PJEE

Premise: Journal of English Education and Applied Linguistics
e-ISSN: 2442-482x, p-ISSN: 2089-3345
https://fkip.ummetro.ac.id/journal/index.php/english
DOI: 10.24127/pj.v12i2.7502

ACCURACY ORDER OF THE ACQUISITION OF THE ENGLISH ARTICLE SYSTEM: EVIDENCE FROM THE WRITTEN AND ORAL PRODUCTION OF UNDERGRADUATE ARAB EFL LEARNERS

by

Muneera Muftah*
Department of English, Faculty of Arts, Thamar University, P.O. Box. 87246, Dhamar, Yemen
munmef5@gmail.com
ORCiD https://orcid.org/0000-0002-5898-5426

***Corresponding author**

(Article History: Received:2023-04-01.Revised1:2023-05-15.Revised2:2023-05-19. Accepted:2023-06-01. Published: 2023-06-30).

Abstract

The present study investigates the accuracy order and underlying processes of acquiring English articles by 45 undergraduate Arab EFL learners drawn from three proficiency groups. The theoretical approach adopted here is Bickerton's (1981) semantic model that focuses on the features ±Specific Referent(±SR) and ±Assumed Known to the Hearer (±HK). The study aims to identify what the three measures, SOC (Supplied in Obligatory Contexts), TLU (Target-Like Use), and UOC (Used in Obligatory Contexts), reveal about the accuracy order of the acquisition of English articles in the written and oral production tasks. It also seeks to determine the semantic contexts in which the articles are overgeneralized and the most challenging context to acquire. The results show that the SOC measure reveals an accuracy order of a >the >Ø and the >a >Ø, whereas the TLU measure reveals an accuracy order of the>a>Ø across the groups. In addition, the UOC measure reveals that Ø goes through a flooding-then-trickling process, then experiences a U-shaped behavior highlighted by an overgeneralization stage, and follows the by undergoing a U-shaped development. Furthermore, the learners have difficulty distinguishing [±HK] and [±Countability]. The accuracy comparison in article use showed that [+SR, −HK] and [−SR, +HK] were the most challenging contexts to acquire.

***Keywords**: accuracy order, article acquisition processes, Bickerton's semantic model, EFL Learning, English articles,*

Abstrak

Studi ini menyelidiki urutan akurasi dan proses yang mendasari perolehan artikel bahasa Inggris oleh 45 pelajar EFL Arab sarjana yang diambil dari tiga kelompok kemahiran. Pendekatan teoritis yang diadopsi di sini adalah model semantik Bickerton (1981) yang berfokus pada fitur ±Specific Referent(±SR) dan ±Assumed Known to the Hearer (±HK). Penelitian ini bertujuan untuk mengidentifikasi apa yang tiga ukuran, SOC (Disediakan dalam Konteks Wajib), TLU (Penggunaan Seperti Target), dan UOC (Digunakan dalam Konteks Wajib), mengungkapkan tentang urutan akurasi akuisisi artikel bahasa Inggris dalam tugas produksi tertulis dan lisan. Ini juga berusaha untuk menentukan konteks semantik di mana artikel terlalu umum dan konteks yang paling menantang untuk diperoleh. Hasil penelitian menunjukkan bahwa ukuran SOC mengungkapkan urutan akurasi >>Ø dan >a >Ø, sedangkan ukuran TLU mengungkapkan urutan

How to cite this article

Muftah, M. (2023). Accuracy order of the acquisition of the English article system : evidence from the written and oral production of undergraduate Arab EFL learners. *Presmie; Journal of English Education and Applied Linguistics*, *12*(2), 740–766. https://doi.org/10.24127/pj.v12i2.7502

akurasi >a>Ø di seluruh grup. Selain itu, ukuran UOC mengungkapkan bahwa Ø melewati proses banjir-kemudian-menetes, kemudian mengalami perilaku berbentuk U yang disorot oleh tahap generalisasi berlebihan, dan mengikuti dengan menjalani perkembangan berbentuk U. Selain itu, peserta didik mengalami kesulitan membedakan [±HK] dan [±Countability]. Perbandingan akurasi dalam penggunaan artikel menunjukkan bahwa [+SR, −HK] dan [−SR, +HK] adalah konteks yang paling menantang untuk diperoleh.

Kata kunci: artikel bahasa Inggris, EFL, urutan akurasi, proses akuisisi artikel, model semantik Bickerton,

INTRODUCTION

Learning English articles *a (n), the,* and the *zero articles* is one of the important issues for learners of English as a foreign language (EFL). Articles are very commonly used morphemes, and their usage is complex for EFL learners, causing even the most advanced non-native speakers (NNS) of English to make errors. These errors occur even when other elements of the language seem to have been mastered (Ekiert, 2004; 2007). Part of the complexity can be attributed to the fact that the English article system not consisting of one-to-one form and meaning relationships but stacks multiple functions onto a single morpheme. When processing language primarily for meaning, function words, unlike content words, are generally overlooked by learners.

Articles are generally unstressed and almost inaudible. Consequently, they are tough for Arab EFL learners to discern, affecting input availability in the spoken mode (Bataineh, 2005). Nevertheless, given that they are among the most frequent words in English, it is of the utmost significance that high proficiency level students have some control over their usage.

The English articles are challenging to acquire for EFL learners and children learning English as a first language (L1). L1 Children master the article system early by the age of four (Park, 2006). However, second language (L2) learners need help mastering the article system. In particular, adult L2 learners in the EFL environment have persistent difficulty (Abudalbuh, 2016; Martinović & Balenović, 2019; Muftah, 2023a, 2023b; Sun, 2016; Suniphan, 2022; Testa, 2019). L1 children acquire the article system naturally, employing the language-specific operation called Language Acquisition Device (LAD) (Chomsky, 1968). However, adult L2 learners may depend on a mechanism different from what L1 children employ. Many proposals argue that adult L2 learning may be based on cognitive mechanisms instead of the LAD. They create conscious deductive rules and rely on the mechanism in selecting an article. Particularly, L2 learners in the EFL context may depend largely on explicit knowledge learned through

formal instruction given in classroom settings(Muftah, 2016b, 2016a; Muftah & Eng, 2011; Muftah & Rafik-Galea, 2013).

The complexity of choosing accurate articles poses several challenges for EFL learners (Andersen, 1978). According to Pienemann (1998), the difficulty of the meaning expressed by an article is determined by the novelty and abstractness of the concept, not to mention learners' changing hypotheses about article usage at different stages in interlanguage development and the potential influence of the native language which may further complicate the task.

Semantically, the Arabic article system is similar to English; however, the forms are highly varied. While the Arabic system manifests a binary distinction between the defined and the undefined, the English system exhibits a tripartite distinction. The Arabic defined (marked by the definite article *al*) and the undefined (marked by the absence of *al*) correspond to the English defined (marked by the definite article *the*) and the undefined (marked by the indefinite articles *a(n)* and *zero*). In other words, even though the concept is present in the two languages, indefiniteness in English is marked by lexical items such as *a* and *a*, while it is marked in Arabic by affixes such as the prefix *al* and the suffix *–n*, both to mark definiteness and indefiniteness respectively (Lyons, 1999). For example, the Arabic and English sentences below are translation equivalents (Bataineh, 2005):

Dahara rajulun filbalde.
*Appeared man in town
The man appeared in town. Or
A man appeared in town.

Celce-Murcia and Freemen-Larsen (1999) claim that the problematics of using the article system is partly due to whether the lexical classification into countable versus uncountable nouns corresponds in the native and target languages. For example, while *furniture* and *equipment* are uncountable in Arabic and English, *chalk* and *information* are countable in Arabic and uncountable in English. This mismatch may add to the complexity of the learner's task, for they need to learn both the article systems and other noun distinctions.

A good number of the studies which yielded significant findings (Hakuta, 1976; Huebner, 1983a, 1983b; E. E. Tarone, 1985) were explicitly conducted to examine grammatical morphemes rather than article learning and acquisition per se. Only Ekiert (2004), Lu (2001), Master (1997), Parrish (1987), Tarone and Parrish (1988), and Thomas (1989) specifically

studied the acquisition of articles by employing Bickerton's (1981) binary semantic system [±Specific Referent, ±Hearer Knowledge] for noun phrase reference.

Master (1987) was the first to point out that articles seem to be acquired differently, depending on whether or not they occur in the learner's native language. Overall, acquiring the definite article precedes acquiring the indefinite article (Huebner, 1983; Master, 1997; Parrish, 1987; Thomas, 1989). Several studies (Chaudron & Parker, 1990; Lu, 2001; Huebner, 1985; Thomas, 1989) reveal overuse of the definite article, but higher proficiency learners improved accuracy with the indefinite *a*. Although both Huebner (1983) and Master (1997) referred to the phenomenon of '*the*-flooding' in which *the* is overgeneralized with a dramatic rise in usage, Thomas (1989) found the *zero* articles overgeneralized across proficiency levels.

For learners whose native languages lack articles, researchers (Ekiert, 2004, 2007; Master, 1997; Parrish, 1987) reported that *zero* articles dominate in all environments for articles in the early stages of language learning. Parrish (1987) suggested an order of acquisition in which the *zero*, *definite*, and *indefinite* articles are acquired consecutively. On the other hand, Wong and Quek (2007) reported in their study of two (-Art) groups of L1 Malay and L1 Chinese learners that the indefinite article *a* dominates in all environments suggesting an accuracy order of *a* > *the* > *Ø*.

However, Master (1997) argued that L2 learners of English seem to acquire the *zero* article first, although he warns that one cannot tell the difference between the *zero* article and the omission of the article. Master's data showed that *zero* article accuracy is close to 100% for the low-ability level participants, which then drops and rises to nearly 100 % again for the high-ability level participants. He further reported that the overuse of *zero* articles decreases with increased proficiency. However, the overuse of *zero* articles persists more than the overuse of the other articles. Liu and Gleason (2002) re-examined Master's data and offered a new interpretation of the overuse of the *zero* articles and underusing *a* and *the*. They found that 'this overuse of the *zero* articles and the underuse of *the* at the advanced stage would suggest that the two articles are acquired rather late.'

Liu and Gleason's (2002)hypothesis was supported by Young's (1996) data on the use of articles by Czech and Slovak learners of English; for a while, definiteness was not encoded at the early stages of acquisition; it persisted even at the more advanced stages. However, participants encoded indefiniteness using the indefinite article *a*t all proficiency levels with rising frequency as acquisition progressed.

Research findings show similarities in the problems facing EFL learners, some of which are believed to be more severe for learners from specific language backgrounds. The findings of comparative studies of first and second language acquisition are widely varied. Some morpheme studies report similar stages of development, while others report apparent variability in the order of acquisition of different groups. A third group yet limits the similarity to natural learning situations. Corder (1973) maintains that, unlike natural language learning, where learners make and test their hypotheses about the language, L2 learners in tutored situations follow an externally imposed syllabus.

Article choice is complicated, context-specific, and sometimes beyond simple rules. Therefore, to study English articles' accuracy order and underlying processes by Arab EFL learners, Bickerton's (1981) semantic wheel for noun phrase (NP) reference, marked by the features, [±Specific Referent (±SR)] and [±Assumed Known to the Hearer (±HK)] was employed in this study. In his model, English NPs are classified by two features of referentiality—namely, specific reference [±Specific Referent (±SR)] and Hearer's knowledge [±Assumed Known to the Hearer (±HK)]. These two aspects of referentiality thus give rise to four primary NP contexts that determine article use. The four basic NP contexts are herein denoted as uses of type 1 ([–SR, +HK], generics), type 2 ([+SR, +HK], referential definites), type 3 ([+SR, –HK], referential indefinite), and type 4 ([–SR, –HK], non-referential).

The lexical properties of the noun (singular or plural, mass or count) determine article choice from among the possibilities available in a given environment (Thomas, 1989). Table 1 presents a classification system for two binary features (Thomas, 1989). Nouns classified as [-SR +HK] are generic nouns marked with *a, the*, *Ø* (zero article). Nouns classified as [-SR -HK] are non-referential nouns marked with a*, Ø*. These articles use nouns that name a class to which another noun is asserted to belong or refer to an unspecified class member (Thomas, 1989). The [+SR -HK] feature includes referential indefinite nouns marked with a*, Ø*. The referent is identifiable not to the Hearer but to the speaker, who is entering the noun into the discourse for the first time. Finally, the [+SR +HK] feature includes referential definite nouns previously mentioned, are specified by entailment or definition, and are unique in all contexts, a given context, etc. (Thomas, 1989). These nouns are marked with *the* (see Huebner, 1983; Thomas, 1989 for more information on these four environments). In addition to these four types, idiomatic expressions and conventional uses were classified as Type 5, as in Butler (2002), Ekiert (2004), and Thomas (1989).

Table 1: Environments for the Appearance of *a, the,* and *Ø* (Adapted from Butler (2002), Huebner (1985), and Thomas (1989))

features	Article	Environment	Example
Type 1 [−SR,+HK]	*a, the, Ø*	Generic nouns	*Ø* Fruit flourishes in the valley. *The* Grenomian is an excitable person. *A* paper clip comes in handy.
Type 2 [+SR,+HK]	*the*	referential definites previous mention specified by entailment specified by the definition unique in all contexts unique in a given contexts	Pass me *the* pen. *The* idea of coming to the US was… I found a book. *The* book was… The first person to walk on *the* moon…
Type 3 [+SR,−HK]	*a, Ø*	referential indefinite first-mention nouns	Chris approached me carrying *a* dog. I keep sending *Ø* messages to him.
Type 4 [−SR,−HK]	*a, Ø*	Non-referential nouns attributive indefinite non-specific indefinite	Alice is *an* accountant. I should buy *a* new car. *Ø* Foreigners would come up with a better solution.
Type 5	*a, the, Ø*	Idioms Other conventional uses	*All of a sudden*, he woke up. *In the 1950s*, there weren't many cars. His family is now living *Ø hand to mouth.*

Bickerton's binary semantic system was the system put forward by Bickerton (1981) to categorize the articles according to semantic function. Several writers adopted this system to assess data such as that generated by the present study. In his work on Creole languages, Bickerton noticed that these language types invariably divide NPs according to notions of specific/non-specific, which he argues is an innate division that has significant implications for the language learner (Bickerton, 1981). Since English provides no such clear-cut marking of specific/non-specific items (generics, for example, may take any one of the three articles) and is governed by the additional criteria of supposed-known-to-listener and supposed-unknown-to-listener, Bickerton (1981) proposed the 'semantic space' for English articles which was later adapted by Huebner (cited in Parrish (1987) into "... a system of analysis that accounts for article use in all contexts, that is to say, in all pre-noun positions".

Much research has been conducted about the processes of L2 acquisition of English articles. Research on article acquisition in English language learning falls into two areas: pedagogy and its effectiveness on the one hand and the process of acquisition on the other. The present study examines the acquisition orders and underlying processes in terms of the accuracy of article use. Therefore, the discussion focuses on the three research questions:

1. What do the three measures, SOC, TLU, and UOC, reveal about the accuracy order of English

articles by Arab EFL learners?

2. In what semantic contexts (according to Bickerton's wheel model 1981) are the articles *the*, *a*, and *Ø* overgeneralized?
3. Which one of the four semantic contexts (Bickerton's 1981) is most difficult for L1 Arabic speakers to acquire and why?

This study may add another perspective to the current literature on learning and acquiring English articles. The findings of the study are expected to enable Arab EFL learners to be aware of the importance of the English article system, to avoid misuse, to be aware of the different linguistic features in the English and Arabic languages, and to try to trace the reasons behind inaccuracy, if any.

METHODOLOGY

Design

In this research, a quantitative research design was intended to investigate acquisition orders and underlying processes regarding article accuracy and use by undergraduate Arabic EFL learners. The research's design included the following data analysis techniques: SOC (Supplied in Obligatory Contexts), devised by Brown (1973); TLU (Target-Like Use), formulated by Pica (1983); and UOC (Used in Obligatory Contexts), formulated by Master (1987). The study was therefore devised in order to elicit data for these variables. (1) The two independent variables are article use (three articles: *the*, *a*, and *Ø*) and group (with three levels: advanced, upper-intermediate, and lower-intermediate). (2) The dependent variables are the scores for using *the*, *a*, and *Ø*, which SOC and TLU will measure. A repeated-measures design is necessary because the same groups of participants will be examined for their usage of each article type. Along with the multiple-choice cloze test, an oral production task was conducted to elicit data. This data allows for a comparison between the accuracy of article usage in both written and oral production.

Participants

Forty-five (45) undergraduate Arab EFL learners participated in this study. Thirty-nine (35 males and four females) participated in the written production, and the remaining six (4 males and two females) participated in the oral production. All participants took TOEFL (Test of English as a Foreign Language). According to the TOEFL scores, all the participants were divided into three groups: the Advanced group scores were in the range of 570-645 (M = 601.19, SD = 28.9); the Upper-Intermediate group, 500-560 (M = 11.62, SD = 19.0); and the Lower-Intermediate group, 400-490 (M = 455.15, SD = 27.4).

More specifically, sixteen (16) participants were at the advanced level (14 learners participated in the written production task, and only two participated in the oral production task), sixteen (16) were at the upper-intermediate level (14 of the learners took part in the written production task and two learners attended the oral production task) and thirteen (13) were at the lower-intermediate level (11 learners participated in the written production task and two learners took part in the oral production task). They were nineteen (19) to twenty-four (24) years old. The average age of first exposure to English was seven to fifteen years old (M = 11.62, SD = 1.40).

Instrumentation

Data were collected using two instruments; the first instrument is a written task, and the second is an oral production task. The written task was a multiple-choice cloze test. The cloze test, adapted from Master (1994), comprises fifty-eight (58) items in two parts: discrete sentences and a descriptive paragraph. The participants were asked to fill in the blanks by choosing the most appropriate article from among *a*, *an*, *the*, and *Ø* on an answer sheet. In addition, the subjects were asked not to consult dictionaries or textbooks during the exercise, nor to collaborate with anyone else, to ensure as much as possible that the data was a true reflection of the subjects' unaided ability.

As for validity, Lu (2001) mentioned that the Master's article test is considered to be a 'legitimate instrument' because the test covers the entire range of article usage, including the four semantic categories of (±SR) and (±HK). The same test was designed to test article usage for non-native speakers of English, so it is also suitable for Arab EFL learners. However, examining students' speech from different proficiency levels using both written and oral production tasks would have added further validity to the claims made in this research.

As for reliability, the Cronbach alpha (α) for this test in this study was .75. The estimate of .75 was close to .77 found by Lu (2001) which in its case was very close to the K-R20 of .79 reported for Master's (1994) pilot test that 75 L2 learners took at five different levels of the ESL courses. Following Lu (2001), the Cronbach α, instead of the K-R20 formula, was used in this study because the participants, with their TOEFL scores of 400-645, homogeneously scored above 50% accuracy on the article test, so the restricted variance of the test scores would affect the K-R20 estimate. Therefore, Cronbach α, a split-half procedure for internal consistency, was chosen to measure test reliability. Statistically, the K-R20 and Cronbach α values underestimate the actual reliability of the test (Brown, 1996), so this test with a Cronbach α of .75 can be accepted as reasonably reliable.

In the oral production task, the participants were told a story twice. Then they were asked to retell the same story. The story contained the four different semantic contexts (1) ([–SR, +HK], generics), (2) ([+SR, +HK], referential definites), (3) ([+SR, –HK], referential indefinite), and (4) ([–SR, –HK], non-referential) where the three articles *the*, *a*, and *Ø* could be used. The participants' oral production was recorded. Then a transcription of each participant's oral production was made. Data were analyzed using

the three measures: Brown's (1973) SOC (Supplied in Obligatory Contexts), Pica's (1983) TLU (Target-Like Use), and Master's (1987) UOC (Used in Obligatory Contexts). On average, it took approximately twenty to thirty (20 to 30) minutes to complete the test, although there was no time limit. ANOVA procedures were also performed to identify the acquisition orders. The oral data were compared with the written data to identify the accuracy of article usage in both models.

Data analysis procedure

The measures employed for data analysis were SOC (Supplied in Obligatory Contexts), devised by Brown (1973); TLU (Target-Like Use), formulated by Pica (1983); and UOC (Used in Obligatory Contexts), formulated by Master (1987).

SOC (Supplied in Obligatory Contexts) has been widely used in various morpheme studies (Andersen, 1978; Dulay & Burt, 1974; Hakuta, 1976; Larsen-Freeman, 1975; Lu, 2001; Master, 1987; Parrish, 1987; Thomas, 1989; Upor, 2021; Wong & Quek, 2007) to estimate an L2 learner's accuracy level. It was formulated by Pica (1983) as follows:

$$\text{SOC} = \frac{\text{Number of correct suppliances in obligatory contexts}}{\text{Number of obligatory contexts}}$$

However, SOC has been criticized for its failure to consider the over-suppliance of a morpheme in non-obligatory contexts (Andersen, 1978; Hakuta, 1976; Hatch, 1978; Pica, 1983). SOC will overestimate the learner's accuracy if the morpheme is over-supplied or overgeneralized. So, suppliance in non-obligatory contexts, i.e., morpheme overgeneralization, should be considered in the accuracy measure (Lu, 2001). Thus, Pica (1983) devised another measure which is TLU, to redress the inflation of SOC, and it was formulated as follows:

$$\text{TLU} = \frac{\text{Number of correct suppliances in obligatory contexts}}{(\text{Number of obligatory contexts}) + (\text{Number of suppliances in non-obligatory contexts})}$$

UOC (Used in Obligatory Contexts), which is used to measure article use, was used by Master (1987) as a complementary measure to observe the learner's overuse or underuse of the article. It was formulated as follows:

$$\text{UOC} = \frac{\text{The total number of suppliances in both obligatory and non-obligatory contexts}}{\text{Number of obligatory contexts}}$$

Like TLU, suppliances in non-obligatory contexts are also considered in UOC, so the learner's overall use of a particular morpheme can be inspected. UOC and SOC share the same denominator, so comparisons between accuracy and use can be easily spotted. Statistically, SOC and TLU cannot exceed 100%, but UOC can. So UOC can indicate overuse or underuse of the morpheme. On the other hand,

SOC indicates simple but potentially overestimated accuracy, TLU reveals a de-inflated estimate of accuracy level, and UOC shows learners' actual use or overuse of the morpheme (Lu, 2001).

None of the previous studies employed the improved accuracy measure, TLU. Parrish (1987), Thomas (1989), and Yamada and Matsuura (1982) looked at SOC, and Master (1987) used SOC and UOC. Only Lu (2001) and Wong and Quek (2007) used the three measures SOC, TLU, and UOC for data analysis.

The present study focused on the distinction in use between *the, a,* and *Ø*. Both *a* and *a* were tallied correctly in the indefinite article contexts, even if *a* was incorrectly replaced by *an* or vice versa, just as Lu (2001), Master (1987), and Thomas (1989) did for their studies.

First, frequencies of *the, a,* and *Ø* in each of the [±SR, ±HK] contexts were counted for each participant. Next, the SOC, TLU, and UOC for *the, a,* and Ø were calculated for each participant and proficiency group for both the written and oral production tasks. Then, each group's means of SOC and TLU were compared to identify the acquisition order in terms of article accuracy. Likewise, the means of UOC for each group were compared to examine the overuse or underuse of the articles. The means of SOC and TLU for both written and oral tasks were compared to identify the acquisition order in terms of article accuracy for both modes. Furthermore, the means of UOC for written and oral tasks were compared to examine the overuse or underuse of the articles. Besides, ANOVA procedures and Scheffé and Bonferroni follow-up tests were performed to identify the accuracy orders.

RESULTS AND DSICUSSION

Comparisons of Means of SOC for the Written and the Oral Production Tasks

The SOC average for the written task shows that the performance of the advanced level is the highest compared to the other groups, advanced > upper-intermediate > lower-intermediate (77.60% > 59.79% > 47.04%), with an overall average of 61.48%. Similarly, the SOC measure average for the oral production task shows that the performance of the advanced level was the highest compared to the other two groups, advanced > upper-intermediate > lower-intermediate (86.72% > 74.91% > 73.14%), with an overall average of 78.26%.

The means for the articles in the written and the oral production tasks are presented in Fig. 1. The lines for the three articles are more comprehensive apart at *the* and *Ø*, which indicates that those greater mean differences result in a significant interaction. However, the lines for the written and the oral production tasks nearly meet each other at *a*, which indicates that the narrow mean difference of 1.28% may not be significant. In addition, the lines for the articles are not systematic in a similar pattern. Also, the mean differences between *the* and *a*n are10.93% and 13.74% for the written and the oral

production tasks, respectively, and the mean differences between *a* and *Ø* are 44.19% and14.68% for the written and the oral production tasks, respectively.

Therefore, the SOC measure reveals an accuracy order of *a* > *the* > *Ø* for the written task and an acquisition order of *the* > *a* > *Ø* for the oral production task. The obligatory use of *a* is acquired earlier than *the* zero article *Ø* for the written task. In contrast, the obligatory use of a is acquired earlier than *a* and, finally, the zero article *Ø* for the oral production task. Overall, it seems that the performance of the participants in the oral production task 92.31% > 78.57% > 63.89% is better than that of the participants in the written task 68.92% > 79.85% > 35.66%.

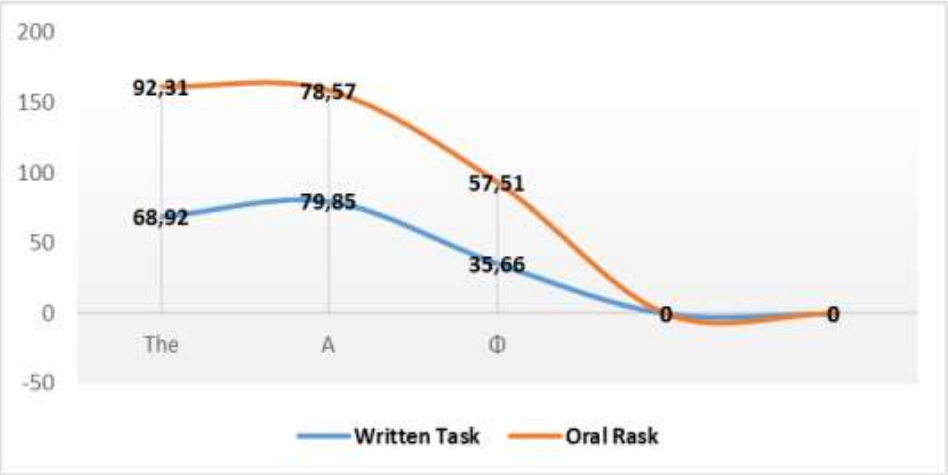

Figure 1: Comparison of Means of SOC for the Written and the Oral Production Tasks

The source table for the two-way repeated-measures ANOVA for the written task, shown in Table 2, indicates that the main effects for group and article use are both significant, $p < .05$, power = 1.00 (The recommended power is .80 up to 1.00, see Kirk, 1982). The interaction effect for article use and the group was also significant, $p < .05$, power = 1.00.

Table 2. Two-way repeated-measures ANOVA for group and article use on SOC in the written task

Source	SS	df	MS	F	P	Partial Eta²	Power
Between-Subjects Effects							
Group	18292.328	2	9146.164	70.921	.000	.798	1.000
Error	4642.678	36	128.963				
Within-Subjects Effects							
Article	39887.419	2	19943.710	112.484	.000	.758	1.000
Article × Group	3817.419	4	954.355	5.383	.001	.230	.965
Error	12765.780	72	177.303				
Total	79405.62						

The one-way ANOVA for the group effect on the overall SOC means of the three articles combined was significant, $p < .0125$, power= 1.00. The Scheffé post hoc results also indicated that the

mean differences between all possible pairs of the groups were significant, $p < .0125$. These results suggest that article accuracy increases with proficiency in a similar pattern across the groups.

In addition, the one-way repeated-measures ANOVA for the article's effect on the overall SOC means of the three groups combined was also found to be significant, $p < .0125$, power= 1.00. The Bonferroni test in Table 3 showed that all pairwise comparisons between the articles were significant. Therefore, it can be confirmed that the SOC measure reveals the accuracy order of *a* > *the* > *Ø* across the groups in the written task.

Table 3. Bonferroni test for pairwisepairwise comparisons of SOC in the written task

SOC (I)	SOC (J)	Mean Difference (I-J)	Standard Error	*p*
the	*a*	18.2181	2.47813	.000*
	Ø	30.9699	2.64169	.000*
a	*the*	-18.2181	2.47813	.000*
	Ø	12.7518	2.64169	.000*
Ø	*a*	-30.9699	2.64169	.000*
	the	-12.7518	2.64169	.000*

Comparisons of Means of TLU for the Written and the Oral Production Tasks

The TLU average for the written task shows that the performance of the advanced group is more nativelike than the upper-intermediate group, which is more nativelike than the lower-intermediate group (62.84% > 38.83% > 28.83%) with an overall average of 43.5%.

The results for the three groups are parallel and systematic, and no significant interaction can be found, which reveals only one acquisition pattern for TLU for the three groups, which is not similar to that of the SOC measure. The TLU measure reveals an acquisition order of *the* > *a* > *Ø* across the groups, which means that the use of *the* is more target-like than *a*, which is more target-like than *Ø*.

Likewise, the TLU measure average for the oral production task shows that the performance of the advanced group is more nativelike than the upper-intermediate group, which is more nativelike than the lower-intermediate group (76.15% > 63.55% > 60.35%) with an overall average of 66.68%. The results for the three groups are parallel to some extent, with an overlap that can be found between the upper-intermediate and the lower-intermediate groups towards the performance of *Ø*. However, the three groups have only one accuracy pattern for TLU, similar to that of SOC.

Figure 2 shows that the lines for the three groups are parallel and systematic. As such, no significant interaction can be found, which suggests that participants in the written and oral production tasks might have only one acquisition pattern for TLU, which is different from that of SOC. Therefore,

it is hypothesized that the TLU measure reveals an accuracy order of *the* > *a* > *Ø*, 57.73% > 43.28% > 29.50%, and 75.39% > 67.16% > 57.51% for the participants in the written and the oral production tasks, respectively, which means that the use of *the* is more target-like than *a*, which is more target-like than *Ø*.

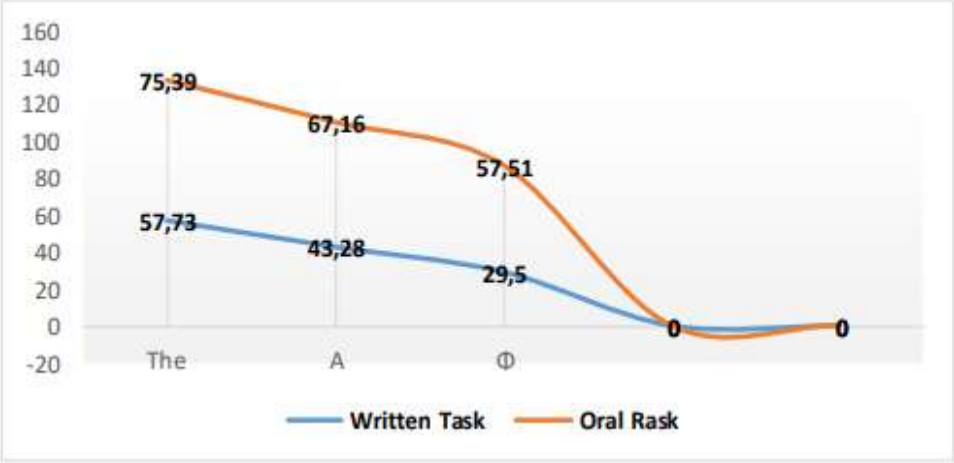

Figure 2: Comparison of Means of TLU for the Written and the Oral Production Tasks

As shown in Table 4, the two-way repeated-measures ANOVA for the main effects of group and article on the means of TLU was significant, $p < .05$, power = 1.00. However, the interaction effect for the article and group was insignificant, $p > .025$, perhaps because of a lack of power (power = .390).

Table 4. two-way repeated-measures ANOVA for group and article on TLU in the written task

Source	SS	df	MS	F	P	Partial Eta²	Power
Between-Subjects Effects							
Group	24488.179	2	12244.089	76.968	0.000	0.810	1.000
Error	5726.889	36	159.080				
Within-Subjects Effects							
Article	15686.184	2	7843.092	108.688	.000	.751	1.000
Article × Group	378.802	4	94.700	1.312	.274	.068	0.390
Error	5195.626	72	72.161				
Total	51475.68						

Similarly, to ensure the accuracy order of *the* > *a* > *Ø* for the TLU measure, two one-way ANOVAs and the Scheffé and Bonferroni follow-up tests were administered, respectively. The one-way ANOVA for the group effect on the overall TLU means of the three articles combined was significant, $p < .05$, power = 1.00. The Scheffé post hoc test was then used. The results also showed that the mean differences between all possible pairs of groups were significant, $p < .05$. Next, the one-way repeated-measures

ANOVA for the particle effect on the overall TLU means of the three groups combined was also found to be significant, $p < .05$, power = 1.00. The Bonferroni test was then used, and the results in Table 5 show that all pairwise pairwise comparisons between the articles were significant, $p < .05$. Therefore, it can be confirmed that the TLU measure reveals an acquisition order, *the* > *a* > Ø, across the groups.

Table 5. Bonferroni test for pairwise pairwise comparisons of TLU in the written task

TLU (I)	TLU (J)	Mean Difference (I-J)	Standard Error	*p*
the	*a*	24.8738	2.75232	.000*
	Ø	34.4640	2.93398	.000*
a	*the*	-24.8738	2.75232	.000*
	Ø	9.5902	2.93398	.007*
Ø	*a*	-34.4640	2.93398	.000*
	the	-9.5902	2.93398	.007*

Comparisons of Means of UOC for the Written and the Oral Production Tasks

The UOC measure discloses that *a* is overused at 127.14%, 207.14%, and 199.09% for the advanced, upper-intermediate, and lower-intermediate groups. The definite article is overused at 103.14% for the advanced group and is underused at 84.00% and 80.73% for the upper-intermediate and the lower-intermediate groups, respectively. On the other hand, *Ø* is underused at values of 73.08%, 34.07%, and 60.84% across the groups.

Moreover, the data shows that the value for the overuse of *a* is more for the upper-intermediate group (207.14%) than for the lower-intermediate group (199.09%), which decreases further as proficiency increases with the advanced group (127.14%). The value for underuse of *Ø* for the advanced group is 73.08%, which decreases with the lower-intermediate and the upper-intermediate groups (60.84% and 34.07%), respectively. Because the UOC measure for *the* is not very stable, that is, it was overused with the advanced group (103.14%) and underused with both the upper-intermediate and the lower-intermediate groups (80.73% and 84.00%) respectively, it makes sense that UOC for *a* goes up and down, that is it increases then decreases (127.14% ↑ 207.14% ↓ 199.09%) while UOC for *Ø* goes down and up that is it decreases then increases (73.08% ↓ 34.07% ↑ 60.84%) across the groups. The upper-intermediate and the lower-intermediate groups overuse *a* with a mean difference of 8.05%. This suggests that the upper-intermediate and lower-intermediate groups tend to overuse *a* and underuse *Ø*. In contrast, the advanced group tends to overuse *a* and thus increase the use of *Ø* towards target-like use. Thus, we can say that the lower-intermediate and the upper-intermediate groups have more difficulty in the article choice of *a* and *Ø*. Furthermore, in most cases, they tend to overuse in *Ø* contexts.

The UOC measure for the oral production task discloses that it is overused at values of 103.85%, 123.08%, and 119.23% for the advanced, upper-intermediate, and lower-intermediate groups, respectively, with the advanced speakers being better in their performance. The indefinite article *a* is overused for the advanced group at the rate of 114.29% and underused at 92.86% and 78.57% for the upper-intermediate and lower-intermediate groups, respectively. The zero article *Ø* is underused (83.33%, 58.33%, and 83.33%) across the groups. The fluctuation of the overuse between *the* upper-intermediate and the lower-intermediate groups and the overuse of *a* for the advanced group can be easily noted. This has resulted in an inflated SOC rate. Because the SOC measure does not consider overuse, the means of SOC are 96.15% and 88.46% for the upper-intermediate and the lower-intermediate groups, respectively. These values are much higher than the means of TLU, 75.76% and 67.65% for the upper-intermediate and lower-intermediate groups, respectively. On the other hand, the means of SOC for *a* is 92.86% for the advanced group, which is much higher than the means of TLU for *a*, which is 76.47%.

Moreover, the overuse increases with the upper-intermediate group (123.08%) than with the lower-intermediate (119.23%) but decreases as proficiency increases with the advanced group (103.85%). Only the advanced group overused *a* (114.29%), and the upper-intermediate and the lower-intermediate groups underused the indefinite article at 92.86% and 78.57%, respectively. Because UOC for *a* is not very stable (78.57%–114.29%), it makes sense that UOC for *the* goes up and down, that is, increases then decreases (103.85% ↑ 123.08% ↓ 119.23%) while UOC for *Ø* goes down and up, that is decreases then increases (83.33% ↓ 58.33% ↑ 83.33%) across the groups. The upper-intermediate and lower-intermediate groups overuse *the*m with a mean difference of 3.85%. This suggests that the upper and lower intermediate groups tend to overuse *the* underuse *Ø*. In contrast, the advanced group tends to decrease overuse of *the* target-like use. Thus, we can say that the lower-intermediate and the upper-intermediate groups have more difficulty in article choice between *the* and *Ø*. Furthermore, in most cases, they tend to overuse *the* in *Ø* contexts. Overall, the UOC measure indicates that TLU is a more reliable acquisition measure.

In Figure 3, the UOC measure discloses that *a* is overused at a value of 177.79% in the written task and underused at 95.24% in the oral production task. In contrast, it is overused at 115.39% in the oral production task and underused at 89.29% in the written task. On the other hand, *Ø* is underused at values of 56.00% and 75.00% in both tasks. Overgeneralization of the written and oral production tasks, which can be easily noted in Figure 3, resulted in an inflated SOC rate. Because the SOC measure does not consider overuse, the means of SOC for (79.85% see Figure 1) are much higher than the mean of TLU for (43.28%, see Figure 2) in the written task, and the means of SOC for *the* (92.31%, see Figure 1) are much higher than the means of TLU for *the* (75.39%, see Figure 2) in the oral production task. This proves that the SOC measure tends to overestimate accuracy and needs to be revised by a more

accurate measure. This was the TLU measure, as Pica (1983) has suggested. Recall that the eta² values in the source tables for the two-way repeated-measures ANOVAs (see Tables 2 and 4) also suggest that the TLU measure reflects proficiency better than SOC, so TLU is probably a more reliable accuracy measure. Therefore, the accuracy order of the > a > Ø identified by TLU is probably more reliable than the a > the > Ø order, which SOC identified.

Figure 3: Comparison of Means of UOC for the Written and the Oral Production Tasks

Overall, the means comparisons of SOC, TLU, and UOC for the written and oral production tasks show that the performances of the three groups in the oral task are better than those in the written task.

Use of the in [+SR] and [+HK] Contexts

Concerning Research Question 2, since there was disagreement in the previous studies about whether *the* is overgeneralized to all specific nouns in [+SR] contexts (Parrish, 1987; Thomas, 1989) or all known nouns in [+HK] contexts (Huebner, 1983a; Master, 1987), the contexts for *the*-overgeneralization are first examined for both the written and the oral production task. Table 6 shows SOC, TLU, and UOC for *the* [+SR] and [+HK] contexts. The UOC indicates that *the* is slightly overused in the [+SR] contexts at the rate of 103.70% for the advanced group and is underused at the rate of 87.76% and 82.68% for the upper-intermediate and lower-intermediate groups, respectively. In the [+HK] contexts, *the* is also slightly overused at the rate of 101.29% for the advanced group and is underused at 83.43% and 54.43% for the upper-intermediate and lower-intermediate groups, respectively. Examining the use of *the* in both [+SR] and [+HK] contexts shows that *the* is slightly overused in [+SR] contexts (103.70%, 87.76%, and 82.68%) compared to [+HK] contexts (101.29%, 83.43%, and 54.43) for the three group respectively (see Table 6). Therefore, the present study is consistent with Parrish and Thomas, rather than Huebner and Master: *the* is more associated with [+SR] than with [+HK] contexts. The juxtaposition of SOC, TLU, and UOC in Table 5 suggests that *the*y are overused in some [+SR] contexts where it is not supposed to occur. For example, *I once read the story* [a story] *about the courage and strength of these wild pigs* (item 55). On the contrary, the use of *the* in

[+HK] contexts is stable to some extent on the three measures within each group, and this indicates that *the* was not overgeneralized in [+HK] contexts.

On the other hand, the UOC measure for both [+SR] and [+HK] contexts shows that the definite article is underused at values of 82.68% and 54.43% for the lower-intermediate group and 87.76% and 83.43% for the upper-intermediate group for [+SR] and [+HK] contexts respectively. On the other hand, the UOC measure shows that it is overused at values of 103.70% and 101.29% for the advanced group for [+SR] and [+HK] contexts, respectively. However, the SOC and TLU measures in Table 6 are stable in both [+SR] and [+HK] contexts for the three groups; for example, the SOC values for the upper-intermediate group are 70.07% and 65.14%, and the TLU values are 59.54% and 55.07% for both contexts (see Table 6).

Table 6. Use of the in [+SR] and [+HK] Contexts

	Advanced (*n* = 11)			Upper-Intermediate (*n* = 19)			Lower-Intermediate (*n* = 15)		
Context	SOC	TLU	UOC	SOC	TLU	UOC	SOC	TLU	UOC
[+SR]	86.05	78..41	103.70	70.07	59.54	87.76	57.14	45.52	82.68
[+HK]	78.14	75.68	101.29	65.14	55.07	83.43	54.18	43.57	54.43

[+SR] combines [+SR, +HK] and [+SR, −HK];
[+HK] combines [+SR, +HK] and [−SR, +HK].

Since the feature [+SR, +HK] demands the article, the key feature that causes *the* overgeneralization in [+SR] contexts lies in [+SR, −HK], where *a* or *Ø* is required. So the marked [+SR, −HK] contexts must be examined closely on an item-by-item basis to discover the difficulties underlying article choice. Table 7 presents the frequencies and item facilities of *the*, *a*, and *Ø* in [+SR, −HK] contexts. It also shows that *the* was misused in *a* context at the rate of 6.25%, 2.68%, and 11.36% for the advanced, upper-intermediate, and lower-intermediate groups, respectively, and that *the* was misused in *Ø* contexts at a much higher rate of 23.21%, 29.17%, and 37.12% for the three groups, respectively. Generally, based on item facility (IF), Items 12, 21, 27, 46, and 55 (IF = .27 – 1.00) are more challenging to use, and Items 25, 32, 37, 45,51, 52, 53, and 58 (IF = .00 – .71) are more difficult for *Ø* usage across the groups.

What was found about the misuse or overgeneralization of *the* in *a* or *Ø* contexts is that Arabic speakers had difficulty in distinguishing [±HK]. For example, in Item 12 (*earns $25,000 a year*), Item 25 (*Ø bottles of Pepsi*), Item 27 (*a length of 12 meters*), Item 37 (*Ø large, catlike tracks)*, Item 46 (*a favorite food*), Item 51 (*in Ø bands of fifteen to twenty*) and Item 52 and 53 (have *Ø great courage and*

Ø strength), Arabic speakers tended to substitute *the* for *a* or Ø in *of*-phrase structures, which is in tune with the findings in Lu (2001) and Takahashi (1997). In other words, Arabic speakers, like Chinese and Japanese learners, seemed likely to overgeneralize *the* in-*of*-phrase structures, regardless of whether the context is [+SR, +HK] or [+SR, −HK]. Moreover, the lower-intermediate group misused *the* for *a* or *Ø* (IF = .18 – .64) in Item 5 (*drink Ø water with meals.*), Item 15 (*smoke Ø cigarettes*) and Item 55 (*I once read a story*), but the advanced group had more knowledge (with correspondingly higher IFs of .71 – 1.00 for the same items) to distinguish [+SR, −HK] from [+SR, +HK] contexts. Therefore, it can be said that overgeneralization of *the* in [+SR, −HK] contexts can be attributed to the speaker's difficulty in distinguishing the [±HK] contexts.

Table 7. Frequencies and Item Facilities (IF) of the, a, and Ø in [+SR, −HK] Contexts for the Written Task

Target Article	**Item #**	**Advanced (*n* = 14)**				**Upper-Intermediate (*n* = 14)**				**Lower-Intermediate (*n* = 11)**			
		IF	*the*	*a*	Ø	IF	*the*	*a*	Ø	IF	*the*	*a*	Ø
a	1	0.93	0	13	1	1.00	0	14	0	0.91	0	10	1
	2	0.93	0	13	1	0.86	0	12	2	0.91	0	10	1
	11	1.00	0	14	0	1.00	0	14	0	1.00	0	11	0
	12	1.00	0	14	0	0.93	0	13	1	0.27	0	3	8
	21	1.00	0	14	0	0.93	1	13	0	0.45	4	5	2
	27	0.71	2	10	2	0.79	0	11	3	0.45	2	5	4
	46	0.79	1	11	2	0.93	0	13	1	0.73	1	8	2
	55	0.71	4	10	0	0.79	2	11	1	0.64	3	7	1
Total	***k*=112**		**7**	**99**	**6**	***k* =112**	**3**	**101**	**8**	***k* =88**	**10**	**59**	**19**
	Percentage(%)		**6.25**	**88.39**	**5.36**		**2.68**	**90.18**	**7.14**		**11.36**	**67.05**	**21.59**
Ø	5	0.93	0	1	13	0.79	2	1	11	0.18	4	5	2
	15	1.00	0	0	14	0.36	3	6	5	0.27	1	7	3
	23	0.71	4	0	10	0.71	2	2	10	0.45	4	2	5
	25	0.71	3	1	10	0.21	5	6	3	0.18	6	3	2
	28	0.73	6	0	8	0.21	9	2	3	0.27	8	0	3
	32	0.07	7	6	1	0.14	1	11	2	0.27	1	7	3
	37	0.29	0	10	4	0.00	1	13	0	0.27	2	6	3
	45	0.64	5	0	9	0.14	12	0	2	0.27	7	1	3
	51	0.79	3	0	11	0.07	4	9	1	0.18	8	1	2
	52	0.36	2	7	5	0.07	1	12	1	0.27	1	7	3
	53	0.57	2	4	8	0.29	5	5	4	0.27	2	6	3
	58	0.43	7	1	6	0.29	4	6	4	0.09	5	5	1
Total	***k*=168**		**39**	**30**	**99**	***k* =168**	**49**	**73**	**46**	***k*=132**	**49**	**50**	**33**
	Percentage(%)		**23.21**	**17.86**	**58.93**		**29.17**	**43.45**	**27.38**		**37.12**	**37.88**	**25.00**

In addition, *the* was not the only article misused in [+SR, −HK] contexts; *a* was also misused for *Ø*, or *Ø* for *a*. As shown in Table 6, *a* was misused in *Ø* contexts at values of 17.86%, 43.45%, and 37.88% for the Advanced, Upper-Intermediate, and Lower-Intermediate groups, respectively. Conversely, *Ø* was misused in *a* context at values of 5.36%, 7.14%, and 21.59% for the three groups,

respectively. For instance, in Item 46 (*a favorite food*), Item 27 (*a length of 12 meters*), and Item 52 (*have Ø great courage*), Arabic speakers tended to substitute *Ø* for *a*, or vice versa, even when they could distinguish [+SR, −HK] (i.e., *a* or *Ø*) from [+SR, +HK] (i.e., *the*) contexts. This suggests that Arabic speakers have difficulty judging the countability of noun phrases, which results in the misuse of *a* or *Ø* in [+SR, −HK] contexts. Overall, it can be concluded that Arabic speakers' difficulties in article choice lie in the distinctions between [±HK] and [±Countability].

Comparison of Accuracy in Article Use for the Four Contexts [±SR, ±HK] in Both Tasks

Table 7 compares the four contexts [±SR, ±HK] for both the written and the oral production tasks. In type1 context [−SR, +HK], which refers to the generic use of articles, the scores of the three groups are 75.00%, 50.00%, and 66.67% for the oral production task and the scores are 71.43%, 26.79%, and 19.32% for the written task for the advanced, upper-intermediate, and lower-intermediate respectively. Thus, the participant's performance in the oral task is higher than that of the participants in the written task. Moreover, the type 1 context reveals a significant difference of 47.35% for the lower-intermediate group and a difference of 23.21% for the upper-intermediate group in both tasks. However, this difference reduces to 3.57% for the advanced group. These scores indicate that the ability to distinguish the [−SR, +HK] context increases and becomes more stable with the increase in the proficiency level (see Table 8).

In type 2 context [+SR, +HK], which refers to referential definites, the performance in the oral production task is also higher with scores of 92.31%, 96.15%, and 88.46% compared to the scores in the written task- 86.05%, 70.07% and 57.14%- for the three groups respectively. Once again, the difference between the two tasks for the Lower-intermediate group is 31.32%. However, this difference decreases as the learners' proficiency increases to 26.08%, then 6.26%, with the upper-intermediate and the advanced group, respectively.

In type 3 context [+SR, -HR], which refers to referential indefinite, the performance of the oral task is higher at scores of 80.00%, 70.00%, and 70.00% than that of the written task-73.66%, 58.78%, and 46.03%. The difference between the two tasks for the lower-intermediate group is 23.97%, and this difference decreases as proficiency increases to 11.22% and 6.34% in the upper-intermediate and advanced groups, respectively.

The performance of the oral task is higher in the type 4 context [−SR, −HK], which refers to non-referential, at scores of 100.00%, 100.00%, and 50.00%, while the scores of the written task are 82.14%, 78.57%, and 59.09% for the advanced, upper-intermediate and lower- intermediate groups respectively.

Thus, the participant's scores in the oral production task are higher in all four contexts, but this is not necessarily true. Moreover, the participants' accuracy of article use across the groups increases with proficiency and becomes more stable with the advanced groups (see Table 8).

Overall, the accuracy order of the four contexts in both tasks is as follows: Type 4 > Type 2 > Type 3 >Type 1. In other words, Arabic speakers find type 4 easier than type 2 and type 2 easier than type 3, which is easier than type 1. Thus, type 3 and type 1 are the most challenging contexts for L1 Arabic speakers to acquire.

Table 8. Comparison of accuracy in article use for the four contexts [±SR, ±HK] in the written and the oral tasks

		Advanced	Upper-Intermediate	Lower-Intermediate
Type 1 [−SR, +HK]	Written Task	71.43%	26.79%	19.32%
	Oral Task	75.00%	50.00%	66.67%
Type 2 [+SR, +HK]	Written Task	86.05%	70.07%	57.14%
	Oral Task	92.31%	96.15%	88.46%
Type 3 [+SR, −HK]	Written Task	73.66%	58.78%	46.03%
	Oral Task	80.00%	70.00%	70.00%
Type 4 [−SR, −HK]	Written Task	82.14%	78.57%	59.09%
	Oral Task	100.00%	100.00%	50.00%

DISCUSSION

In light of the comparison between the written and oral production tasks, the present study is consistent with that of Tarone (1985) and Tarone and Parrish (1988). To some extent, article accuracy was much better in spontaneous oral production than in written tasks. In other words, participants in the oral production task produced lower error rates than those in the written task. However, the accuracy within one type of article would change across different tasks.

To sum up, the discussion for *research Question 1*, the three measures, SOC, TLU, and UOC, reveal the following about the accuracy order of the English articles:

1. SOC reveals the order, *a* > *the* > *Ø* for the written task and *the* >*a* > *Ø* for the oral production task,
2. TLU informs us of the accuracy order in both tasks, the > a > Ø.
3. UOC confirms the same patterns proposed by Chaudron and Parker (1990), Huebner (1983a), and Master (1987), that is, *Ø* first goes through a flooding stage. Then a trickling stage for hypothesis testing experiences a U-shaped behavior for the written task, highlighted by an

overgeneralization process, and follows *the* to undergo the U-shaped overgeneralization process as well.

As to *Research Question 2*: In what semantic contexts (Bickerton, 1981) are the articles, *the, a,* and *Ø,* overgeneralized? This investigation should help identify the areas of difficulty underlying article choice for Arab EFL learners. As revealed in Table 5, the present study supports Parrish and Thomas' claim that *the* is associated with [+SR] contexts rather than with [+HK]. So, [+SR, −HK] context was recognized as the key marked feature that causes *the* overgeneralization. Similarly, there is evidence in research on both L1 and L2 acquisition of articles that shows overgeneralization of *the* [+SR, −HK] contexts: for example, in Cziko (1986), L1 children could distinguish [+SR, +HK] and [−SR, −HK] by using *the* and *a* correctly at an early stage, but they tended to substitute *the* for *a* with the first-mention NPs in [+SR, −HK] contexts. As for L2 acquisition of articles, Suniphan (2022) concluded that [+SR, +HK] context, which requires *the* as a correct article, was used most accurately by Malaysian ESL learners. Zhao and Shirai's (2022) study indicated that definite articles show a significantly higher type ratio than indefinite, proposing that learners develop a more complex and heterogeneous profile of definite article usage. Parrish (1987) and Thomas (1989) found that *the* was overgeneralized prominently in [+SR, −HK] contexts. In Chaudron and Parker (1990), Japanese learners overgeneralized *the* to a greater degree in first-mention [+SR, −HK] contexts. Also, Takahashi's (1997) study shows that Japanese college students *used the* instead of *a*n in specific structures in [+SR, −HK] contexts, such as indefinite prepositional phrases and indefinite relative clauses. For example, in the sentence "Robert is listening to *a* record of Mozart music." *a* tended to be mistakenly replaced by *the* to mark a specific referent that is assumed unknown to the Hearer (i.e., [+SR, −HK]). Moreover, recall that the participant in Huebner (1983a) withdrew from [+SR, −HK] contexts at the final stage. Therefore, it can be concluded that *the* tends to be overgeneralized in [+SR, −HK] contexts. In addition, Tarone and Parrish (1988) found that L2 learners' accuracy in article use was significantly lower in [+SR, −HK] contexts than in [+SR, +HK] and [−SR, +HK] contexts. So the feature [+SR, −HK], positioned right between the two apparent features [+SR, +HK] and [−SR, −HK] in Bickerton's semantic wheel model, is much more marked and problematic for both L1 and L2 learners.

In addition, based on an item-by-item investigation in [+SR, −HK] contexts, the areas of difficulty underlying article choice for the learners in this study were identified: Arab EFL learners misuse *the* for *a* or *Ø* because they have difficulty distinguishing [+SR, −HK] from [+SR, +HK] contexts (i.e., the distinction of [±HK]). Even when they can distinguish [±HK], they misuse *a* for *Ø,* or vice versa, due to their difficulty distinguishing [±Countability]. Recall that the zigzag pattern of usage in Master (1987) was ascribed to the judgment of [±Countability]. Master (1997; 1987) points out that L2 speakers fail to use *a* or Ø correctly because they have difficulty judging the countability of noun phrases. Likewise, Yoon (1993) found that Japanese learners had trouble with the article choice between

a and *Ø* due to insufficient knowledge of countability. Moreover, in Bickerton's semantic wheel model, *a* and *Ø* share the same contexts but differ in the feature [±Countability]. Also, it is evident in Master's binary system that *a* and *Ø* are used to mark the feature "classification," whereas *the* is used to mark the feature "identification." In other words, [±Countability] is the subset underlying the choice of *a* or *Ø* within the same feature "classification." To sum up, Arab EFL learners' difficulties in the distinctions of [±HK] and of [±Countability] echo Master's (1997) assertion that three elements are required for article choice: "In the article system, the elements are countability, number, and definiteness, which must all be considered in arriving at the correct choice of the article."

Regarding *Research Question 3*, several facts emerge from the data in Table 7. When the accuracy of article use for the four contexts [±SR, ±HK] in the written task is compared with that in the oral production task, certain similarities become apparent. The accuracy order of the four contexts in both tasks is as follows:

Type 4 [−SR, −HK] > Type 2 [+SR, +HK] > Type 3 [+SR, −HK] > Type 1[−SR, +HK]

That is to say that Arab EFL learners acquire type 4 easier than type 2 and acquire type 2 easier than type 3, which is more accessible than type 1. Thus, type 4 and type 2 are easier than type 3 and type 1, which are the most challenging contexts to acquire. Different types of errors are identified in these four contexts.

CONCLUSION AND RECOMMENDATIONS

Conclusion

The results suggest that the majority of errors made by the groups are the result of common learning processes. Therefore, the present study hopes to shed some light on article pedagogy in view of the accuracy order, patterns in acquisition processes, and actual difficulties in article choice for Arab EFL learners. Although the results achieved in this study are sound and significant, more research is needed. A longitudinal study using a larger sample size throughout the study may prove invaluable for these purposes, incorporating oral as well as written data in the analysis.

Regarding the measure for article acquisition, although TLU is the best accuracy measure, none of the previous studies employed TLU: Parrish (1987), Thomas (1989), Yamada and Matsuura (1982) used SOC only, and Master (1987) used SOC and UOC, but not TLU. Only Lu (2001) and Wong and Quek (2007) used all three measures. It is hoped that TLU will be employed in future studies on the article or other morpheme acquisitions. The patterns of article acquisition processes (i.e., the flooding-then-trickling process for *Ø*, and the U-shaped behavior for *the* and then *a*) need further confirmation in future research. In addition, due to a lack of data from beginners and low-level learners, it is hard to know definitively whether Arabic speakers at different proficiency levels exhibit the same acquisition order,

as found in Thomas (1989) and this study, or whether each level exhibits a different acquisition order, as found in Master (1987) and Yamada and Matsuura (1982).

Moreover, a larger number of participants in the oral production task should be used in continued research to compare the results with those obtained from a written cloze test, as was used in the present study. Tarone and Parrish (1988) found that L2 learners' accuracy rates for article use varied when different tasks were performed. Therefore, further studies should be conducted with lower-level Arab EFL learners and with a large sample of oral tasks to build a complete profile of article acquisition for Arab EFL learners.

Implication

As previously stated, the findings assume that the English article system is so complex that instructors should devote more attention to it. Instructors and students should know the problematic areas they may encounter in their teaching and learning contexts. Therefore, emphasizing quality over quantity in course materials would help students avoid these drawbacks through reinforcement and practice. Moreover, exposing learners to authentic materials to improve observing as a learning assumption would be a good teaching technique to overcome these error types. Likewise, students might learn the English article system more successfully if audio-visual aids are introduced.

BIO-PROFILE:

Dr. Muneera Muftah is an Associate Professor in the Department of English at the Faculty of Arts, Thamar University, Yemen. Her main research interests include translation technologies, syntactic and morphological mental representation and development, vocabulary development in SLA, generative syntax and morphology, discourse studies, and second language assessment. Her current research focuses on information and communication technologies (ICTs) in English language teaching and learning. She has published numerous articles related to these areas in scholarly journals. Corresponding email: munmef5@gmail.com

REFERENCES

Abudalbuh, M. (2016). The Acquisition of English Articles by Arabic L2-English learners: A Semantic Approach. *Arab World English Journal*, *7*(2), 104–117. https://doi.org/10.24093/awej/vol7no2.7

Ali Muftah, M. Y., & Eng, W. B. (2011). The Acquisition of English be Auxiliary and Thematic Verb Constructions by Adult Arab ESL Learners. *International Journal of English Linguistics*, *1*(2), 91. https://doi.org/10.5539/ijel.v1n2p91

Andersen, R. W. (1978). An implicational model for second language research. *Language Learning*, *28*(2), 221–282. https://doi.org/10.1111/j.1467-1770.1978.tb00134.x

Bataineh, R. (2005). Jordanian undergraduate EFL students' errors in the use of the indefinite. *The Asian EFL Journal Quarterly*, *7*(1), 56–76.

Bickerton, D. (1981). Roots of language. Karoma. *Ann Arbor*, *10*, 268–284.

Brown, J. D. (1996). Testing in language programs. *Upper Saddle River, NJ: Prentice-Hall*.

Brown, R. (1973). *A first language: The early stages*. Harvard University Press.

Butler, Y. G. (2002). Second language learners'theories on the use of English articles: An Analysis of the Metalinguistic Knowledge Used by Japanese Students in Acquiring the English Article System. *Studies in Second Language Acquisition*, *24*(3), 451–480.

Chaudron, C., & Parker, K. (1990). Discourse markedness and structural markedness: The acquisition of English noun phrases. *Studies in Second Language Acquisition*, *12*(1), 43–64.

Chomsky, N. (1968). Language and Mind New York. *Brace and World*.

Corder, S. P. (1973). The elicitation of interlanguage. *Errata: Papers in Error Analysis*, 36–48.

Cziko, G. A. (1986). Testing the language bioprogram hypothesis: A Review of children's acquisition of articles. *Language*, *62*(4), 878. https://doi.org/10.2307/415175

Dulay, H. C., & Burt, M. K. (1974). Natural sequences in child second language acquisition 1. *Language Learning*, *24*(1), 37–53. https://doi.org/10.1111/j.1467-1770.1974.tb00234.x

Ekiert, M. (2004). Acquisition of the English article system by speakers of Polish in ESL and EFL settings. *Studies in Applied Linguistics and TESOL*, *4*(1).

Ekiert, M. (2007). The acquisition of grammatical marking of indefiniteness with the indefinite article a in L2 English. In *Teachers College, Colombia University Working Papers in TESOL & Applied Linguistics* (Vol. 7, Issue 1, pp. 1–43).

Fen-Chuan Lu, C. (2001). *The acquisition of English articles by Chinese learners*.

Freeman, D. E. L. (1975). The Acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly*, *9*(4), 409. https://doi.org/10.2307/3585625

Hakuta, K. (1976). Becoming bilingual: A case study of a Japanese child learning English. *Language Learning*, *26*(2), 321–351.

Hatch, E. (1978). Acquisition of syntax in a second language. In *Understanding second and foreign language learning* (pp. 34–70).

Huebner, T. (1983a). *A longitudinal analysis of the acquisition of English*. Karoma Pub.

Huebner, T. (1983b). Linguistic systems and linguistic change in an interlanguage. *Studies in Second Language Acquisition*, *6*(1), 33–53. https://doi.org/10.1017/S0272263100000280

Huebner, T. (1985). System and variability in interlanguage syntax. *Language Learning*, *35*(2), 141–163. https://doi.org/10.1111/j.1467-1770.1985.tb01022.x

Kirk, R. (1982). *Experimental design: Procedures for rhe behavioral sciences . Belmont. CA: Brooks/Cole*. StatChoice.

Liu, D., & Gleason, J. L. (2002). Acquisition of the article theby nonnative speakers of english: An Analysis of Four Nongeneric Uses. *Studies in Second Language Acquisition*, *24*(1), 1–26.

Lyons, C. (1999). *Definiteness*. Cambridge University Press.

Martinović, A. K., & Balenović, K. M. (1970). A study of English article usage among Croatian university students. *Методички Видици*, *10*(10), 95. https://doi.org/10.19090/mv.2019.10.95-112

Master, P. (1997). The English article system: Acquisition, function, and pedagogy. *System*, *25*(2), 215–232. https://doi.org/10.1016/S0346-251X(97)00010-9

Master, P. (2012). The effect of systematic instruction on learning the English article system. In *Perspectives on Pedagogical Grammar* (pp. 229–252). https://doi.org/10.1017/cbo9781139524605.011

Master, P. A. (1987). *A cross-linguistic interlanguage analysis of the acquisition of the English article system*. University of California.

Muftah, M. (2016a). English copular be in the interlanguage of adult l1 Arabic Speakers. *British Journal of Applied Science &Technology*, *16*(6), 1–13.

Muftah, M. (2016b). English past tense morphology in adult Arab EFL learners: Mental representation mechanism and types of errors. *JOURNAL OF ADVANCES IN LINGUISTICS*, *6*(3), 1038–1046. https://doi.org/10.24297/jal.v6i3.4666

Muftah, M. (2023a). Communication apprehension and self-perceived communication

competence: a study of undergraduate students in their final year. *Higher Education, Skills and Work-Based Learning*. https://doi.org/10.1108/HESWBL-08-2022-0174

Muftah, M. (2023b). Data-driven learning (DDL) activities: do they truly promote EFL students' writing skills development? *Education and Information Technologies*, 1–27. https://doi.org/10.1007/s10639-023-11620-z

Muftah, M., & Rafik-Galea, S. (2013). Error analysis of present simple tense in the interlanguage of adult Arab English language learners. *Canadian Center of Science and Education*, *6*(2), 146–154.

Murcia-Celce, M., & Freemen-Larsen, D. (1999). *The grammar book: An ESL/EFL teacher's course*. Newbury House.

Park, T.-S. (2006). The effect of natural exposure on advanced Korean EFL learners' English Article Use. 영어학, *6*(2), 279–305.

Parrish, B. (1987). A new look at methodologies in the study of article acquisition for learners of ESL. *Language Learning*, *37*(3), 361–384. https://doi.org/10.1111/j.1467-1770.1987.tb00576.x

Pica, T. (1983). Methods of Morpheme Quantification: Their Effect on the Interpretation of Second Language Data. *Studies in Second Language Acquisition*, *6*(1), 69–78. https://doi.org/10.1017/S0272263100000309

Pienemann, M. (1998). *Language processing and second language development: Processability theory* (Vol. 15). John Benjamins Publishing.

Sun, G. (2016). The Acquisition of English articles by second language learners. *SAGE Open*, *6*(1), 215824401663571. https://doi.org/10.1177/2158244016635716

Suniphan, T. (2022). An Analysis of English article usage by ESL students: A case study of Malaysian students. *Journal of Research and Development Institute Rajabhat Maha Sarakham University*, *9*(1), 121–138.

Takahashi, T. (1997). Japanese learners' acquisition and use of the English article system. *Edinburgh Working Papers in Applied Linguistics*, *8*, 98–110.

Tarone, E. E. (1985). Variability in interlanguage use: A study of style-shifting in morphology and syntax. *Language Learning*, *35*(3), 373–403.

Tarone, E., & Parrish, B. (1988). Task-related variation in interlanguage: The case of articles. *Language Learning*, *38*(1), 21–44. https://doi.org/10.1111/j.1467-1770.1988.tb00400.x

Testa, M. (2019). The acquisition of Spanish L3 articles: What can be learned from a simple linear regression analysis? *Research in Corpus Linguistics*, *7*, 84–112.

Thomas, M. (1989). The acquisition of English articles by first-and second-language learners.

Applied Psycholinguistics, *10*(3), 335–355.

Upor, R. A. (2021). Rater agreement and disagreement in the measurement of English article acquisition suppliance and accuracy. *LLT Journal: A Journal on Language and Language Teaching*, *24*(1), 126–148. https://doi.org/10.24071/llt.v24i1.2603

Wong, B. E., & Quek, S. T. (2007). Acquisition of the English definite article by Chinese and Malay ESL Learners. *Electronic Journal of Foreign Language Teaching*, *4*(2), 210–234. http://e-flt.nus.edu.sg/

Yamada, J., & Matsuura, N. (1982). The use of the English article among Japanese students. *RELC Journal*, *13*(1), 50–63. https://doi.org/10.1177/003368828201300104

Yoon, K. K. (1993). Challenging prototype descriptions: Perception of noun countability and indefinite vs. zero article use. *IRAL - International Review of Applied Linguistics in Language Teaching*, *31*(4), 269–290. https://doi.org/10.1515/iral.1993.31.4.269

Young, R. (1996). Form-function relations in articles in English Interlanguage. In *Second language acquisition and linguistic variation* (Vol. 135, p. 135). https://doi.org/10.1075/sibil.10.07you

Zhao, H., & Shirai, Y. (2022). Frequency effects in Chinese learners' acquisition of the English article construction. In *Challenges Encountered by Chinese ESL Learners* (pp. 237–263). Springer. https://doi.org/10.1007/978-981-16-5332-2_10