



Content list available at <http://ijltr.urmia.ac.ir>

*Iranian Journal  
of  
Language Teaching Research*

ORIGINAL ARTICLE



Urmia University

## The Age of Exposure to English Medium Instruction, Atypical Contexts, and Vocabulary Knowledge in a Second Language

Muhammad Asif Qureshi<sup>a,\*</sup>

<sup>a</sup> *Zayed University, UAE*

### ABSTRACT

The current study investigated differences in lexical knowledge of Arabic learners whose age of onset (AO) of exposure to English medium instruction (EMI) was at the elementary, secondary, and tertiary educational levels. Ninety undergraduate students enrolled in a public university in the UAE took part in the study. Data collection involved a background questionnaire, a vocabulary size task, and a vocabulary depth task. Using the background questionnaire, the participants were separated into the early, middle, and late learners - those exposed to EMI in elementary, secondary, and tertiary levels, respectively. The results revealed a significant multivariate main effect for the AO to vocabulary knowledge. The post hoc analysis confirmed a significant effect for vocabulary size only; no such effects were observed for vocabulary depth. Theoretical, methodological, and pedagogical implications are discussed.

**Keywords:** onset age of exposure; English medium instruction; vocabulary size; vocabulary depth; vocabulary knowledge

© Urmia University Press

### ARTICLE HISTORY

**Received:** 5 Apr. 2021

**Revised version received:** 3 July 2021

**Accepted:** 15 Dec. 2021

**Available online:** 1 Jan. 2022

\* Corresponding author: Department of Language Studies, Zayed University, UAE

Email address: [asif.qureshi@zu.ac.ae](mailto:asif.qureshi@zu.ac.ae)

© Urmia University Press

10.30466/ijltr.2022.121119

The importance of the English language cannot be over-emphasized in an increasingly globalized world. With the need to learn English, comes the question of when learning should start. Although several studies have examined the effects of AO on second language learning, results in this area are yet inconclusive. Findings in the second language (SL) contexts — where an L2 like L1 is used as an active language outside the classroom — support age effects for language learning wherein early starters almost always perform better than late starters on measures of second language proficiency (e.g., DeKeyser et al., 2010; Granena, 2012). However, in the foreign language (FL) environments, where exposure to a second language is mostly through instruction and limited to a few hours in school-settings, findings contradict the generally held belief of ‘earlier is better’ (Muñoz, 2011, Qureshi, 2018), indicating that age effects may play out differently in SL and the typical FL contexts. Atypical contexts – L2 settings where an L2 is used as a medium of instruction and widely available as a lingua franca in the society (e.g., United Arab Emirates) - are under-explored.

Moreover, most of the previous research that explored the impact of AO on L2 acquisition (L2A) focused either on morphosyntax (e.g., DeKeyser et al., 2010; Hartshorne et al., 2018) or phonology (e.g., Abrahamsson; 2012; Abrahamsson & Hyltenstam, 2009), while lexis as a variable of interest generally remained less explored. This lack of research on the AO to EMI and L2 lexical attainment needs to be addressed, especially because previous research shows significant contributions of vocabulary knowledge to various language abilities, including reading comprehension (Schmitt et al., 2011), scores on major aptitude tests (e.g., GRE, GMAT; Gardner & Davies, 2014), and overall academic attainment (Goldenberg, 2008). In terms of the four skills, command of vocabulary appears to be the major determinant of proficiency in writing, reading, listening, and speaking skills (Miralpeix & Muñoz, 2018). It also positively contributes to grammar knowledge, class grades, and inferencing skills (Schmitt, 2010). Notwithstanding the valuable contributions of lexical knowledge to academic life, its command remains a challenge for L2 learners (Shaw, 1991). Considering the significant contribution vocabulary makes to various aspects of academic life, it is vital to explore how exposure to EMI at different levels might influence ultimate attainment in this vital aspect of the second language. The present study addresses this gap by investigating the effects of AO on second language vocabulary knowledge in the UAE – an FL setting where English is offered as a medium of instruction and is commonly spoken in society. The following section first describes language context in the UAE, and later, it presents an overview of research exploring AO effects on L2 vocabulary knowledge in SL and FL settings.

### **Context: United Arab Emirates (UAE)**

The UAE has both Arabic and English medium strands in mainstream education. The onset age of exposure to English starts as early as the first grade for those attending EMI. However, students in Arabic medium schools predominantly receive their education in Arabic, while studying English as a Foreign Language (EFL) for only a few hours per week until they enter college when the medium of instruction changes to English for all the courses. The EFL experience of these students has been reported as predominantly teacher centered (Fareh, 2010), and deficient in several aspects, such as “limited instruction time, large class sizes, . . . and lack of effective and efficient assessment instruments” (Hussain et al., 2013, pp. 112-120). Recognizing this inadequate exposure to EFL, students are reported to experience hardship in learning subject matter in English (Belhiah & Elhami, 2015).

The medium of instruction situation, on the face of it, appears very similar to other FL contexts, such as Pakistan, where both L1 (i.e., Urdu-medium) and EMI schooling options are available. However, unlike most FL contexts, the UAE offers more opportunities for learners to immerse themselves in the target language use outside the classroom. In contrast to the typical FL settings, the UAE presents a distinctive picture where 90% of the population is expatriate (Lewis et al.,

2016) who use English for everyday communication. Acknowledging this, *The Ethnologue: Languages of the World* reports English as the “De facto national working language” in the UAE (Lewis et al., 2016). As such, its use can be noticed in offices, airports, parks, sports grounds, shopping malls, etc. This immense presence of English in the UAE offers opportunities for exploring the effects of EMI at different academic levels. Previous research supports positive contributions of immersive L2 experience, both formal and informal, to second language development (Qureshi, 2016; Lee, 2019; Moyer, 2009; Muñoz & Singleton, 2007; Pfenninger, 2014; Pfenninger & Singleton, 2016). Considering the contextual reality of the UAE – FL learners have greater exposure to English – it is important to investigate the extent to which AO of exposure to English medium instruction (AoEMI) in this particular context would account for differences in learners’ vocabulary knowledge in English.

## Literature Review

### *The Age of Onset and L2 Vocabulary Knowledge*

The effects of AO – the age of first exposure - on L2 vocabulary knowledge are generally inconclusive. The studies that have explored this impact on L2 vocabulary knowledge can be categorized as coming from either a target language-rich (SL) or a target language deficient (FL) background based on the amount of target language use present in the society. The following section presents an overview of these studies in SL settings first, followed by their discussion in the FL context. Within the FL settings, the studies will be divided between those that explored the impact of exposure to L2 as a foreign language, and those that investigated its influence as a medium of instruction, the latter being generally less common.

### *The Age of Onset and L2 Vocabulary Knowledge in SL contexts*

In the SL environments, learners, particularly those who immigrate to an English-speaking country as children, receive an immense amount of exposure to the target language, both in school and in society. In these contexts, the use of the target L2 is ubiquitous, and learners are left with no option except to use the target language for their most everyday interactions. For those who attend schools, their exposure to the target language - English in most cases - is augmented by receiving all instructions in English, which is further enhanced by the language support programs schools provide to non-English speaking students. In contrast, adult learners in these settings generally do not have the luxury of attending schools or receiving similar instructed exposure to the target language mainly due to their age and other obligations that accompany adulthood. In this context, previous research supports early learners’ superiority over late learners on various aspects of L2 vocabulary, including multiword units (Spadaro, 1998), lexical accuracy (Hyltenstam, 1998), and reaction time (Kim, 1997). However, several observations have been made about the quality of these studies (for review, see Hellman, 2008). These studies suffered from various methodological flaws such as missing information about participants’ AO in Hyltenstam (1998), use of peculiar lexical items or phrases for assessing lexical knowledge, like ‘jinks’, ‘gab’, and ‘throw his size’ (size = weight) in Spadaro (1998), or use of very basic pairs for assessing lexical size (man-women, black-white) in Kim (1997). It is obviously hard to make valid claims about the impact of age of exposure on lexical knowledge based on studies that have missing information about participants’ AO or those using very basic or peculiar items for assessing L2 lexical attainment. A later study by Hellman (2008) adapted more reliable measures to assess adult L2 learners’ command of English vocabulary as compared to native speakers. Her instruments included a Peabody Picture Vocabulary Task (4th edition; Dunn & Dunn, 2007), a Self-rated Vocabulary Size Task (Zechmeister et al., 1993), and a Word Association Task (Read, 1998). All the measures had been previously validated. The results showed that the majority of

adult-onset learners (76%) in this study performed at the native level (p. 123). Hellman used this as a piece of evidence to reject age as a constraining element for L2A of vocabulary. However, like Hyltenstam (1998), Hellman also reports no data about the participants' AO. Moreover, her study included only adult-onset learners; therefore, any comparison to childhood learners might be hard to make. To summarize the impact of AO on L2 vocabulary knowledge in the SL settings, it would be safe to conclude that findings thus far have been inconclusive, indicating a need for further exploration.

### *The Age of Onset and L2 Vocabulary Knowledge in FL contexts*

Similar to the findings in the SL context, studies conducted in the foreign language settings also do not offer a clear picture of AO effects on L2 vocabulary knowledge. Research in this setting can be divided into two categories: (a) that conducted in the instructed foreign language (IFL) setting and (b) those exploring the impact of extended exposure to the target L2 (i.e., EMI or extra courses besides FL classes) on L2 vocabulary attainment. Research in the first group represents a typical foreign language scenario where learners are exposed to an FL only for a few hours every day, while their medium of instruction for academic purposes remains their first language. The L2 learners in this setting have limited exposure to the IFL, for example, between 726 hours to 800 hours in three years (Miralpeix, 2007). In most cases, both early and late learners fall below adulthood, for instance, early and late learners' AO is 8 and 11 years in Cenoz (2000) and Miralpeix (2007), and in Xue et al. (2019), the mean age for six groups is ten years. In most cases, these learners can be referred to as early learners. Not only do studies in this setting fail to show any advantage for an early start, but also the overall gains in learners' lexical attainment are modest. For example, in Miralpeix (2007), after three years of exposure to IFL, 90% of the words all the groups used fell under the 1K list, while less than 1% of words used were from the 3K list. Nonetheless, the findings of the studies in this setting do not support early advantage either for receptive (Miralpeix, 2007; Pfenninger & Singleton, 2016) or productive vocabulary knowledge (Pfenninger & Singleton, 2016).

The only exception to the 'no early advantage' outcome in this setting is Xue et al. (2019), who explored the AO effects on 5th grade Chinese L1 speakers' L2 proficiency in English on a Peabody Picture Vocabulary Task (PPVT), a word identification task, and a lexical access task. The results of the multivariate analysis revealed the main effects of AO on the Peabody Picture Vocabulary Task only, while no significant effects were observed for the two tasks. However, it is important to note that, despite a main effect for the PPVT, participants' overall performance on the task was quite lackluster; they obtained a mean score of 49.45 out of 220 possible points. Moreover, several observations can be made about this study. For example, Xue et al. (2019) investigated the impact of AO on L1 (Chinese) and L2 acquisition in general, without specifically exploring ultimate attainment. Besides, the tasks used in data collection were basic, for example, for the lexical access, the participants were supposed to name five pictures (i.e., a tree, a car, ...). It can be argued that tasks that measure the size and depth of lexical knowledge might better account for the impact of AO on L2 learning. Moreover, findings by Xue et al. (2019) do not conform to the previous findings on the PPVT by Pfenninger and Singleton (2016) and Unsworth (2016), who found no significant main effect for the AO for early and late learners on this task. Unsworth (2016) observed a main effect for the amount of exposure as a covariate, indicating that the early starting group with greater exposure to the target language outperformed the late exposure group. In contrast to the studies that find no edge for early exposure (Miralpeix, 2007; Pfenninger & Singleton, 2016; Unsworth, 2016) or offer a mixed picture (Xue et al., 2019), Cenoz (2002) rather supports a late advantage. In Cenoz (2002), late learners (AO 10-11 years) outperformed early learners (AO 7-8 years) on the oral and written vocabulary tasks, a finding that supports AO effects for L2 vocabulary knowledge in favor of late learners. An advantage for the later AO is further supported by Agustín-Llach and Jiménez-Catalán (2018). In their study, while the amount of exposure to IFL was kept constant, adult learners outperformed young

learners attending 5th grade in the number of types and tokens produced in a written task. To summarize, previous research in this context, at best, offers a mixed picture, indicating three distinct outcomes: (a) those that report no advantage for early exposure to a foreign language, (b) those that report mixed findings, and (c) those that support a late-start advantage for vocabulary learning.

Similar to the IFL environment, findings in the extended contexts - settings where learners are exposed to an L2 either as an EMI or receive extra support in addition to their typical FL instruction - are inconclusive. Vidal and Jarvis (2018), who explored the impact of EMI on 195 undergraduate students' lexical diversity, found no significant improvements in three years of EMI. In contrast, Swain (1981) supports a late advantage for three sets of studies exploring L2 achievement in French. The three contexts these studies were conducted in included: (a) FL contexts where French was taught in non-French speaking countries (b) immersion context in Ontario where English speaking children were taught entirely in French for the early elementary schooling, followed by bilingual instruction in late-elementary years, and (c) extended French programs in Ottawa that allowed English speaking students to take a few courses in French along with having a daily class of French as an FL. The results revealed immersion programs as the most beneficial for L2 development, followed by the extended programs, which in turn were better than the FL programs. The study also confirmed the amount of exposure as the main contributor in learner's L2 proficiency and particularly acknowledged late starters as more efficient learners of a second language. The late starters in this review appeared to have a greater vocabulary size as compared to the early learners when the amount of exposure (i.e., time) was kept constant, a finding supported by the other studies (e.g., Cenoz, 2002; Cummins & Swain, 1986).

To sum up, research on the AO and L2 vocabulary knowledge is inconclusive and underexplored. Some of the potential reasons why the results in SL and FL settings have so far been inconclusive might include those factors that could influence the results of individual studies. These might involve measuring productive vs. receptive vocabulary knowledge (Pfenninger & Singleton, 2016); using different measures of lexical knowledge (i.e., density, variation, & sophistication; Hyltenstam, 1998); task effects (e.g., word recognition vs. understanding of meaning); and interaction between AO and amount of exposure (e.g., Swain, 1981). To better account for the AO effects, the type of data collection instruments should not only differentiate between the early and late starters but should also tap into their proficiency levels. Moreover, there is a general lack of research concerning the effects of AO on L2 lexical knowledge. This lack of research is despite the fact that vocabulary knowledge is central in language acquisition (Clark, 1993). It has a high correlation with other language skills, such as reading, writing, and grammar (Schmitt, 2010). In foreign language contexts where students are exposed to English medium courses at different levels of education (i.e., elementary, secondary, tertiary), the need for vocabulary knowledge becomes more acute. In these settings, communication takes place entirely, in the case of reading and writing, or mainly, for lectures and class conversations, through the medium of English. An investigation of the effects of the AO on two vital aspects of vocabulary knowledge - breadth and depth - might have policy implications for the appropriate age of exposure to EMI in instructed contexts. To address the lack of previous research exploring the impact of EMI on L2 vocabulary knowledge, the current study specifically investigated the following research question:

To what extent does the difference in the onset age of exposure to English medium instruction in an atypical context impact learners' vocabulary size and depth?

## Method

### *Participants*

To ascertain the number of participants needed for the study, an a priori power analysis was computed using GPower (Erdfelder et al., 1996). A priori power analysis requires three things: (a) a power level, usually .80, a reference effect size, typically based on the previous research, and a significance level of .05. Conducting a priori power analysis helps in determining the number of participants needed to correctly reject a null – no difference between the groups. Studies with smaller sample sizes are underpowered; hence, these might give an impression of false negative, leading to inaccurate interpretations of the results of a study (Farrokhyar et al., 2013). To ensure accuracy of results, a priori power analysis was conducted before the actual data collection in the current study. Keeping in view the possibility of emergence of different number of groups based on their exposure to EMI, two types of a priori analyses were conducted: one for two groups, and another for three groups. The results revealed that a sample size of 64, with 32 participants in each group should suffice if two groups emerge (i.e., elementary & tertiary), and a total of 42 participants would be needed if three groups (i.e., elementary, secondary, & tertiary) appear. A sample of 100 participants was sought to fulfil requirements for both of the two possible scenarios. The participants in the study were randomly selected from the undergraduate program, and later on, based on the data in the background questionnaire, they were divided into early, middle, and late groups.

### *Instruments*

#### *Background Questionnaire (BQ)*

To collect participants' background information, a BQ was used. The BQ contained 18 questions about participants' educational history, the academic level of exposure to EMI, age at testing, and their IELTS score at the entrance to the undergraduate program. The data obtained from the BQ are reported in Table 1.

Table 1  
*Contextual Information for the Early, Middle, and Late English Learners*

Characteristics	n	Age at Testing <i>M (SD)</i>	Male	Female	LangCent	<i>IELTS</i>
Early Learners [Grade 1]	64	20.00 (1.55)	23	39	3	5.97
Middle learners [grade 6]	12	20.60 (1.65)	3	9	3	5.57
Late learners [Undergrad]	16	20.45 (1.03)	7	9	2	5.42
Educational level						
1 <sup>st</sup> Year	26					
2 <sup>nd</sup> Year	50					
3 <sup>rd</sup> Year	14					

Based on the AO of exposure to EMI, three groups emerged for analysis. The early learners received instruction in English for all courses, except Islamic and Emirati studies, since they joined the elementary school. In contrast, the middle and late learners were exposed to EMI at the secondary and tertiary levels. Before this, the early and later learners had attended English as a subject (i.e., prose, poetry, etc.) for about 45 minutes every day in the elementary and secondary schools, respectively. The English courses were predominantly taught by non-native English

speakers. After joining college, all the middle and late learners attended one year of required intensive English program. At the time of the data collection - the end of the semester - all the late learners, except two, were finishing their third year of college. Hence, the late learners' length of significant exposure to English was four years – a year in the foundation program and three years of exposure in college. All the early learners were attending the first year of college at the time of data collection. Participants' proficiency in English, as depicted through their IELTS scores, was approximately the same for the early and middle groups but differed from the late learners by a mean difference of .5 only. Data about participants' socioeconomic status were not collected.

### *Vocabulary Size Task (VST)*

The vocabulary level test developed by Schmitt et al. (2001) was used for assessing learners' vocabulary size. This task measures learners' receptive vocabulary size. The test is based on an academic vocabulary list and frequency bands of word families at four frequency levels (i.e., 2000, 3000, 5000, & 10,000). The test has been validated by Schmitt et al. (2001) and includes ten sets for each frequency level. In each set, there are three stems and six options. Overall, the ten sets include thirty items. An example of the test instructions and format is provided in figure 1:

**Figure 1**

*Vocabulary Size Task (Schmitt et al., 2001, p. 58).*

**Directions:** This is a vocabulary size test. You must choose the right word to go with each meaning. Write the number of that word next to its meaning. An example is provided below.

**Example:**

1 business		
2 clock	<u>        6        </u>	part of a house
3 horse	<u>        3        </u>	animal with four legs
4 pencil / writing	<u>        4        </u>	something used for
5 shoe		
6 wall		

For the current study, the VST attained a KR-20 reliability coefficient of .93.

### *Vocabulary Depth Task (VDT)*

The Word Associate Test by Read (1993) was used for measuring learners' depth of vocabulary knowledge. On this task, the participants need to match a word with other words having similar meaning or that are used together. Overall, eight words are provided in two boxes, and the participants need to select four, without the restriction of any specific number from each box. The test has been revised (c.f. Read, 1998) and validated (c.f., Qian, 2002). An example of the test is provided below.

**Figure 2**

*Vocabulary Depth Task (Read 1998, p. 46).*

**Directions:** This is a vocabulary depth task. You are supposed to match a word with four options that have similar meaning or can be used together. There is no restriction on the number of choices you make from either side.

**Example****Sudden**

beautiful  quick  surprising  thirsty |  change  doctor  noise  school

For the current research, the VDT attained a KR-20 reliability coefficient of .95.

**Data Analysis**

On both the instruments, participants' responses were coded dichotomously: 1 point for each correct and 0 points for every incorrect selection. For the statistical analysis, International Business Machines: Statistical Package for Social Sciences (IBM-SPSS) version 24 was used. A summary of the scoring criteria for the instruments is provided in Table 2.

Table 2  
*Instruments, Number of Items, Scoring Criteria, and Total Possible Scores*

	Vocabulary Size Task	Vocabulary Depth Task
Number of items	5 lists 30 items/list	40-items 4 correct options for each
Scoring	✓ = 1 ☒ = 0	✓ = 1 ☒ = 0
Total possible score	150	160

***Independent and Dependent Variables, and Statistical Analysis***

The current study involved three levels of age (i.e., early, middle, & late) as the independent variables and two vocabulary tasks (i.e., VST & VDT) as the dependent variables. As the study involved multiple DVs, a MANOVA analysis was run, which compares groups with multiple dependent variables. For the post hoc investigations, when necessary, univariate analyses were run.

**Results**

The current study investigated the effects of AoEMI on vocabulary knowledge as represented through learners' performance on vocabulary size and depth tasks. Table 3 presents descriptive scores on the VST and VDT.

Table 3  
*Descriptive Scores for the Early, Middle, and Late Learners on Vocabulary Size and Depth*

Age of Exposure	<i>n</i>	<i>M</i>	<i>SD</i>
VST			
Early Learners	62	52.80	24.62
Middle Learners	12	38.66	23.85
Late Learners	16	30.62	25.03
VDT			
Early Learners	62	54.91	15.68
Middle Learners	12	46.71	14.68
Late Learners	16	50.12	14.19

\* VST had 150 items and VDT had 160 items. Scores for both were normed at 100.

According to Table 3, the early learners had the highest mean scores on the vocabulary size task, followed by the middle learners. The late learners obtained the lowest mean scores on this task. On the vocabulary depth task, a slightly different picture emerged; the early learners attained the highest means, followed by the late learners, while the middle learners secured the lowest mean scores.

To statistically evaluate the differences, a MANOVA was run. Assumptions of multivariate normality and variance were checked. Data on the dependent variable were bell-shaped, and skewness and kurtosis values for the VST .05 ( $SE = .25$ ), -1.00 ( $SE = .50$ ) and VDT -.42 ( $SE = .25$ ) and -.44 ( $SE = .50$ ) were less than the Z value of  $\pm 3.29$  ( $p < .001$ , two-tailed test) (Tabachnick & Fidell, 2007); hence, data were considered normally distributed. A Pearson's correlation of .15 ( $p = .14$ ) indicated the absence of multicollinearity and a non-significant *Bax'* M value of 6.87 ( $p = .37$ ) confirmed the assumption of homogeneity of variance. The results of the Mahalanobis distance showed the absence of multivariate outliers ( $p = .29$  to  $.94$ ), which were less than the standard  $p$  of .001. The results are provided in Table 4.

Table 4  
*Multivariate test for AoEMI and Vocabulary Knowledge*

	$\lambda$	<i>F</i>	<i>df</i> <sub>1</sub>	<i>df</i> <sub>2</sub>	<i>p</i>	$\eta^2$
AOE	.853	3.571	4	172.00	.008	.077

The analyses revealed a significant multivariate main effect of group (AoEMI) for vocabulary knowledge, *Wilks'*  $\lambda = .853$ ,  $F(4, 172.00) = 3.57$   $p < .008$ ,  $\eta^2 = .077$ . The power to detect the effect was .864. Thus, the results confirmed a significant effect of AoEMI on L2 learners' vocabulary knowledge, and showed a 7.7 % variance in learners' responses on the vocabulary tasks that could be accounted for by their AO of exposure to EMI. However, as the three groups had a difference in sample size, which could affect the results, a random sample of 16 participants was taken from the early learners and rerun for MANOVA with the pre-existing middle and late learner groups. The results still confirmed a significant multivariate main effect for the vocabulary knowledge, *Wilks'*  $\lambda = .627$ ,  $F(4, 80.000) = 5.26$   $p < .001$ ,  $\eta^2 = .135$ .

As the overall test was significant in both conditions – with larger and smaller early learner group, a univariate analysis on the larger data set was run. The univariate main effects were examined after checking the assumptions of independence, normality, and homogeneity. The resultant data

were bell-shaped, and the skewness and kurtosis values were .05 ( $SE = .25$ ), -1.01 for the VST, and -.42 ( $SE = .25$ ) and -.41 ( $SE = .50$ ) for the VDT; hence, data were considered normal. The Levene's test was not significant for VST' ( $p = .50$ ) and VDT' ( $p = .87$ ). The results of the analysis are presented in Table 5.

Table 5  
*Univariate F-tests Based on Learners' Age of Significant Exposure*

Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>P</i>	$\eta^2$
VST						
Between Groups	2	7213.861	3606.931	5.960	.004	.12
Within-groups	87	52652.94	605.196			
Total	89					
VDT						
Between Groups	2	833.493	416.746	1.761	.178	.01
Within-groups	87	20583.760	236.595			
Total	89					

To account for multiple ANOVA, a Bonferroni correction was made in the alpha by following the formula,  $\alpha = \alpha / c$  (number of comparisons of paired means)  $.05 / 3 = .017$ . The results showed a significant difference in participants' vocabulary size  $F(2,87) = 5.960, p = .004, \eta^2 = .12$ . Considering that the 10,000 frequency band could have affected learners' overall performance as it contained several words that seemed very uncommon (i.e., alabaster, swagger, squirm), the analyses were re-run excluding this band. The results still showed a significant difference  $F(2,87) = 6.126, p = .003, \eta^2 = .12$ . In contrast to vocabulary size, no significant difference was observed on vocabulary depth task,  $F(2,87) = 1.761, p = .178, \eta^2 = .01$ . As the ANOVA revealed a significant difference in the vocabulary size, a post hoc Tukey HSD analysis was computed to investigate the between-group differences. This analysis indicated that early and late learners differed significantly in their vocabulary size ( $p < .001, d = .89$ ), a finding that was also confirmed when a harmonic mean sample size ( $n = 18.52$ ) was used for comparing the groups.

To further investigate differences in learners' vocabulary size across five vocabulary levels (i.e., 2000, 3000, Academic Vocabulary, 5000, & 10,000), ANOVA and post hoc comparisons were computed. The descriptive scores for the three groups are displayed in table 6.

Table 6  
*Descriptive Scores for the Early, Middle, and Late Learners across five Vocabulary Levels*

	Early Learners ( $n = 62$ )			Middle Learners ( $n = 12$ )			Late Learners ( $n = 16$ )		
	%	<i>M</i>	<i>SD</i>	%	<i>M</i>	<i>SD</i>	%	<i>M</i>	<i>SD</i>
2000	79.51	23.85	6.53	68.61	20.58	8.36	50.20	15.06	10.25
3000	59.62	17.88	9.23	53.33	16.00	8.12	34.58	10.37	9.26
AcadVocb	57.09	17.12	9.35	43.05	12.91	7.85	31.17	9.31	9.30
5000	44.35	13.03	9.67	27.77	8.33	6.18	22.08	6.62	9.35
10,000	21.99	6.60	6.05	19.72	5.92	5.03	16.04	4.81	5.73

\* VST had 30 items in each band, so the total maximum score was 30 for each band.

The early learners achieved the highest mean scores across all the five vocabulary levels, followed by the middle, and late learners. All the three groups secured the highest scores on the lowest vocabulary band (i.e., 2000 word level), and as the bands progressed, participants' scores displayed a declining trend. Interestingly, results for the early group are surprising, considering, on average, they had only 80% correct responses at the 2000 word level after more than ten years of EML. This means they may not be able to recognize 20 out of 100 words on every page if they are reading a text at this basic level. Another notable observation was that all the groups obtained very humble scores at the 10,000 frequency band, indicating that the amount/length of learning even in the early group was not sufficient for them to obtain any significant number of words at that frequency band. It might be conjectured that the participants' scores on the 10,000 band may in fact be partly due to guessing on that test.

However, for a better understanding of variation across the five vocabulary levels, univariate analyses were computed. The assumption checks were performed and confirmed. These produced a bell-shaped distribution; the skewness and kurtosis values ranged from -.033 to 1.387, and -1.259 to 2.156, respectively; and the Levene's test was not significant (range = .54 to 3.25). The results of the ANOVA are presented in Table 7.

Table 7  
*Univariate F-tests for the Early, Middle, and Late Learners across five Vocabulary Levels*

Source	df	SS	MS	F	P	$\eta^2$
2000						
Between Groups	2	1005.57	502.78	8.838	.00	.16
Within-groups	87	4949.54	56.891			
Total	89					
3000						
Between Groups	2	718.94	359.47	4.336	.01	.09
Within-groups	87	7211.96	82.89			
Total	89					
Academic Vocabulary						
Between Groups	2	847.83	423.91	5.044	.00	.10
Within-groups	87	7311.32	84.03			
Total	89					
5000						
Between Groups	2	702.72	351.36	4.273	.01	.08
Within-groups	87	7153.59	82.22			
Total	89					
10,000						
Between Groups	2	41.51	20.75	.600	.55	.01
Within-groups	87	3008.27	34.57			
Total	89					

The ANOVA analysis showed a significant difference across four vocabulary levels (i.e., 2000, 3000, AcadVocab, & 5000); no significant difference was observed for the 10,000 word list. A post hoc Tukey HSD test revealed a significant difference only between the early and the late learners across all the four vocabulary levels (2000,  $p < .01$ ,  $d = 1.0$ ; 3000,  $p < .01$ ,  $d = .81$ ; AcadVocab,  $p < .01$ ,  $d = .83$ ; 5,000,  $p = .02$ ,  $d = .67$ ). No significant difference was observed in the comparison between early and middle learners and middle and late learners.

## Discussion

The current study investigated the effects of AO of exposure on early, middle, and late learners' lexical knowledge. The results revealed that learners who were exposed to EMI at the elementary level appeared to have better lexical proficiency as compared to those whose exposure started later (i.e., at secondary or tertiary levels). However, this difference was significant only for vocabulary size (i.e., how many words they know) but not for vocabulary depth (i.e., how well they know various aspects of a word). Overall, the AoEMI accounted for 7.7% to vocabulary knowledge for the omnibus outcome. This finding does not conform to the previous research that rejects AoEMI as a constraint for vocabulary knowledge in FL contexts (e.g., Miralpeix, 2007). The findings of the current study can be compared with Hellman (2008). Although Hellman (2008) is conducted in an SL context, both her study and the current investigation observed no differences in learners' vocabulary depth based on the AO of exposure to the target language. However, in contrast to Hellman (2008), which did not find any significant difference between early and late learners' vocabulary size, the current study reports a significant difference in favor of early learners across the frequency bands, except for the 10,000-word list, which did not reveal a significant outcome. The findings of the current study are also in contrast to Swain (1981) and Cummins and Swain (1986), who report a larger vocabulary size for late learners. In the current study, AoEMI accounted for 12% variance in favor of early learners, indicating a larger lexical size for them. For vocabulary depth, no significant difference was observed across the groups.

The different outcomes in the UAE as compared to the typical FL contexts might be attributed to the availability of informal exposure to the target language in society. Previous research supports the positive effects of informal exposure to a target language (Moyer, 2009; Pfenninger & Singleton, 2016). In the UAE, learners do have opportunities for language immersion in society, which is not typical of other FL settings. Perhaps it is because of this availability of English in the society that early learners exceeded the middle and the late learners in vocabulary size. This outcome does not conform to previous findings in other FL contexts (e.g., Álvarez, 2006; Cenoz, 2002; Muñoz, 2011; Qureshi, 2018), indicating a need to consider the variations across FL settings. In SL contexts, early learners outperform late learners mainly because of a rich language environment and their implicit learning ability (DeKeyser, 2003). In the UAE, English is frequently spoken in daily life, so it is reasonable to assume that although the three groups of participants had different AoEMI, they all had similar AO of exposure to English language in the society. Therefore, what could possibly make the three groups differ may be more about the differences in the total amount of exposure rather than the different AoEMI (Unsworth, 2016).

## Conclusion

The findings of the current study have theoretical, methodological, and pedagogical implications for research on the AO of exposure to EMI and lexical development in a second language. On theoretical grounds, this research makes an original contribution to our understanding of AO effects in the UAE context, which is dissimilar to other FL settings investigated in the past. In the UAE, roughly 90% of residents are expatriates, and English is used as the common functional language in society (Lewis et al., 2016). Hence, the findings of the current study might have implications for the theory of age effects and second language vocabulary development in similar settings. On methodological grounds, this study examined AO effects on L2 learners' lexical knowledge, a dimension of L2 proficiency generally underexplored in research on age effects. The exploration of AO effects on L2 lexis, using measures of vocabulary size and depth advances our understanding of this vital feature of second language learning. The results of these tasks facilitated the creation of a more in-depth profile of L2 learners' lexical proficiency. On the

pedagogical level, the study indicates weaknesses in both EMI and the limited EFL instruction provided to students in this context. The study revealed that despite varying lengths of exposure, the early, middle, and late learners scored 79%, 69%, and 50% at the 2000 levels, indicating that the learners did not know about 20% to 50% words at this basic level. The results for the early group were particularly surprising, considering, on average, they had only 80% correct responses at the 2000 word level after more than ten years of English medium instruction. The range of known words showed a decreasing trend as the frequency lists ascended. At the academic vocabulary level, which came after the 3,000 frequency band, learners knew only 57% (early learners), 43% (middle learners), and 31% (late learners) words. This limited proficiency at the level of academic word list might have serious implications for learners' reading comprehension, among other areas. Previous research suggests that, with instructional support, L2 learners need to know 98% of words in an academic text to read it successfully (Schmitt et al., 2011). A strong correlation between vocabulary knowledge and reading comprehension is well established. For example, Jeon and Yamashita (2014) explored the relationship between ten different variables and reading comprehension, adopting a meta-analytic approach. Their findings unveiled vocabulary knowledge as the strongest correlate ( $r = .79$ ) of second language reading comprehension. Recognizing the significance of vocabulary knowledge for second language reading comprehension, and considering the findings of the current study, it might be safe to assume that the lack of vocabulary knowledge might be one of the reasons L2 learners in the Gulf, including UAE, score lowest in reading tests (c.f., IELTS as cited in Gallagher, 2016). Moreover, as learners' overall proficiency in vocabulary was quite lackluster – averaging around 50%, including those who were exposed to EMI at the elementary level – the quality of EMI imparted might need serious reconsiderations.

Nonetheless, the current study did not make a concerted effort to elicit effects of the availability of English in the society on L2 learners' language proficiency. Hence, any generalization should be made with caution; instead, these effects should be investigated through empirical investigations. Moreover, in the current study, late learners had a shorter exposure to English medium instruction, as they were exposed to it for four years – three years in college and one year in the foundation program. Previous research conducted in SL settings suggests a length of exposure of five years, beyond which the duration of exposure does not seem to significantly affect the outcome on second language grammar knowledge (DeKeyser et al., 2010). No similar threshold for the duration exposure is suggested for the FL settings. In the current study, the late learners had a short length of immersive exposure to English (i.e., 4 years). However, they had attended an English class of 45 minutes every day during their elementary and secondary schooling, which might offset the impact of lack of length of exposure. Nonetheless, future research with a longer length of exposure to EMI for the late learners might help in better understanding the effects of AO on L2 learners' lexical proficiency.

## References

- Abrahamsson, N. (2012). Age of onset and nativelike L2 ultimate attainment of morphosyntactic and phonetic intuition. *Studies in Second Language Acquisition*, 34, 187-214.
- Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning*, 59, 249-306.
- Agustín-Llach, M. P., & Jiménez-Catalán, R. M. (2018). Teasing out the role of age and exposure in EFL learners' lexical profiles: A comparison of children and adults. *International Review of Applied Linguistics in Language Teaching*, 56(1), 25-43. <https://doi.org/10.1515/iral-20150075>

- Álvarez, E. (2006). Rate and route of acquisition in EFL narrative development at different ages. In C. Muñoz (Ed.), *Age and the rate of foreign language learning* (pp. 127–155). Multilingual Matters.
- Belhiah, H., & Elhami, M. (2015). English as a medium of instruction in the Gulf: When students and teachers speak. *Lang Policy*, 14, 3–23.
- Cenoz, J. (2002). Age differences in foreign language learning. *Review of Applied Linguistics*, 13, 125–142.
- Clark, E. (1993). *The lexicon in acquisition*. Cambridge University Press.
- Cummins, J., & Swain, M. (1986). Linguistic Interdependence: A Central Principle of Bilingual Education. In J. Cummins & M. Swain (Eds.), *Bilingualism in education: Aspects of theory, research and practice* (pp.80-95). Longman.
- DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22, 499-533.
- DeKeyser, R. (2003). Implicit and explicit learning. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 313-348). Blackwell.
- DeKeyser, R., Alfi-Shabtay, I., & Ravid, D. (2010). Cross-linguistic evidence for the nature of effects in second language acquisition. *Applied Psycholinguistics*, 31, 413-438.
- Dunn, L.M., & Dunn, D.M. (2007). *The peabody picture vocabulary test* (4<sup>th</sup> ed.). NCS Pearson, Inc.
- Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments & Computers*, 28(1), 1–11.
- Fareh, S. (2010). Challenges of teaching English in the Arab world: why can't EFL programs deliver as expected? *Procedia*, 2, 3600-3604.
- Farrokhyar, F., Reddy, D., Poolman, R. W., & Bhandari, M. (2013). Why perform a priori sample size calculation? *Canadian Journal of Surgery. Journal canadien de chirurgie*, 56(3), 207–213. <https://doi.org/10.1503/cjs.018012>.
- Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics*, 35 (3), 305-327.
- Goldenberg, C. (2008). Teaching English language learners: What the research does—and does not—say. *American Educator, Summer*, 8–44.
- Gallagher, K. (2016). From 'late-late' to 'early-early' immersion: Discontinuities and dilemmas in medium of instruction policies and practices. In L. Buckingham (ed). *The place of English in Gulf Arab societies*. Multilingual Matters.
- Granena, G. (2012). *Age differences and cognitive aptitudes for implicit and explicit learning in ultimate second language attainment* (Doctoral dissertation, University of Maryland, College Park). <http://libproxy.nau.edu/docview>.

- Hartshorne, J. K., Tenenbaum, J. B., & Pinker, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition*, 177, 263–277. <https://doi.org/10.1016/j.cognition.2018.04.007>
- Hellman, B. A. (2008). Vocabulary size and depth of word knowledge in adult-onset second language acquisition. *International Journal of Applied Linguistics*, 21, 162-182.
- Hussain, M., Nasseef, O., & Shah, S. (2013). Factors impacting EFL teaching: an exploratory study in the Saudi Arabian context. *Arab World English Journal*, 4(3), 104-123.
- Hyltenstam, K. (1998). Lexical characteristics of near-native second-language learners of Swedish. *Journal of Multilingual and Multicultural Development* 9 (1-2), 67-84.
- IBM Corp. (2016). IBM SPSS Statistics for Windows, Version 24.0. IBM Corp.
- Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning*, 64, 160-212.
- Johnson, J. S., & Newport, E. (1989). Critical period effects in second language learning: The influence of the maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21, 60-99.
- Kim, E. J. (1997). The sensitive period for second-language acquisition: A reaction-time study of maturational effects on the acquisition of L2 lexico-semantic and syntactic systems. *Dissertation Abstracts International*, 57 (12), 5133A. (UMI No.9717292).
- Lee, E. (2019). Age of onset, type of exposure, and ultimate attainment of L2 morphosyntactic sensitivity. *Studies in Second Language Acquisition*, 1-23. <http://doi:10.1017/S0272263119000731>.
- Lewis, M., Gary, F., & Charles, D. (2016). *Ethnologue: Languages of the World*. <http://www.ethnologue.com>
- Miralpeix I. (2007). Lexical knowledge in instructed language learning: The effects of age and exposure. *International Journal of English Studies*, 7(2), 61-83.
- Miralpeix, I. & Muñoz, C. (2018). Receptive vocabulary size and its relationship to EFL language skills. *International Review of Applied Linguistics in Language Teaching*, 56(1), 1-24. <https://doi.org/10.1515/iral-2017-0016>
- Moyer, A. (2009). Input as a critical means to an end: Quantity and quality of experience in L2 phonological attainment. In M. Young-Scholten & T. Piske (Eds.), *Input matters in SLA* (pp. 159–174). Multilingual Matters.
- Muñoz, C. (2011). Input and long-term effects of starting age in foreign language learning. *International Review of Applied Linguistics in Language Teaching*, 71, 197-220.
- Muñoz, C., & Singleton, D. (2007). Foreign accent in advanced learners: two successful profiles. *The EUROL2A Yearbook*, 7, 171-190.

- Pfenninger, S. (2014). The misunderstood variable: Age effects as a function of type of instruction. *Studies in Second Language Learning and Teaching*. <https://doi.org/10.14746/ssl.2014.4.3.8>
- Pfenninger, S., & Singleton, D. (2016). Affect trumps age: A person-in-context relational view of age and motivation in SLA. *Second Language Research*. Prepublished 15 January 2016; doi: 10.1177/0267658315624476
- Qian, D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance. *Language Learning* 52(3), 513–536.
- Qureshi, M. A. (2016). A meta-analysis: Age and second language grammar acquisition. *System*, 60, 147–60.
- Qureshi, M. A. (2018). Age and knowledge of morphosyntax in English as an additional language: Grammatical Judgment and Error Correction. *International Review of Applied Linguistics*. DOI <https://doi.org/10.1515/iral-2015-0062>
- Read, J. (1998). Validating a test to measure depth of vocabulary knowledge. In Kunnan, A. (ed.), *Validation in language assessment* (pp. 41–60). Lawrence Erlbaum.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillan.
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *Modern Language Journal* 95, 26–43.
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the Vocabulary Levels Test. *Language Testing*, 18(1), 55–88.
- Shaw, P. (1991). Science research students' composing processes. *English for Specific Purposes*, 10, 189–206.
- Spadaro, K. (1998). Maturation constraints on lexical acquisition in a second language. *Unpublished Doctoral Dissertation*. University of Western Australia, Australia.
- Swain, M. (1981). Time and timing in bilingual education. *Language Learning*, 31, 1–15.
- Tabachnick, B. G. & Fidell, L. S. (2007). *Using multivariate statistics* (5<sup>th</sup> ed.). Allyn and Bacon.
- Unsworth, S. (2016). Early child L2 acquisition: Age or input effects? Neither, or both? *Child Language*, 43, 608–634.
- Vidal, K., & Jarvis, S. (2020). Effects of English-medium instruction on Spanish students' proficiency and lexical diversity in English. *Language Teaching Research*, 24(5), 568–587. <https://doi.org/10.1177/1362168818817945>
- Xue, J., Hu, X., Yan, R., Wang, H., Chen, Xi., & Li, M. (2019). Onset age of language acquisition effects in a foreign language context: Evidence from Chinese–English bilingual Children. *Journal of Psycholinguistic Research*. <https://doi.org/10.1007/s10936-019-09637-y>

Zechmeister, E. B., D'Anna, C. A., Hall, J.W., Paus, C. H., & Smith, J. A. (1993). Metacognitive and other knowledge about the mental lexicon: Do we know how many words we know? *Applied Linguistics*, 14(2),188–206.

### **Acknowledgment**

This research could not have been possible without the generous support I received from the Zayed University Through their Research Incentive Fund (RIF) for which I am deeply grateful.

**Muhammad Asif Qureshi** is an associate professor in the department of Language Studies at Zayed University. Dr. Qureshi's main research interest is second language acquisition with a particular focus on the effects of age of first exposure on second language grammar and vocabulary knowledge. Besides, he explores the validity of data collection tasks and the types of knowledge these elicit, as well as the impact of translanguaging and text-based vocabulary knowledge on reading comprehension.