought to be arbitrary (e.g., \emptyset 2 occurs with *parks*, while *the* occurs with *rivers*), can be included in a principled fashion by means of the null article. In this regard, Chesterman (1991), arguing that \emptyset 2 occurs with *parks* because they have clear natural boundaries, while *the* is required with *rivers* because they do not (i.e., rivers have a point of origin but sometimes varying courses of flow and vague mouths and deltas), said, "It is surely a weakness of [the] standard description that proper nouns are *not* incorporated into the description in any systematic way" (p. 7, emphasis in original).

The existence of the zero and null articles with their diametrically opposed significations adds yet another layer of difficulty in the acquisition of the article system, especially for [-ART] learners of English. However, the lack of a marked distinction between either the accuracy (SOC) of the two forms or the differential performance of the [\pm ART] groups analyzed in this study raises the possibility that the null and zero articles were learned 1) as a single application (e.g., use \emptyset with noncount and plural count nouns, both common and proper) rather than as a rule-governed system with differential applications for \emptyset 1 and \emptyset 2, 2) as lexical items, i.e., learned as chunks or lexical phrases, as was also suggested in Master (1995), or 3) as a mixture of the two. The confusion that the HM Chinese learner evinced may be the product of an intersection between lexical knowledge (e.g., known chunks such as *Taiwan* and *China*) and other rules for article selection (e.g., subsequent

mention). On the other hand, it may be an indication of the dawning (unconscious) realization that Ø functions as both an indefinite and a definite marker.

While the majority of null usage is in fact removed by the general tendency to ignore proper nouns in article research, without null there will always be some uses of the zero article with common nouns that are misclassified. It is therefore hoped that article researchers will take the zero-null distinction into account in future research.

ENDNOTES

¹Despite the recency of this description, the distinction was actually first proposed by Palmer (1939).

²In the 5736 instances of the undifferentiated zero article in the 16 research articles mentioned in Master (1997, Table 2, p. 221), 51.7% occurred with plural count nouns, 40.5% occurred with noncount nouns, and 7.9% occurred with singular count nouns).

³The Japanese LM learner flooded heavily with *the* (*the* UOC = 141.2%, almost 50% more than any other learner), which brought about a concomitant decrease in Ø UOC.

REFERENCES

- Andersen, R. (1977). The impoverished state of cross-sectional morpheme acquisition/accuracy methodology (Or: The leftovers are more nourishing than the main course). *Proceedings of the Los Angeles Second Language Research Forum*. Los Angeles: Department of English, University of California at Los Angeles.
- Andersen, R. (1990). Models, processes, principles and strategies: Second language acquisition inside and outside the classroom. In B. VanPatten & J. Lee (Eds.), *Second language acquisition/Foreign language learning* (pp. 45-68). Philadelphia: Multilingual Matters.
- Bates, E., & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 173-218). Cambridge: Cambridge University Press.
- Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.
- Cazden, C., Cancino, E., Rosansky, E., & Schumann, J. (1975). *Second language acquisition sequences in children, adolescents, and adults*. Washington, D.C.: Final report submitted to the National Institute of Education.
- Chesterman, A. (1991). *On definiteness: A study with special reference to English and Finnish*. Cambridge, UK: Cambridge University Press.
- Correman, A., & Kilborn, K. (1991). Functionalist linguistics: Discourse structure and language processing in second language acquisition. In T. Huebner & C.A. Ferguson (Eds.), *Crosscurrents in second language acquisition and linguistic theories* (pp. 195-224). Amsterdam: John Benjamins.
- Dietrich, R., Klein, W., & Noyau, C. (1995). *The acquisition of temporality in a second language*. Amsterdam: John Benjamins.
- Dulay, H. & Burt, M. (1974). Natural sequences in child second language acquisition. *Language Learning*, 24, 37-53.
- Greenbaum, S., & Quirk, R. (1990). *A student's grammar of the English language*. Harlow, UK: Longman.

- Hakuta, K. (1976). A case study of a Japanese child learning English as a second language. *Language Learning*, 26, 321-351.
- Hiki, M. (1989). *The acquisition of the indefinite article and zero article by advanced Japanese EFL learners: The acquisition of the concept of countability*. Manuscript.
- Huebner, T. (1983). *A longitudinal analysis of the acquisition of English*. Ann Arbor, MI: Karoma Publishers.
- Klein, W. (1990). A theory of language acquisition is not so easy. *Studies in Second Language Acquisition*, 12, 219-231.
- Master, P. (1987). *A cross-sectional interlanguage analysis of the acquisition of the English article system*. Unpublished doctoral dissertation, UCLA.
- Master, P. (1988). X-bar syntax and the unity of and-coordinated noun phrases, *Proceedings of the Western Conference on Linguistics, Vol. 1*, Fresno, CA: California State University, 174-184.
- Master, P. (1990). Teaching the English articles as a binary system. *TESOL Quarterly*, 24, 461-478.
- Master, P. (1995). Consciousness raising and article pedagogy. In D. Belcher & G. Braine (Eds.), *Academic writing in a second language* (pp. 183-204). Norwood, NJ: Ablex.
- Master, P. (1997). The English article system: Acquisition, function, and pedagogy. *System*, 25, 215-232.
- Mellow, J. D., Reeder, K., & Forster, E. (1996). Using time-series research designs to investigate the effects of instruction on SLA. *Studies in Second Language Acquisition* 18, 325-350.
- Odlin, T. (1989). *Language transfer: Cross linguistic influence in language learning*. Cambridge: Cambridge University Press.
- Oller, J. & Redding, E. Z. (1972). Article usage and other language skills. *Language Learning*, 21, 85-95.
- Palmer, H. E. (1939). *A grammar of spoken English on a strictly phonetic basis* (2nd ed.) Cambridge: Heffer.
- Pfaff, C. (1987). Functional approaches to interlanguage. In C. Pfaff (Ed.), *First and second language acquisition processes* (pp. 81-102). New York: Newbury House.
- Schumann, J. (1978). *The pidginization process: A model for second language acquisition*. Rowley, MA: Newbury House.
- Tomlin, R. (1990). Functionalism in second language acquisition. *Studies in Second Language Acquisition*, 12, 155-177.
- Zobl, H. (1982). A direction for contrastive analysis: The comparative study of developmental sequences. *TESOL Quarterly*, 16, 169-83.

Peter Master is Professor Emeritus in the Department of Linguistics and Language Development at San José State University in San José, California. His primary research focus has been the English article system. Further information can be found at his website under "The English Article System" at <<http://www2.sjsu.edu/faculty/pmaster/>>.