



## **Article Errors in the English Writing of Advanced L1 Arabic Learners: The Role of Transfer**

**Peter Crompton**

*American University of Sharjah*

### **Bio Data:**

Peter Crompton has been teaching at AUS since 2006. He has previously taught in universities in England, China, Saudi Arabia, Brunei, and Lithuania. He holds a PhD in Linguistics from Lancaster University. He teaches English for academic purposes and applied linguistics and his research interests are written discourse analysis, corpus linguistics, learner corpora, and academic writing.

### **Abstract**

The problems encountered by English learners whose mother tongue does not have an article system have been researched extensively. The problems encountered by learners whose mother tongue does have an article system, such as Arabic, have been studied less. This article describes an enquiry into article system errors in a corpus of English writing by tertiary-level L1 Arabic speakers. Frequencies of articles are compared with those in native English and non-native English speaker corpora. A detailed account is given of the commonest types of errors, classified according to the mis-use of each article. It is found that the commonest errors involve mis-use of the definite article for generic reference. The strong likelihood that these errors are caused by L1 transfer, rather than an interlanguage developmental order, is argued by a comparison of the forms of generic reference in English and Arabic. It is suggested that even for learners of English with mother-tongues which have an article system, such as Arabic, L1 transfer may be a problem and as such could be usefully addressed in language instruction.

**Keywords:** article error, Arabic learners, generic reference, transfer, English writing

### **Introduction**

The paper describes research into the scale and nature of article system errors made by tertiary-level L1 Arabic learners of English in a corpus of written text. It begins with reviews of research on the role of transfer in language acquisition in

general and in article acquisition in particular. It then reviews the literature on errors in article use by L1 Arabic learners, reports some corpus-based research by the author which brings together both acquisition and use perspectives, discusses the findings, and suggests some implications for future research and pedagogy.

As the paper gives considerable attention to *generic reference*, that term requires clarification. In the literature not all NPs are described as referential nor all non-generic reference as specific reference (e.g. Huebner, 1983; Thomas, 1989). For simplicity's sake, however, this paper follows the practice of Quirk, Greenbaum, Leech and Svartvik (1985) in treating all NPs as referential and using the term *specific* as a synonym for *non-generic*.

#### *Status of transfer studies*

Corpus studies show that *the* is the commonest word in English and *a* in the top five (Master, 2002) and together *the* and *a* make up 8.5% of all English text (Berry, 1993).<sup>1</sup> Given both the ubiquity of the article system and the comparatively long history of ELT in many Arabic speaking countries, surprisingly little has been written on English article system errors made by L1 Arabic speakers. Overall, there is less literature on errors by learners categorized on the basis of their L1 than one might expect: this can be attributed to a temporary eclipse of contrastive studies and error analysis (James, 1998; Mohammed, 2000; Odlin, 1989). In brief, the dominant view in SLA research for many years was that second and first language acquisition were essentially similar. This view, based on studies into first language acquisition (Brown, 1973), child second language acquisition (Dulay & Burt, 1974) and adult second language acquisition (White, 1977), explicitly opposed the contrastive approach (Lado, 1957) of the 1960s. As summarized by Dulay, Burt and Krashen (1982, p. 198) the “great majority of grammatical errors found in the language output of L2 learners is similar to those made by L1 learners of the target language rather than to the structure of the L2 learner’s mother tongue.” Interestingly for our purposes, the article system was one of the morphemes analyzed in these studies. Dulay and Burt diagnosed only one of the 74 article system errors in their data as “interlingual”, i.e. originating from L1; White (1977), for example, found that of the 31 article system errors in their data none was interlingual. The proper focus of such study were errors

conceived of as “intralingual” or developmental errors, evidence of a process of creative construction in a learner’s interlanguage. In the light of this, the term “error” in itself was later conceived of as unhelpful and misleading, in that what teachers registered as errors in target language production were in fact system features of learner interlanguage: to describe such features as errors was to commit the “comparative fallacy” (Bley-Vroman, 1983) of describing any language as a deficient form of another.

However, as noted in James (1998), approaches to analysis of learner language which expected and looked for errors arising from L1 influence survived, albeit rebranded slightly, as studies of cross-linguistic or language *transfer* (Odlin, 1989). As James (1998) argues, the concept of transfer now enjoys “renewed acceptance as a crucial component in modern L2 learning theories.” In a recent review of the evidence on transfer, Ellis (2006b) goes further than most in arguing for the essential correctness of the position originally taken by Lado’s (1957) contrastive analysis hypothesis; Ellis marshals evidence from a meta-analysis of SLA studies on morpheme acquisition and “associative and connectionist theories of animal and human learning” to argue that “second language learning is...faced with maximal transfer and interference from L1” (2006b, p. 186). Indeed, all the literature on article system acquisition in the last twenty years recognizes L1 transfer as a variable at least worthy of investigating (e.g. Butler, 2002; Jarvis, 2002; Master, 1990, 1997; Thomas, 1989; Zdorenko & Paradis, 2008; Zegezac, 2004).

#### *Studies of English article acquisition*

In English the article system is a marker of definiteness. The extensive literature on definiteness is usefully summarized in Lyons’ (1999) cross-linguistic study. According to Lyons, definiteness is not a semantic concept but a grammaticalization of the semantic/pragmatic concept of identifiability: a marker of definiteness, such as a definite article, “directs the hearer to the referent of the noun phrase by signaling that he is in a position to identify it” (5-6). Like other grammatical categories there is not a simple one-to-one meaning-form correspondence. Definiteness is manifested in other forms, such as pronouns and proper nouns, and diachronically the prototypical (identifiability) meaning can be both added to and subtracted from. As Lyons notes,

“some languages will require generics to be definite while others do not” (1999, p. 278). Arabic, like French, is in the former class, while English is in the latter. In the course of the research described in this paper it becomes apparent that this particular divergence appears to have serious consequences for Arabic speakers learning English.

Unfortunately, much of the language acquisition literature on articles offers little of direct interest to those investigating L1 Arabic learners’ acquisition of English, for two main reasons: (i) the broad-grained nature of their specification of the L1 from which transfer is analysed; (ii) the broad-grained view nature of their specification of the functions of the article system.

Turning to broad-grained L1 formal specification first, in general, SLA research on English article acquisition has divided L1s into a simple binary typology, languages which have an article system [+article] and those which do not [-article]. Arabic is a [+article] language but unlike English in that there are definite articles but no indefinite articles (Schulz, 2004). In general, researchers have found that subjects with [-article] L1s, such as Japanese and Russian, take longer to acquire the system than those with [+article] L1s (Chaudron and Parker, 1990; Liu & Gleason, 2002; Master, 1997; Thomas, 1989). Using a corpus-based study of the morphemes in Dulay and Burt’s (1973) “morpheme order” studies in the Japanese component of the Longman Learners Corpus McEnery, Xiao, and Tono, (2006) found that articles were the most difficult to acquire, with even proficient learners not having achieved the acquisition rate of 90% specified in Dulay and Burt’s Bilingual Syntax Measure. This differs from the order predicted by Dulay and Burt (1973), prompting McEnery et al to speculate that this is because Japanese is a [-article] language. Ionin, Ko, and Wexler (2004) and Zdorenko and Paradis (2008) use a further binary distinction, based on a Universal Grammar principles and parameters model, according to which [+article] languages may be further categorized into languages in which the article system marks the semantic category of definiteness (e.g. English), and languages in which the system marks the category of specificity (e.g. Samoan). However, their research focuses on comparative error rates for learners from [-article] and [+article] L1s rather than types of errors made by learners from [+article] L1s. Two much-cited longitudinal studies into morpheme acquisition which included articles (Huebner,

1983; Parrish, 1987) also had subjects with East Asian [-article] L1s, limiting generalizability of findings to Arabic – unless a universal order of acquisition is assumed.

Turning to broad-grained functional analysis of article use next, the instruments used to elicit data, whether obligatory occasion or discursive tasks, are based on narratives (Chaudron & Parker, 1990; Jarvis, 2002; Liu & Gleason, 2002; Thomas, 1989) in which referents are concrete rather than abstract and specific rather than generic. Ionin, Ko and Wexler (2004) and Zdorenko and Paradis (2008) reduce options further by setting up contexts in which  $\emptyset$  is never accurate. This has helped to focus on basic semantic features of the article system, such as previous mention, of interest in the early levels of language acquisition, but reduces the applicability of their findings on overall accuracy for later levels.

In general, for the kind of errors made by intermediate to advanced learners a more fine-grained analysis of the functions of the articles is required. Rather than merely recording correct and incorrect instances of article use, some analysis of the contexts in which errors occur is required. One such study is by Liu and Gleason (2002), who note that previous studies have neglected to investigate the fact that “certain uses of [non-generic] the might be more difficult than others” (5), and find evidence in their own data for an order of acquisition for these uses. Their data was mostly from subjects with East Asian [-article] L1s – so its applicability to learners from other L1 backgrounds is an open question. A further gap in the data used in the literature, noted by Thomas (1989), is that relating to use of the articles for generic referents.

Two somewhat more fine-grained discourse-oriented studies, Robertson (2000) and Jarvis (2002), looked for systematicity in the errors learners made and found it; some of this systematicity was attributable, as expected, to discourse factors and some to linguistic context, i.e. variation in the syntactic forms of the host NP. However, both researcher also found evidence for “unsystematic variation” (Robertson, 2000, p. 135), i.e. individuals didn’t always make the same article choices in the same contexts. Jarvis attributes this to individual learners simultaneously entertaining “multiple hypotheses about article use” (Jarvis, 2002, p. 416). Robertson posits a “remapping principle” according to which learners have to effect an adjustment from

(a) how semantic and pragmatic features are mapped onto syntactic and lexical resources in L1 to (b) how this mapping is done in L2: the variation is a reflex of the difficulty of this task. Robertsons's subjects were L1 Chinese, and Jarvis's L1 Finnish and L1 Swedish.

#### *Studies of Arabic ESL learners' article errors*

Observations on likely errors are provided in collections by experienced ESL teachers (Kharma & Hajjaj, 1997; Smith, 2001; Thompson-Panos & Thomas-Ruzic, 1983). More formal investigations are described in Scott and Tucker (1974), Kharma (1981) and Bataineh (2005), and a study of translator errors appears in Farghal and al-Zou'bi (2004). Scott and Tucker (1974) found article errors among the top four types of error among high school graduate L1 Arabic learners: the predominant error was omission of a. Kharma (1981) looked at tertiary level students' errors in a cloze test and found a mean error rate of over 25% for all articles (*a*, *the* and  $\emptyset$ ), highest for *a* and lowest for *the*. However, in production data – a corpus of essays – Kharma (1981) reports half of all errors being in the use of *the*. Bataineh (2005) also looked at errors associated with a in a corpus of tertiary essays. Focusing only on the 319 syntactic/semantic errors in her data (i.e. leaving aside spelling and orthographic errors) and collapsing some of Bataineh's error categories, 61% of errors were instances of under-use:- "*∅-for-a*" (46%), "*the-for-a*" (15%). Overuse counted for the remainder with the largest types of error being "*a-for-the*" (27%) and "*a-for-∅*" (8%).

These studies do not all classify errors in the same way or count the same things; Bataineh's (2005) only looks at *a*, for example. Unfortunately Scott and Tucker (1974) and Bataineh (2005) do not tally accurate use instances so an overall success or failure rate for each article cannot be calculated. However, under-use of a in obligatory contexts is common in all three: for both the studies which considered all three articles, Scott and Tucker, and Kharma, this is the single most common type of article error. In Scott and Tucker's (1974) and Bataineh's (2005) data, "*∅-for-a*" is the more frequent error subtype; in Kharma's data "*the-for-a*" is slightly more frequent.

Kharma's (1981) data is the most detailed about types of article use; he identifies 5 types of use for the and a and 3 for  $\emptyset$ . Of the ten most common mistakes in his

cloze test six are categorized as instances of generic or general use (3 where the correct answer was Ø, and 3 where it was a), and 3 as mis-renderings of idiomatic phrases (e.g. *for an example*).

The guides for teachers, Kharma and Hajjaj (1997) and Smith (2001), point out similarities and differences between articles in Arabic and English. As for positive transfer, Smith (2001) notes that the definite article in Arabic, is used as in English to refer back to indefinite nouns in previous discourse and for unique references (*the floor*). As for negative transfer, both studies have similar lists of likely problems:

- Because of the absence of an indefinite marker in Arabic, <sup>2</sup> initial underuse of *a* (*This is book*) is to be expected and is likely to be followed by overuse (*These are a books*)
- Differing patterns of definiteness for the nouns in genitive constructions are likely to transfer (*Car the teacher*)
- In a range of idiomatic uses learners are likely to “reinstate” definite articles omitted in English (*I went to the bed*)
- Proper nouns in Arabic often contain the article (*He lived in the India*).

Kharma and Hajjaj (1997) add four other likely transfer problems:

- non-ellipsis of articles in compound noun phrases (*the salt and the pepper*)
- use of the definite article (obligatory in Arabic) in generic plural noun phrases (*The horses are useful animals*)
- use of the definite article for abstract nouns (*All men fear the death*)
- use of the definite article for mass nouns (*The milk is nutritious to the body*).

An interesting confirmation that generic uses of the article system might be one of the main stumbling blocks for advanced L1 Arabic English learners is Farghal and al-Zou'bi (2004), who compare three different English translations of the Koran and show evidence that none of the translators has fully acquired how English handles generic reference, using *the* instead of Ø with generic plural NPs (*The lions are ferocious animals*) and using *the* in generic singular NPs where English would require *a* (*The wolf ate him for A wolf ate him*).

In terms of the role of transfer in causing errors, Scott and Tucker (1974) attributed half their subjects' article errors to transfer, and Kharma (1981) attributes the majority of his subjects' errors to transfer but adds the compounding effect of

“inadequate instruction”. Bataineh (2005) considers transfer responsible only for “Ø-for-a” errors and even then only as a possibility, citing overgeneralization, training transfer, learning and communication strategies as other possible causes.

### *Teachability*

All the literature attests to the difficulty of the task L2 learners face in acquiring the English article system. Another pedagogic consideration is error gravity: if learners can communicate satisfactorily without mastery, is it worth devoting much time to article instruction? Master (1997, 2002) notes that non-mastery of the system damages the rhetorical credibility of those using English for academic purposes and argues that for these learners pedagogic intervention is called for: he cites research suggesting that instruction can make a difference to proficiency (Master, 1994, 1995, 2002) and in Master (1997) makes specific suggestions about instruction for learners at the main stages of proficiency. Liu and Gleason (2002) also suggest an instructional order appropriate to the order of acquisition they posit for non generic uses of the.

On the attested difficulty of acquiring high frequency L2 morphemes such as articles, in spite of the vast exposure to them learners receive, Ellis (2006b) argues that such difficulty is evidence that L1-like, naturalistic acquisition does not take place. He argues that this is because the form-function mappings of L1 inhibit or “block” learners perceiving those of L2, i.e. they fail to notice differences, and his conclusion is that without instruction fossilization is inevitable.

### **Research questions**

There have recently been a few corpus-based studies of article use in naturalistic L2 English. Ringbom (1998) noted that advanced learners of English with western European L1s all use the less frequently than native English speaker (NES) writers. Kachru (2003) looked at definite article use in acrolectal world English varieties and found little variation from NES norms. As we saw, McEnery, Xiao and Tono (2006) found evidence suggesting that even proficient L1 Japanese learners had not acquired articles. The research described in this article is a corpus-based enquiry into article system errors in the writing of intermediate to advanced level L1 Arabic English

learners. It tallies and categorizes the commonest errors in a naturalistic data corpus (i.e. a corpus of language not produced to elicit data on a specific linguistic item) and seeks to identify how much error can be attributed to transfer. It seeks to combine the delicacy of Kharma's (1981) categorization of types of error, Master's (1997) dual description of error frequencies in terms of form supplied and form required<sup>3</sup>, and Kharma and Hajjaj's (1997) categorization of transfer errors with a larger and hopefully more representative set of up to date naturalistic data than Kharma's (1981) to answer the following questions:

RQ1 How common are article system errors in university level L1 Arabic English learners?

RQ2 What are the commonest types of errors?

RQ3 To what extent are these errors likely to have been caused by transfer?

The article will conclude with a discussion of the research and pedagogic implications of the answers to these questions.

## **Method**

### *Data*

The corpus used was a subcorpus from a larger corpus of argumentative essays, modeled on the International Corpus of Learner English (ICLE) (Granger, Dagneaux & Meunier, 2002) written by first and second year students (aged 18-20) at the American University of Sharjah, an English medium university in the United Arab Emirates. The subcorpus, entitled the Arabic Learner English corpus (hereafter the ALE corpus), comprises 95 essays totaling 42,391 words (mean essay length 446 words) written by students giving Arabic as their L1. Only 21% are Emirati, the others coming from a cross-section of the many different Arab nationalities resident in the Emirates. The essays were word-processed and submitted as regular class assignments; they were often completed in the student's own time and not produced under test conditions. The English proficiency level of the students varied considerably; all had had to score a minimum of 500 in the TOEFL exam to enter the university, a few may be bilingual with one Anglophone parent, and secondary school education may have been in English, Arabic, or a mixture.

### *Procedure*

All tokens of *the* and *a* in the data were identified. The data was analyzed with the help of *Wordsmith Tools 4* (Scott, 2006). To identify  $\emptyset$  tokens and to speed analysis (by sorting KWIC concordances into recurrent patterns) the corpus was annotated with part of speech (POS) tags using the software application *Qtag* (Mason, 2008). As identifying tokens of  $\emptyset$  proved very time-consuming (see Mason and Uzar [2000] on the technical problems), it was decided to analyze a sample of these and extrapolate a likely overall frequency for the corpus. The sample of  $\emptyset$  tokens was obtained as follows: a random sample of 1000 of the 10,075 POS noun tokens was analysed. Instances (i) where noun tokens formed heads of NPs which included *the* and *a* and (ii) where noun tokens served as modifiers of other NP heads were removed. The remainder of the instances (373) were identified as  $\emptyset$  tokens.

Once the article tokens had been identified, the totals of tokens for each article type were converted to standardized frequencies and compared with those in other L1 and L2 English corpora. Next, each token was analyzed in context for correctness or incorrectness. This often required searching prior discourse for specific reference to the referent of the NP of which the article was part – directly or indirectly. For example, if the article token *the* was part of an NP *the student* did the prior discourse contain prior mention of *a student* or invoke a general situation (Quirk et al., 1985, p. 266) of which *the student* forms part? Conversely, if the article token *a* was part of an NP *a student* was this the first mention of this referent? If not, were these plausibly generic references? Once each “incorrect” token had been identified, an analytical decision was made as to which article would have been correct in its context. Usually this was unproblematic but occasionally it was concluded that no article at all fitted the context – the noun phrase needed a different determiner to be correct or no correction could be made at all. Examples of the former error are given in “Results” below. Tokens of the different error types for each article, e.g. “the-for-a”, were tallied. Finally, within each error type errors were further classified into typical patterns. In the next section the findings are presented and compared with those discussed in the literature review.

## Results and discussion

Detailed frequencies of article tokens in the ALE corpus – and, if used in error, their corrections – are given in Table 1. Figures 1 to 5 in Sections 4.1 and 4.2 are intended to represent visually the main tendencies revealed in Table 1.

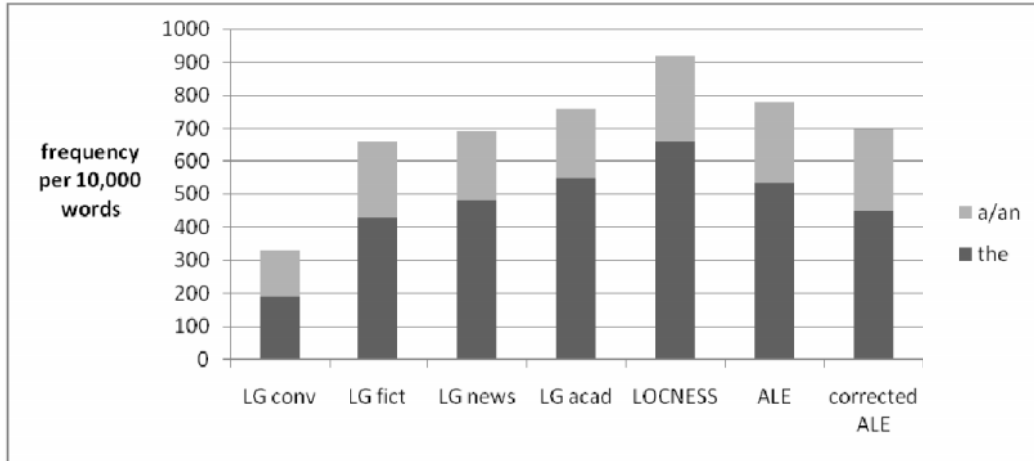
**Table 1: Frequencies of articles supplied in ALE by article type and correction type**

form required in context	article supplied				
	The	a/an	Ø*	<i>total visible articles</i>	<i>total all articles</i>
Ø	288	27	3686	<i>315</i>	4001
a/an	61	995	50	<i>1056</i>	1106
The	1872	6	20	<i>1878</i>	1898
another determiner	43	8	0	<i>51</i>	51
<b>Total</b>	<b>2264</b>	<b>1036</b>	<b>3756</b>	<b><i>3300</i></b>	<b>7056</b>
Error	392	41	70	<i>433</i>	503
correct use	1872	995	3686	<i>2867</i>	6553

\* projected

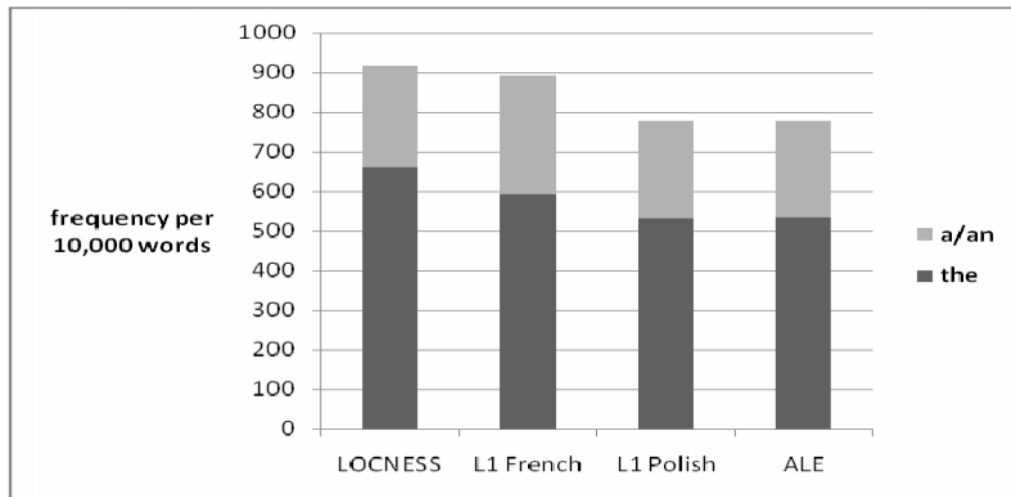
### *Article form frequencies*

Comparison of frequency counts of linguistic features in corpora give a useful overall picture of ways in which variation might occur. As Hunston (2002) notes, several studies of learner corpora show that learner writing has features more typical of spoken than of written academic English, such as more first and second person pronouns and more emphatic particles. Figure 1 compares frequency data from the ALE corpus, originally and after correction, with data from four registers in the Longman Grammar Corpus (LGC) as shown in Biber et al. (1999), and the Louvain Corpus of Native English Essays (LOCNESS). Why the LOCNESS corpus should have higher frequencies overall and a greater proportion of *the* than the other NES corpora is not known. The uncorrected ALE corpus appears to be much closer to the LGC academic register than either the spoken register or fiction, both in terms of overall frequency and proportion of definite to indefinite articles, but when corrected it resembles the LOCNESS corpus even less than before and the LGC academic writing register less than those of fiction and news.



**Figure 1: Frequencies of the visible articles in various NES corpora and ALE**

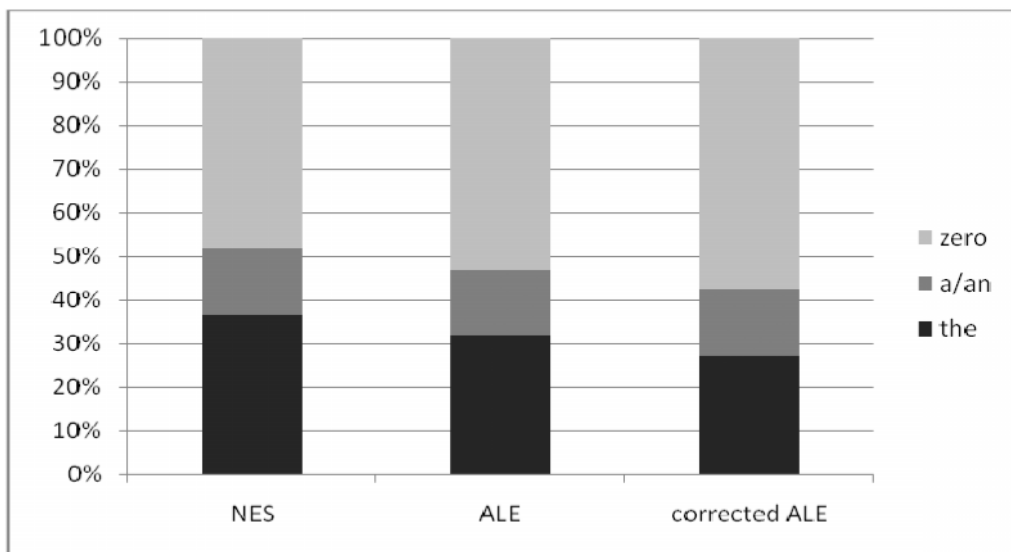
Figure 2 compares the same LOCNESS and uncorrected ALE frequency data with data from two other learner English corpora from the ICLE corpus. Ringbom (1998) notes that NES writers use the more than any of the NNES writers in the ICLE corpus and the ALE corpus seems typical in this respect. The frequencies for the French corpus more nearly resemble those of the NES corpus. Interestingly, from a transfer perspective, although Arabic is a [+article] language the ALE corpus frequencies more closely resemble those of the Polish corpus, Polish being a [-article] language.



**Figure 2: Frequencies of the visible articles in LOCNESS and other learner corpora**

Frequencies of  $\emptyset$  in NES language are harder to obtain, presumably because of the difficulty in automatizing identification mentioned earlier, and they are not

supplied either for the Longman corpus in Biber et al. (1999), or so far as I am aware, for the LOCNESS corpus. Master (1997) contains relative frequencies for the, a, and Ø counted in a corpus of 200,000 words from an NES corpus of EST articles. For the ALE corpus, extrapolating from the sample, we can estimate a figure of 888 Ø tokens per 10,000 word tokens of running text. Figure 3 compares Master's NES figures with the proportions in ALE, uncorrected and corrected. The uncorrected proportions would suggest that the L1 Arabic writers are underusing the and overusing Ø. However, the proportions in the corrected ALE corpus resemble the NES corpus even less. It may be the case that the variation here is at least partly due to register.

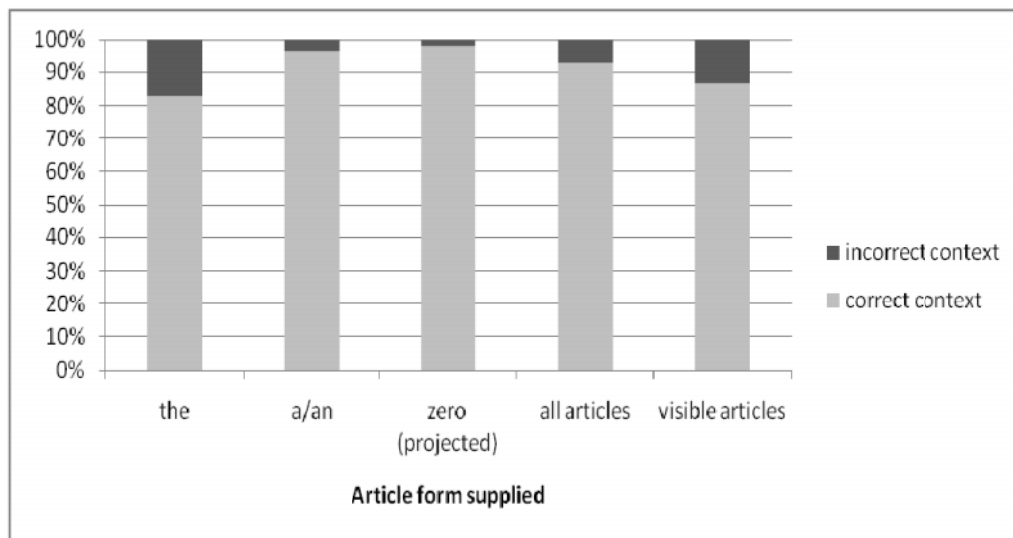


**Figure 3: Proportions of articles in NES, ALE and the corrected ALE**

In summary, one might have expected that with corrections, Non-Native English Speaker (NNES) text would have resembled what could be deemed target types of NES text more rather than less. These findings suggest that Ø is over-used generally by L1 Arabic writers, but that variation between article system use in NES and NNES text is not simply explicable in terms of linguistic proficiency: the reduced ratio of the to a suggests that the ALE corpus is conforming to Hunston's (2002) pattern of learner corpora resemblance to speech rather than writing. Use of articles in general and Ø in particular in NES language requires further research.

*Proportions of correct and incorrect use*

Turning from the overall frequencies of the articles, the issue of accuracy will now be addressed. Errors in article use can be conceptualized as a mismatch between a form and a context. There are two perspectives for viewing such mismatches – that of the article that has been supplied and that of the article which the context requires (or in some cases, the several article forms which the context permits). Figure 4 shows accuracy from the perspective of the article supplied. As can be seen, the errors are far from evenly spread between *the*, *a* and  $\emptyset$ . Do the error rates suggest that these L1 Arabic learners have acquired the English article system? Taking Dulay and Burt's (1974) accuracy rate (after Brown, 1973) of 90% or higher as a sign of acquisition, if we include  $\emptyset$  as part of the system, the answer would be "yes", the accuracy rate is 93%. If we leave out  $\emptyset$  and only consider the visible articles *the* and *a*, however, the accuracy rate drops to 87%. Given that this production is self-edited and largely written without time constraints, the issue seems borderline. Considered as subsystems, it would seem fair to conclude that  $\emptyset$  (98%) and *a* (96%) have been acquired but that *the* (83%) has not. That is to say only 83% of all *the* tokens in the corpus are used correctly. To check that *the* errors were not confined to a few low-proficiency authors the spread of the errors was checked: 90% of authors (86/95) mis-used *the* at least once, and the median number of mis-use errors per essay/author was three.

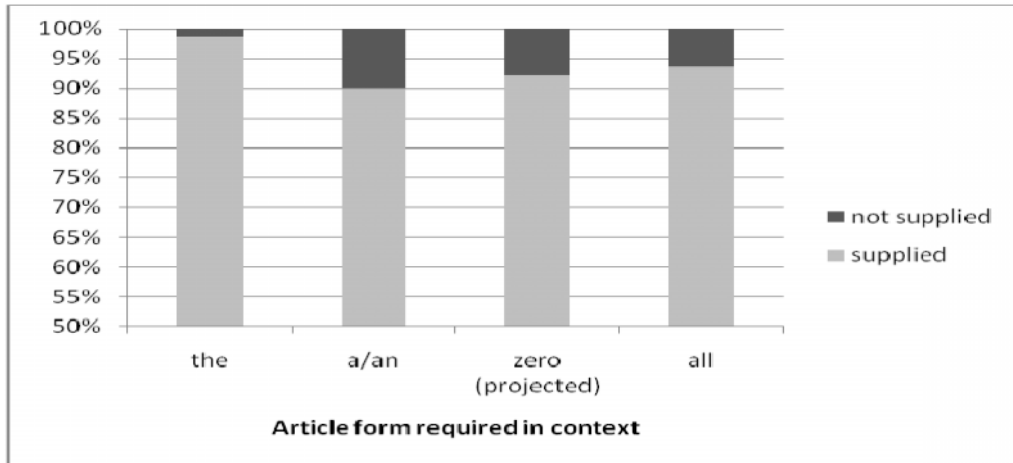


**Figure 4: Proportions of article forms in ALE used correctly**

An alternative way of considering the data, more comparable with the suppliance in obligatory context (SOC) elicitation tasks used in some research (e.g. Kharma, 1981; Master, 1997), is to analyze errors in terms of the article required in context. A slight disadvantage of this perspective is that it ignores overproduction of articles in toto: in this corpus, there were 51 instances of articles being used where the syntax of the NP required something else, for example a possessive or demonstrative. It is worth noting here that the existence of such instances highlights a limitation of some SOC elicitation techniques, i.e. that they will not show up features of interlanguage lacking a cognate in the target L2.

Figure 5 represents the error data from the perspective of context and suggests that in relative terms the article on which learners have the weakest hold is a (90%), followed by  $\emptyset$  (92%), and then the (99%). That is to say, in only 90% of the contexts in which a was required was it supplied. In terms of the 90% accuracy rate, learners can be said to have acquired a. Checking to see the spread of error in the corpus can only be done for “the-for-a” errors as the “ $\emptyset$ -for-a” figures are projected. The “the-for-a” figures show that this error appears in the texts of only 38% (36/95) authors. The more detailed analysis of this error-type below reveals that the error involves one particular use of a, for generic reference. Without a search of and frequency count of correct instances of this particular use, it is impossible to judge whether the relatively low spread of this error is due to either (i) mastery of this particular use by the majority of authors or (ii) a low number of attempts at this use.

To summarize, the answer to RQ1 is complex, then, but appears to be that errors with these are common enough to cast doubt on these learners’ mastery of the form.



**Figure 5: Proportions of article contexts in ALE for which the correct article form was supplied**

*Patterns of incorrect use*

In this section the main patterns of error are reviewed in some detail and the plausibility of L1 transfer being the cause is discussed. Table 2 answers RQ2, listing errors from the perspective of the item supplied, ordering errors from most to least common.

**Table 2: Frequencies of types of article error**

Type of error	N	%
the-for-Ø	288	57
the-for-a	61	12
Ø-for-a*	50	10
the-for-other determiner	43	9
a-for-Ø	27	5
Ø-for-the*	20	4
a-for-other determiner	8	2
a-for-the	6	1
<b>Total</b>	<b>503</b>	<b>100</b>

\* projected

*“the-for-Ø” errors.*

As we have seen *the* is the most frequently mis-used article and its most frequent mis-use is in  $\emptyset$  contexts. These contexts can be further subdivided into the error types shown in Table 3.

**Table 3: Frequencies of subtypes of ‘the-for- $\emptyset$ ’ errors**

Type of construction	N	%
generic non count nouns	128	44
generic plural count nouns	109	38
generic singular count nouns	34	12
Idioms	17	6
<b>Total</b>	<b>288</b>	<b>100</b>

The following generic non-count nouns, usually abstract, all occurred more than once: business, company profit, culture, growth, human nature, imagination, lack, money, nature, environmentalism, ozone depletion, poverty, society, real life, television, war, water, work. It should be noted that all of these could appear non-generically with *the* in a range of contexts, such as previous mention, or invocation of a schematic situation referred to or implicit in the discourse context. However, each analysis of incorrectness was made on the basis of careful inspection of the previous discourse. The society (21), the money (15), and the real life (9) were the top three. Here are some characteristic examples:

- [1] ...These things destroy *the society*.<sup>4</sup>
- [2] Some of us consider *the money* as the force which controls our lives, while others...
- [3] King Fahd University graduates are knowledgeable and ready to join *the real life* from the first day in their business.

The occurrence of the money is perhaps some small evidence in favour of Ellis’s (2006b) “blocking of perception” account of L1 influence, as  $\emptyset$  money actually featured in the prompt used to elicit many of the texts in which it appears (“*Money is the root of all evil*”. Discuss.) The real life is an interesting case where the prompt contained the semantically related expression *the real world*: “*Most university degrees...do not prepare students for the real world.*” Discuss.

Examples of idioms involving generic non-count nouns are the following:

- [4] At the moment a problem arises, the first thing that *comes to the mind* is how to deal with it.
- [5] ...students can memorize some materials and *put it on the paper* the next day.

The next largest error category was generic plurals:

- [6] Others say that *the degrees* are mainly theoretical...
- [7] This invention allowed people to value *the goods and services*.
- [8] ...they are usually good with thinking about the whole including *the individuals*.

Related to this is the next error category – singular nouns which are intended generically but do not satisfy the normal conditions for using *the*, for example:

- [9] *The human* lives in the earth and his existence depend on nature's ...[*Humans live on earth*]
- [10] ...in some way it seems as though *the society* is "baby-sitting" criminals [ *Ø society*]
- [11] The root of all evil is *the man's* desire for status and greed... [ *man's* or *mens' desire*]

Examples of errors which show learners to be unaware of idiomatic NES non-use often involve nouns representing institutions – *jail, prison, university, hospital*, for example:

- [12] *In the university*, students do not gain the experience...

“*the-for-a*” errors.

The vast majority of these errors (55 out of 61) were where the writer intended a generic expression. Another way of correcting these would be to pluralize them and use *Ø*, but because this involved two corrections rather than one they were analysed as singular:

- [13] I believe that the personality of *the person* is what stands out rather than just ...[*a person, people*]
- [14] The only thing a university degree does is provide *the student* with the basics required for him to understand ... [ *a student, students*]

- [15] For example, *the water bottle* needs years to decompose. [*a water bottle, water bottles*]

“Ø-for-a” errors.

There is only a small sample of these. All involve a generic, classificatory meaning, for example:

- [16] Money is [Ø] way to ease a person’s life.  
 [17] For example, you are [Ø] person that do not have enough money...  
 [18] In conclusion, [Ø] theoretical education is very poor...

The head nouns in both the first two examples are countable and would require an article in any context, generic or non-generic, but the last is more complex in that *education* can be a non-count noun and only requires *a* here to carry the generic meaning required in this context – a discussion about types of education.

“Ø-for-the” errors.

The sample of these errors is even smaller. One involves an idiom:

- [19] While in [Ø] contrary, the sentimental environmentalism are spending ...

“a-for-Ø” errors.

The sub-types of error in this category are shown in Table 4.

**Table 4: Frequencies of subtypes of ‘a-for-Ø’ errors**

Type of construction	N	%
generic non-count nouns	16	59
plural count nouns	6	22
Idiom	3	11
Other	2	7
<b>Total</b>	<b>27</b>	<b>100</b>

The majority of errors involved generic non-count nouns, for example:

- [20] Since money is the most accurate measurement of *a status* for individuals and citizens in a country.  
 [21] Every one in that family lives in *a harmony*...

The next largest category involved plural countable nouns, for example:

- [22] People have different point of views regarding if *a university degrees* make individuals more capable or are they just worthless certificates.

This surprising error type is listed by Smith (2001) as to be expected and Bataineh's (2005) data contained 14 such errors. Such errors are presumably the trace of an earlier stage of learning.

4.3.6 *Other determiner required.* Of constructions which require a determiner other than an article, the largest type (22/51) are those requiring a possessive, for example:

- [23] ... and as he's the son of *the father* at that time, crime is therefore an act of instinct [*his father*]
- [24] These groups, jump from the A and B, and started their life from C, and jumping will break *the leg* even after a long time... [*their legs*]
- [25] So they use money to remove what ever that can face *the dream* by using money. [*their dream*]

There are 6 cases of *most of the* where *most* alone is required, e.g.

- [26] This is why *most of the people* fight to increase their fortune.

These can be considered as modified generics, [26] being for example a hedged form of *This is why people fight to increase to their fortune.*

### *Transfer*

Having surveyed the commonest types of errors in some detail, what can we conclude about RQ3, the role of L1 transfer? The English article system is famously complex. Analytical taxonomies of article use in the literature typically list around twenty contexts for the articles *a*, *the* and *zero*, of which only three are generic (Berry 1993; Thomas 1989). Liu and Gleason argue that 'the non-generic use of *the* is much more complex and hence more problematic for ESL students than the generic use' (2002, p. 6). First, let us note the surprising fact that none of the errors involved countable specific referents: most involved either uncountable or countable generic referents. This suggests, as predicted by Smith (2001), positive transfer of use of the definite article for specific reference.

It is, secondly, striking that the largest single category of errors is incorrectly formed generic expressions: looking at the “visible” articles alone (as the small sample size for  $\emptyset$  errors makes the figures unreliable) the proportion of errors which involve an apparent intended generic meaning is 79% (342/433, see Table 5 for a detailed breakdown). Of these errors only 16 did not involve the incorrect use of *the*. In other words, 75% of all the errors in use of *a* and *the* consisted of learners using *the* where it was not required in order to form a generic expression.

**Table 5: Frequencies of errors with ‘visible’ articles in generic reference by NP head type**

Supplied	required	NP head	N
The	$\emptyset$	generic non count noun	128
The	$\emptyset$	generic plural count noun	109
The	$\emptyset$	generic singular count noun*	34
The	A	generic singular count noun	55
A	$\emptyset$	generic non count noun	16
<b>Total</b>			<b>342</b>

\* pluralization also required

Are these errors likely to be due to L1 transfer? As we noted earlier Kharma and Hajjaj (1997) predicted errors with generic noun phrases and with abstract and mass nouns – it happens that in these essays abstract and mass nouns invariably appeared as generics, so these error types are often conflated.

In Arabic, in which there is no indefinite article, indefiniteness is marked by the absence of the definite article (Schulz, 2004).<sup>5</sup> The only forms in Arabic permissible for generic NPs, singular, plural, or non-count, require the definite article (Farghal & al-Zou'bi, 2004). In addition to the fact that there is a number distinction in the English indefinite article and therefore two forms (*a* and  $\emptyset$ ), there are two factors likely to make acquisition of English articles in generic expressions difficult for L1 Arabic speakers. Firstly, recalling Lyon’s (1999) categorization of generics as being identifiable, the Arabic article system does not deviate from the prototypical meaning-to-form mapping of identifiability-definiteness as the English system does. Secondly, the English system does not deviate consistently, but sometimes admits generics as definite. Table 6 compares some feature of the two systems.

**Table 6: Comparison of definiteness marking according to number and countability in English and Arabic**

head of generic NP	pattern	definiteness marking in article	English	Arabic
plural count noun	i	indef.	✓	✗
	ii	def.	✓ *	✓
singular count noun	iii	indef.	✓	✗
	iv	def.	✓ *	✓
non count noun	v	indef.	✓	✗
	vi	def.	✗	✓

\*= sometimes

There are in theory 6 patterns possible. Berry (1993), based on data from the COBUILD corpus, states that the commonest generic reference patterns in English are (i) and (v). If the less common (ii) and (iv) patterns never appeared in English there would be complete formal definite/indefinite asymmetry and it might arguably be cognitively easier for Arabic speakers to notice the differences between English and Arabic. However, there are special cases in which the Arabic patterns are permissible in English (Berry, 1993):

(ii) singular count noun + definite article

regular participants and roles in a situation – *the teacher* in education

animal species – *the red squirrel*

body parts as referred to by doctors – *the head*

inventions – *the computer*

rooms – the kitchen.

(iv) plural count noun + definite article

nationality words – *the Americans*

nominalized adjectives – *the poor*.

One could predict that even without “blocking” effects from L1 acquisition (Ellis, 2006b) acquiring the semantic cues for these special cases will require a lot of exposure.

Table 7 breaks down the errors with generic expressions into the six patterns. It shows that the vast majority of the errors follow patterns which would be correct in

Arabic (ii, iv and vi). The small remainder come from making a wrong choice about the form of the indefinite article (v).

**Table 7: Errors with generic reference by type, ranked from most to least frequent**

<b>Pattern</b>	<b>nature of error</b>	<b>Correction</b>	<b>N</b>
vi	the + non count noun	def. for indef. the to Ø	128
ii	the + plural count noun	def. for indef. the to Ø	109
iv	the + singular count noun	def. for indef. the to a	(55)
		the to Ø + pluralization	(34) 89
v	a + non count noun	wrong indef. Ø to a	16
<b>Total</b>			<b>342</b>

Until L1 Arabic learners notice the differences between the systems in English and Arabic, a strategy for count nouns of sticking to L1 patterns (ii and iv), i.e. routinely using the definite article in generic expressions, will produce correct English some of the time. In other cases, interlocutors (including English teachers) may not recognize the error as a formal one, and may incorrectly assume that the speaker has made a pragmatic error, i.e. is using the definite article infelicitously to make a specific reference - for example to a referent she imagines to be familiar to the hearer - rather than a generic reference. Such misinterpretations may in turn reduce the chance of the learner's attention being drawn to the error.

In terms of causality, generic article errors are arguably both interlingual and intralingual: for example, the learner who writes "*the degrees* are mainly theoretical" has both transferred an L1 form and possibly overgeneralized that this is a permissible form in L2 from the evidence of L1 input, both generic, such as "*the Americans*" above, and non-generic, such as "*the best degrees*". However, in terms of the original interlingual-intralingual distinction these errors are not obviously intralingual because they do not take the form of constructions which would be impossible in L1.

Let us summarize by reviewing the evidence needed to answer RQ3, which concerns the extent to which article system errors in this corpus appear likely to have been caused by transfer. The situation revealed by the data is that learners appear to have mastered all the articles and usually supply *a* and *the* in the contexts in which

they are required. Unfortunately, however, they also tend to use *the* in contexts where it is not required. This over-use of *the* is not scattered equally throughout all possible article contexts, however, but is, in the main, as we have seen, confined to a small minority of contexts, in particular those involving a generic meaning. Although, generic expressions are much less frequent than specific expressions overall, three quarters of all errors made with *the* and *a* occur in what are intended as generic expressions. Examination of the L1 shows that there are only three permissible forms of generic expression, each of which requires use of the L1 cognate of *the*. In short, it would seem reasonable to conclude there is a strong likelihood that transfer from L1 is a major factor in causing most of the article errors found in the corpus. In terms of form-function mapping one would say that Arabic learners have learnt from L1 to map the function ‘marker of generic expression’ onto the form ‘definite article’.

### **Conclusion**

There is not space here to explore the role of transfer in some of the much lower frequency types of error not related to generic reference noted in the corpus. However, in terms of the second research question about the commonest errors, those with generic reference are clearly predominant. The argumentative genre of the texts which comprise the data for this research probably contain a particularly high rate of generic reference; a corpus of narrative texts would presumably have a higher level of specific reference, anaphoric reference in particular, and could thus well have had lower overall error rates in article use. However, texts like those in ALE and LOCNESS are not untypical of assignments in academic writing classes, and it can be expected that much disciplinary academic writing will require students to make general statements about classes of things and thus use generic expressions.

Because the analysis in this research has been confined to errors, we do not know how far the incorrect article use described here exists among individual authors in free variation with correct article use (as found in Jarvis, 2002; Robertson, 2000). My assumption is that even if there were also a high level of correct article use for generic reference, the errors discussed here are evidence of unsystematic variation, and therefore evidence that acquisition has not taken place. Overall then, to answer the research questions on the evidence so far discussed, it seems fair to conclude (RQ1)

that despite an impressive accuracy rate in article use for specific reference a majority of the L1 Arabic advanced learners of English represented in the corpus have not fully mastered the English article system, (RQ2) that most of the errors they make with the visible articles are in the expression of generic reference, most of these errors involving over-use of *the*, and (RQ3) that transfer is a likely factor in such errors. These findings suggest that, despite the considerably increased exposure to English brought about by social, economic, and technological changes in the last quarter century in the Arabian Gulf region, the situation has changed little from when Kharma (1981) tested article acquisition among tertiary level L1 Arabic learners of English.

The broader implications of these findings seem to be as follows. Learners from languages without article systems apparently have more problems (a) overall and (b) earlier in acquiring the English article system (Master, 1997); however, learners from languages with article systems, such as Arabic, may also face significant problems and these problems may survive until relatively advanced stages of learning. Furthermore, the problems faced may be related to specific L1 system features, and not, as some earlier SLA theory suggested, features of a common developmental interlanguage system. In general, in confirming the persistence of such errors, in the face of high frequency exposure to L2 input, the findings also tend to support Ellis's (2006b) argument for maximal transfer and interference from L1 in L2 learning.

Overall, research into learner problems with English articles in general and with generic reference in particular would benefit from more corpus-based research into frequencies of the major reference patterns in native-speaker English, similar to that shown in Biber et al. (1999), but extending to indefinite NPs, including the numerically dominant  $\emptyset$  NPs.<sup>6</sup> The findings of the current study suggest that research into generic reference in the English of learners from L1s with and without articles would be worthwhile, possibly revealing negative transfer effects worthy of pedagogic intervention. In particular, across learners from [+article] L1s it would be interesting to compare similar data from two further subcategories, ELS learners with L1s in which, as in Arabic, generics are marked by definiteness, and those with L1s in which, as in English, generics are marked by indefiniteness, subcategories, which we might designate as [+generic definiteness] or [-generic definiteness]. The research

described here has concentrated on errors: to fully gauge learners' accuracy rates at the different types of article use it would be useful to analyze also frequencies of successful article use in generic expressions in L2 corpora. Research on article use by L1 Arabic speakers at earlier levels of language learning would also help to confirm whether the hypothesis of early positive transfer of specific reference is supported.

The pedagogic implications seem to be as follows. Problems related to the impact of the count/non-count distinction on definiteness marking would presumably be reduced if vocabulary teaching in the earlier stages of learning emphasized this distinction more, as suggested by Master (1997). The commonest form of generic expression for countable nouns in English is that with  $\emptyset$  and plurals. For learners with [-article] L1s this use of  $\emptyset$  will transfer quite naturally. For L1 Arabic learners, and presumably for learners with other [+generic definiteness] L1s, such as French, negative transfer can be expected. At the intermediate level, Master focuses on the discourse dependence of many learner problems but in the case of L1 Arabic learners we have seen that this is not a primary cause of errors. However, Master's suggestions for helping advanced [-article] L1 learners overcome "Ø-for-the" errors may also apply to helping [+article] L1 learners overcome "the-for-Ø" and "the-for-a" errors. He recommends focusing on articles in both reading and writing development, "especially the contrast between minimal pairs of lexical phrases with  $\emptyset$  and *the*" and the need for repetition of the process "with many different lexical items since it can no longer be assumed that learners at this stage will generalize from specific instances" (1997, p. 228). Master assumes that conscious learning has little effect on acquisition but if recent theorizing about the role of consciousness upon implicit cognition (Ellis, 2006a; 2006b; Ellis & Larsen-Freeman, 2006) is correct it would also be worth spending class time explicitly pointing out the various formal possibilities and impossibilities for generic reference in English. Finally, and perhaps most importantly, in order for maximal opportunity to be taken of promoting noticing, it would seem worth making known to those engaged in instructing L1 Arabic ESL learners the likely L1 origins of some persistent errors in their students' production.

### Notes

1. In this paper *a* will be taken to stand for both *a* and *an*. I have avoided using the expression “the indefinite article” to describe *a/an* as this practice is inconsistent with recognizing  $\emptyset$  as an indefinite article
2. In some varieties of Arabic, indefiniteness is marked by a form of affixation, nunation. However, while nunation is a feature of written acrolectal Arabic, in the modern spoken dialects it is described as having “all but disappeared” (Holes, 2004, p. 174) and being “rare” (Abu-Chacra, 2007, p. 18).
3. Master (1997) makes this distinction in terms of accuracy and overuse but I have avoided these particular terms as being susceptible to misinterpretation outside the context of his paper.
4. Errors in quotations from the ALE corpus are not corrected – although suggested alternatives are sometimes presented in square brackets - and all italicization and ellipsis is mine.
5. See note 2 above.
6. Biber et al. (1999, p. 266) gives figures for generic the but not for *a* or  $\emptyset$ .

### References

- Abu-Chacra, F. (2007). *Arabic: An essential grammar*. New York: Routledge.
- Bataineh, R. F. (2005). Jordanian undergraduate EFL students' errors in the use of the indefinite article. Retrieved June 11, 2008 from Asian EFL Journal: [http://www.asian-efl-journal.com/march\\_05\\_rfb.php](http://www.asian-efl-journal.com/march_05_rfb.php)
- Berry, R. (1993). *Articles*. London: HarperCollins.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow, Essex: Longman.
- Bley-Vroman, R. (1983). The comparative fallacy in interlanguage studies: The case of systematicity. *Language Learning* 33 , 1-17.
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

- Butler, Y. G. (2002). Second language learners' theories on the use of English articles. *Studies in Second Language Acquisition* 24, 451-480.
- Chaudron, C. & Parker, K. (1990). Discourse markedness and structural markedness. *Studies in Second Language Acquisition: The Acquisition of English Noun Phrases* 12(1), 43-64.
- Dulay, H. & Burt, M. (1974). Errors and strategies in child language acquisition. *TESOL Quarterly* 8, 129-36.
- Dulay, H., Burt, M., & Krashen, S. (1982). *Language two*. New York: Oxford University Press.
- Ellis, N. C. (2006a). Language acquisition as rational contingency learning. *Applied Linguistics* 27(1), 1-24.
- Ellis, N. C. (2006b). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics* 27(2), 164-194.
- Ellis, N. C. & Larsen-Freeman, D. (2006). 'Language emergence: Implications for applied linguistics - Introduction to the special issue. *Applied Linguistics* 27(4), 558-589.
- Farghal, M. & al-Zou'bi, I. (2004). The Qur'anic definite article in translation: The case of citing animals. In R. Elhamy, H. Mansour, & J.R. Aydelott (Eds.), *Contrastive rhetoric: Issues, insights and pedagogy* (pp. 61-68). Cairo: American University of Cairo Press.
- Granger, S., Dagneaux, E. & Meunier, F. (2002). *International corpus of learner English*. Louvain: Presses Universitaires de Louvain.
- Holes, C. (2004). *Modern Arabic: Structures, functions, and varieties (2nd edn.)*. Washington: Georgetown University Press.
- Huebner, T. (1983). *A longitudinal analysis of the acquisition of English*. Ann Arbor: Karoma.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Ionin, T., Ko, H., & Wexler, K. (2004). Article semantics in L2 acquisition: The role of specificity. *Language Acquisition* 12(1), 3-69.

- James, K. (1998). *Errors in language learning and use: Exploring error analysis*. London: Longman.
- Jarvis, S. (2002). Topic continuity in L2 English article use. *Studies in Second Language Acquisition* 24, 387-418.
- Kachru, Y. (2003). On definite reference in world Englishes. *World Englishes* 22(4), 497-510.
- Kharma, N. (1981). Analysis of the errors committed by Arab university students in the use of English definite-indefinite articles. *International Review of Applied Linguistics* 19(4), 331-345.
- Kharma, N. & Hajjaj, A. (1997). *Errors in English among Arabic speakers: Analysis and remedy*. Beirut: York Press Librairie du Liban.
- Lado, R. (1957). *Linguistics across cultures*. Ann Arbor: Michigan University Press.
- Liu, D. & Gleason, J.L. (2002). Acquisition of the article "the" by nonnative speakers of English. *Studies in Second Language Acquisition* 24, 1-26.
- Lyons, C. (1999). *Definiteness*. Cambridge: Cambridge University Press.
- Mason, O. (n.d.). *Qtag*. Retrieved June 1, 2008 from <http://www.english.bham.ac.uk/staff/omason/software/qtag.html>
- Mason, O. & Uzar, R. (2000). NLP meets TEFL: tracing the zero. In B. Lewandowska-Tomaszczyk & J. Melia (Eds.), *TALC '99 Practical Applications in Language Corpora: papers from the international conference at the University of Lodz, 15-18 April 1999* (pp. 105-116). Frankfurt: Peter Lang.
- Master, P. (1990). Teaching the English articles as a binary system. *TESOL Quarterly* 24(3), 461-478.
- Master, P. (1994). The effect of systematic instruction on learning the English article system. In T. Odlin (Ed.), *Perspectives in pedagogical grammar* (pp. 229-252). Cambridge: Cambridge University Press.
- 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* 2(2), 215-232.

- Master, P. (2002). Information structure and English article pedagogy. *System* 30, 331-348.
- McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. London: Routledge.
- Mohammed, A. M. (2000). Modern standard Arabic vs. non-standard Arabic: Where do Arab students of EFL transfer from? *Language, Culture and Curriculum* 13(2), 126-136.
- Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.
- Parrish, B. (1987). A new look at methodologies in the study of article acquisition for learners of ESL. *Language Learning* 37, 361-383.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. Harlow: Longman.
- Ringbom, H. (1998). Vocabulary frequencies in advanced learner English: A cross-linguistic approach. In Granger, S. (Ed.), *Learner English on computer* (pp. 41-52). London: Longman.
- Robertson, D. (2000). Variability in the use of the English article system by Chinese learners of English. *Second Language Research* 16(2), 135-172.
- Schulz, E. (2004). *A student grammar of modern standard Arabic*. Cambridge: Cambridge University Press.
- Scott, M. (2006). *Wordsmith tools 4*. Oxford: Oxford University Press.
- Scott, M. S. & Tucker, G. R. (1974). Error analysis and English language strategies of Arabic students. *Language Learning* 24(1), 69-97.
- Smith, B. (2001). Arabic speakers. In M. Swan & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems (2nd Ed.)* (pp. 195-213). Cambridge: Cambridge University Press.
- Thomas, M. (1989). The acquisition of English articles by first- and second-language learners. *Applied Psycholinguistics* 10(3), 335-355.
- Thompson-Panos, K. & Thomas-Ruzic, M. (1983). The least you should know about Arabic: Implications for the ESL writing instructor. *TESOL Quarterly* 17(4), 609-623.

- White, L. (1977). Error analysis and error correction in adult learners of English as a second language. *Working Papers in Bilingualism 13*, 42-58.
- Zdorenko, T. & Paradis, J. (2008). The acquisition of articles in child second language English: fluctuation, transfer or both? *Second Language Research 24*(2), 227-250.
- Zegerac, V. (2004). Relevance Theory and "the" in second language acquisition. *Second Language Learning 20*(3), 193-211.

