


## Article

# Psycho-Social Constraints on Naturalistic Adult Second Language Acquisition

Azza Al-Kendi <sup>1</sup>  and Ghada Khattab <sup>2,\*</sup> 

<sup>1</sup> Department of English Language and Literature, Sultan Qaboos University, Al Seeb P.O. Box 42, Oman; azzakindi@squ.edu.om

<sup>2</sup> Speech and Language Sciences, School of Education, Communication and Language Sciences, Newcastle University, Newcastle upon Tyne NE1 7RU, UK

\* Correspondence: ghada.khattab@ncl.ac.uk

**Abstract:** The following study investigated a rare case of adult immersion in a second language context without prior exposure to the language. It aimed to investigate whether Length of Residence (LoR) acts as a strong index of L2 speech performance when coupled with daily exposure and interaction with first language speakers. Twenty-two females from Africa and Asia who worked as Foreign Domestic Helpers (FDH) in Omani homes and with varying LoRs performed an AX discrimination and a production task which tapped into Omani consonants and clusters that are absent from their L1s; their accent was also rated by L1 Omani listeners. Results showed a surprising lack of significance of LoR on all the production and perception measures examined. Discrimination results showed a low sensitivity to Arabic consonantal contrasts that are lacking in the L1 across all participants, and a small positive effect of L1 literacy. Production results exhibited low accuracy on all Arabic consonants and a marked foreign accent as judged by L1 listeners, with a small positive effect of L2 literacy. We argue that the nature of the interactions between FDH and employers, along with uneven power relations and social distance, counteract any advantage of LoR and the immersion setting examined here.

**Keywords:** length of residence; foreign domestic helper; foreign accent; naturalistic adult acquisition; L2 speech performance



**Citation:** Al-Kendi, Azza, and Ghada Khattab. 2021. Psycho-Social Constraints on Naturalistic Adult Second Language Acquisition. *Languages* 6: 129. <https://doi.org/10.3390/languages6030129>

Academic Editors: Robert Mayr and Jonathan Morris

Received: 7 June 2021  
Accepted: 22 July 2021  
Published: 28 July 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Over the past decades, a great deal of second language acquisition (SLA) research has focused on sociolinguistic factors that play a role in successful SLA. In naturalistic second language (L2) settings, age of learning (AoL) and length of residence (LoR) are among the most frequently studied predictors that have been found to affect second language speech learning (Piske et al. 2001). AoL has stood the test of time despite disagreements over whether or not there is a critical age or period for learning, but a fair number of studies have pointed out the unequal opportunities to learn the L2 between younger and older speakers (e.g., Klein and Perdue 1997; Craats et al. 2006). Similarly, LoR has proven robust when 'length' goes hand in hand with exposure. A host of factors have, however, been found to attenuate the effects of LoR, including experience with the L2 prior to arrival in the L2 country, the nature of the input, formal instruction, and first language (L1) literacy/education. Here we review some of this work before we turn our attention to an understudied population of L2 learners who fill a gap in terms of enabling us to test what happens when adult L2 learners receive extensive oral input from L1 speakers from the start and for prolonged periods of time.

SLA studies have long highlighted the difficulty adults face when learning L2 speech (Akahane-Yamada 1995; Best and Strange 1992; Flege 1981; Flege et al. 1995; Iverson and Kuhl 1996). The most widely discussed source for this difficulty is AoL. One of the most influential (and controversial) proposals in this area is that of the critical period hypothesis

(CPH), by [Lenneberg \(1967\)](#), who claimed that the ability to learn a new language declines after puberty due to the completion of neural hemispheric lateralization in the brain. A large body of work has both supported and challenged this claim, with disagreements over the existence of an age cut-off and evidence challenging the loss of neural plasticity in older age (e.g., [DeKeyser 2000](#); [Flege 1987, 2018](#); [Moyer 2014](#)). Nevertheless, and perhaps due to age acting as proxy for other optimal opportunities for L2 acquisition, studies involving naturalistic exposure to an L2 support at least a gradual decline in L2 learning outcomes as a function of age (e.g., [Flege 2018](#); [Flege et al. 1995](#); [Oyama 1976](#); [Pfenninger and Singleton 2017](#); [Seliger 1975](#); [Scovel 1988](#)). One linguistic by-product of age which is often attributed to the difficulty to learn the L2 sounds is increased L1 mastery and subsequent influence on the L2 (e.g., [Best et al. 1988](#); [Escudero and Boersma 2004](#); [Escudero 2006](#); [Flege 1995](#); [Kuhl 1991](#)). As L1 categories become more established with age, difficulties in perceiving and producing the L2 seem to arise ([Flege 1995](#)).

Length of residence (LoR) has also been frequently used in SLA research as an index for ultimate attainment in L2 phonology ([Flege 2009](#); [Moyer 2009](#)). [Flege \(2009\)](#) postulates that if the amount of input learners receive matters, then LoR should be correlated with measures of L2 speech attainment. Likewise, [McAllister \(2001\)](#) states that LoR correlates positively with the amount of input an L2 learner has acquired and that the more L2 input one receives the better the opportunities for the L2 learner to master the L2. A range of studies has provided evidence for the positive role of LoR in L2 performance. For instance, [Flege and Fletcher \(1992\)](#) revealed that Spanish adults who had lived in the United States for an average of 14.3 years received significantly better pronunciation ratings of English sentences than individuals with an average LoR of 0.7 years. Similarly, [Flege et al. \(1997\)](#) reported an effect of LoR on the perception of English vowels by L2 adult speakers of English who varied in their LoR between 0.7 and 7.3 years, albeit a modest one. A more significant effect was found on production accuracy, especially for one of the vowels they examined. In more recent work, [Højen \(2019\)](#) found a significant improvement in L1 Danish females' English pronunciation after 7.1 months of short-term immersion in England, with LoR significantly correlating with the participants' pronunciation gain score.

However, other reports on the significance of LoR effects reveal inconsistent results. For instance, [Oyama \(1976\)](#) and [Flege and Fletcher \(1992\)](#) found no effect of LoR on the L2 phonology of Italian and Spanish English speakers in the United States when the effect of their age of arrival (AoA) was controlled for. Similarly, [Flege \(1988\)](#) reported no difference in foreign accents between two groups of Taiwanese adult immigrants to the United States based on their LoR, which varied between 1.1 and 5.1 years. Further investigation revealed that LoR did not influence the degree of foreign accent after a rapid initial stage of L2 learning.

[Piske et al. \(2001\)](#) discuss a number of reasons for the discrepancies in previous studies. First, for studies that found LoR to have an influence on the degree of foreign accent, LoR was a less significant predictor compared to AoL. Second, LoR is more likely to affect the degree of foreign accent if the mean values of LoR of the L2 learner groups differ greatly. Third, additional years of stay in the L2 community are not likely to lead to a decrease in foreign-accented speech in already experienced L2 learners. However, L2 learners who are in the initial phases of learning the L2 when they arrive in the L2 country might benefit from additional years of experience ([Højen 2019](#)). This once again highlights the importance of input from L1 speakers from the start, as well as the importance of the cumulative effect of L2 exposure, as highlighted by [Flege and Bohn \(2021\)](#) in the revised Speech Learning Model (SLM-r); but most studies have investigated adult immigrants who had previously studied the L2 in their countries of origin, therefore being initially exposed to the L2 in a foreign language context. Little is known about the potential influence of total immersion in the L2 from first exposure for adults, bringing them closer to the experience of the children of immigrants.

While a naturalistic setting can be advantageous for the children of immigrants, who also receive formal instruction in the L2 country, adult immigrants may be disadvantaged due to the difficulty in getting access to formal instruction and/or due to coming from

low educational backgrounds (Klein and Perdue 1997; Craats et al. 2006). The research on oral/aural L2 performance of low educated learners is scarce. The vast majority of studies on SLA make use of convenience sampling and thus show an overreliance on a population which is WEIRD (Western, Educated, Industrialized, Rich, Democratic; Henrich et al. 2010). A significant proportion of the research that has been published in journals including Second Language Research, TESOL Quarterly, and Studies in Second Language Acquisition relies on such samples (Bigelow and Tarone 2004; Craats et al. 2006). Furthermore, few studies include L1 formal schooling as a contributing variable (Craats et al. 2006; Haznedar et al. 2018; Young-Scholten 2013). This bias towards recruiting and examining highly educated L2 learners might have skewed our understanding of second language speech and led to an under-explanation of how adults acquire a new language when we isolate factors such as first language literacy.

A body of work on non-literate and low-educated adult immigrant adults to the USA, Europe and Australia has focused on the challenges this poses for L2 literacy and metalinguistic awareness and calls for further research in this area (Craats et al. 2006; Kurvers et al. 2006; Young-Scholten and Strom 2006). However, difficulties do not necessarily arise in all language domains, with work within the area of morphosyntax suggesting that L2 learners can follow a common route in their L2 development of morphosyntax regardless of their age, educational background, L1 or input (e.g., Gass 2013; Ordem and Bada 2017). More work is still needed on the levels of attainment in all domains, including L2 oral production and perception. Little is known about whether L2 learners who are mainly exposed to extensive oral input from L1 speakers in a naturalistic setting achieve L2 speech outcomes that are more similar to those of L1 children.

To summarize, previous research on L2 speech learning has focused on age and LoR as main external factors in adults' ultimate attainment in the L2. However, the strength as well as vulnerability of these factors is due to their interaction with a multitude of other factors which can increase or reduce opportunities for learning. These include the amount of input from the L2 and opportunities for interaction with L1 speakers, the availability of formal instruction in the L2 and the level of L1 literacy prior to arrival in the L2 country, amongst others. The current study reports on a rare situation of naturalistic adult L2 acquisition through total immersion in the L2 context. It focusses on an understudied group of low-educated migrants from East Africa and South Asia who spend long years in the Arab world as domestic helpers, living with their Arab employer and their family and therefore mainly interacting with and receiving input from speakers who are first language users. The aim of the study is to investigate whether the situation provides an optimal context for L2 speech attainment given that conditions that typically strengthen LoR effects are maximized by the context; the learners have little or no previous exposure to foreign-accented Arabic, their LoR varies a great deal due to constant new arrivals, allowing for a comparison of short and long LoR, and their LoR highly correlates with input and interaction. The focus on Arabic as an L2 here is advantageous for two reasons: compared with English, Arabic is relatively understudied in SLA research (but see Ioup et al. 1994; Alhawari 2018); and by focusing on Arabic rather than English as an L2, there is a smaller chance that learners will have had exposure to it prior to arriving in the Arab world (apart from religious practices for some, which will be described later). On the other hand, social factors such as low L1 literacy and uneven power relations between employer and employee may attenuate LoR effects, but these have not been sufficiently considered in the speech-learning literature.

## 2. Materials and Methods

In what follows we present perception and production experiments that were carried out with foreign domestic helpers (FDH) who were living and working in Oman at the time of the study. The main aim was (1) to investigate the extent to which FDH had acquired Arabic consonants and clusters that were expected to pose a challenge in production and perception due to their articulatory and/or phonological complexity and their absence

from the L1s of the FDH; (2) the extent to which successful acquisition correlated with LoR and L1/L2 literacy. Production analyses were also supplemented with foreign-accent ratings that were carried out by L1 Arabic listeners. Here it is important to note that we are not espousing the view that FDH should sound like L1 speakers of Omani Arabic, or that this should be their aspiration. We concur with research that warns against the native speaker ideal (e.g., [Holliday 2018](#)), and in fact we avoid using the term 'native' wherever possible. The comparison with L1 speakers and the accent rating does, however, allow us to compare how closely L2 speakers approximate the accent patterns of L1 speakers when input is almost exclusively from those speakers. This allows us to address methodological constraints in other studies that have addressed this question, where learners had previously been exposed to accented varieties of the L2 prior to arriving in the L2 country, and/or their residence in the L2 country does not necessarily go hand in hand with increased input.

## 2.1. Participants and Languages under Examination

### 2.1.1. Participants

Twenty-two female FDH who worked for and lived with families in Oman participated in this study. Consent for participation in the study was obtained from the FDH and their employers. The participants completed a questionnaire which elicited information about their demographic and sociolinguistic background. This included information on age, L1(s), age of arrival (AoA) in the Arabic-speaking world, length of residence (LoR) in the Arabic-speaking world, years of formal schooling in the L1 and L2 Literacy (ability in reading or writing in Arabic). The first author read the questionnaire to the FDH and recorded their answers to these questions.

The participants' background was representative of the very diverse background of FHDs who work in the Arab world. For instance, the FDH came from nine L1 backgrounds (Swahili (5), Indonesian (2), Sinhala (4), Tagalog (5), Bengali (2), Telugu (1), Luganda (1), Yoruba (1) and Oromo (1)). They migrated to the Arabic-speaking world as adults (mean AoA = 27.27), and they had varying Arabic experiences based on their LoR that ranged from 0.7 to 21 years (mean LoR = 6.23). Their mean LoR in Oman was 2.36 years. Nine of them had worked in different Arabic-speaking countries before moving to Oman (e.g., Gulf countries and Lebanon). They all reported that they had been addressed to mainly in Arabic by the family members of the household(s) they had lived in and worked for. Fourteen of them had never been exposed to Arabic before migration, while eight had had access to Arabic via Islam and recitation of the Qur'an. It should be noted that when classifying FDH based on their Arabic literacy, only Muslim FDH who reported being able to read in Arabic via recitation of the Qur'an were considered as literate. Those who reported not to be able to read in Arabic or recite the Qur'an were considered as non-literate in Arabic even if they were exposed to Arabic during other rituals of worship. Other than Arabic, 15 of them reported having some knowledge of English. Ten L1 Omani speaker females were recruited in order to obtain comparative information from L1 patterns for the perception and production tasks. They all had a comparable educational background and were between 19 and 40 years old.

### 2.1.2. L1 Consonant Inventory and Target Sounds and Structures

Table 1 shows the consonant chart of Omani Arabic. In order to control for the variability in the FDH's L1, we targeted consonants that were absent in all the L1 sound inventories of the FDH participants and that were likely to pose a challenge in perception and/or production due to their complex articulation; these are highlighted in grey. Table 2 details whether the L1 phonology of the FDH participants allows onset and coda consonant clusters in comparison with the target language. It is worth noting that epenthesis in onset clusters is optional in Omani Arabic, but the examination of adult input suggests a prevalence of non-epenthetic realisations ([Al-Kendi 2021](#)). L1 phonologies of FDHs that have a restricted onset structure only permit a limited combination of CCs. For

example, Tagalog onset consonant clusters are restricted to a consonant and a glide (e.g., /bwan/ ‘month’), while Luganda and Swahili onset structures are restricted to a consonant and a glide or a consonant preceded by a nasal (e.g., /mba/).

**Table 1.** The sound inventory of Omani Arabic (The highlighted consonants are absent from the L1 inventories of the FDH).

	Voicing	Labial	Labio-Dental	Dental	Alveolar	Palatal	Velar	Uvular	Pharyngeal	Glottal
Stop	Voiced	b			d		g			
	Voiceless				t		k	q		ʔ
Emphatic stop	voiceless				t <sup>ʕ</sup>					
Fricative	Voiced			ð	z			ʁ	ʕ	
	Voiceless		f	θ	s	ʃ		χ	ħ	h
Emphatic fricative	Voiced			ð <sup>ʕ</sup>						
	Voiceless				s <sup>ʕ</sup>					
Nasal	Voiced	m			n					
Liquid	Voiced				r l					
Approximant	Voiced	w				j				

**Table 2.** The prevalence of CC clusters in the sound systems under investigation. Languages that permit consonant clusters are indicated with the symbol ✓, while those that do not are indicated with the symbol x.

Language	Initial CC	Final CC
Omani Arabic	✓	✓
Bengali	✓	x
Swahili	✓ (restricted)	x
Luganda	✓ (restricted)	x
Tagalog	✓ (restricted)	x
Yoruba	x	x
Oromo	x	x
Indonesian	x	x
Sinhala	x	x
Telugu	x	x

## 2.2. Examining FDH’s Discriminability of Arabic Consonants

### 2.2.1. AX Task and Stimuli

An AX forced-choice discrimination paradigm was used to elicit FDH’s discriminability of different Arabic consonantal contrasts. In this kind of task, participants are presented with two stimuli in a sequential order whereby the second stimulus is either the same as the first (AA) or different (AB) (Strange and Shafer 2008), and they answer ‘same’ or ‘different’. For the current study, a list of 16 Arabic consonant contrasts was created. The phonemic pairings were created based on their potential confusability for the listeners in terms of perception and/or production, as they only varied in one feature: voicing (e.g., /θ/-/ð/), manner (e.g., /t/-/s/), place (e.g., /χ/-/ħ/; /q/-/k/) or the presence or absence of secondary articulation (e.g., /t<sup>ʕ</sup> /-/t/). The latter refers to Arabic emphatic sounds whose

production is accompanied by a primary articulation at the dental/alveolar area and a secondary articulation that involves a constriction in the upper pharynx.

Two more contrasts were used as control: /r/-/l/, /r/-/w/. Given that the FDH's L1 inventories included these sounds, albeit with potentially different phonetic realizations and phonological patterning, it was likely that all participants would detect these contrasts as different (Aoyama et al. 2004). With the control pairs, a total of 18 consonant contrasts were included in the test. Four test items were created for each of the 18 contrasts (AA, AB, BA, BB), yielding a total of 72 test trials. Each test trial consisted of two monosyllabic pseudo words containing the contrasting sounds in the context Ca:n (where C is a consonant), for instance, /sa:n-<sup>s</sup>a:n/. Thirty-six trials contained consonants that were acoustically different, while 20 trials contained consonants that were acoustically identical (16 trials were excluded from the list, as they were repetitions of existing trials). This reduced the number of trials to 56. To give an example, /θ/ was paired twice, once with /s/ and another with /t/ because [t] and [s] are likely variants of /θ/ in NNS' productions (Lombardi 2003). When the four test trials were created for each of these pairs, one test item was repeated for both contrasts. So, one of the repeated trials was excluded. When all trials were created, they were submitted to an online randomization software (RANDOM.ORG) to ensure that the four test items for each contrasting pair were not following each other. The stimuli were recorded by the first author in a sound-treated lab using an Edirol digital recorder R-09HR by Roland coupled with a Sennheiser radio microphone, with a sampling rate of 44.100 Hz and WAV-16bit recording mode. Another native Omani speaker trained in linguistics listened to the recorded stimuli for a reliability check. She spoke the same dialect as the NSs in this study. She confirmed that all recorded instances were clearly articulated and checked that time intervals between test items were the same and as specified in the present study.

#### 2.2.2. Procedure for the AX Discrimination Task

In the home of their employers, each FDH was presented with the aural stimuli over headphones at a comfortable volume level using the Praat program on a MacBook laptop (Boersma and Weenink 2009). They were instructed, in Arabic, that they needed to decide if the two test items they were about to hear were the same or different. They gave their responses to the first author, who manually entered them on an answer sheet designed specifically for this task. It was not possible to use a full computer version for this task due to the potential difficulty the participants might have faced dealing with technology in case of limited computing education. An inter-stimulus interval (ISI) of 1 s was used between each word in a comparison pair. A longer ISI is shown to facilitate phonemic discrimination rather than phonetic discrimination of contrasts that are absent from the L1 (Werker and Logan 1985). The participants were allowed to listen to the same items again if needed, but they could not change an answer once given (Guion et al. 2000). The trials were presented in two blocks during a twenty-minute session. A three-minute interval separated the two blocks.

To ensure that the participants understood the task procedure, they were presented with a familiarization task prior to the experiment. They were trained to listen to and judge two contrasts (/ʃ/ vs. /s/, /t/ vs. /d/) and were given immediate feedback about the accuracy of their responses. The contrasts presented in the familiarization task were different from those used in the real test to avoid providing the participants with help on the target stimuli (Beddor and Gottfried 1995). Adopting a similar approach to Aoyama et al. (2004), the participants had to respond correctly to at least 90% of the stimuli in order to proceed to the actual task. If a participant did not reach this standard, the practice task was repeated up to four times or until they met the standard. Two FDH who performed below 90% in the familiarization task were excluded from this study because they did not display understanding of the task. A similar procedure was used to elicit NSs' responses to the same task. However, the NSs were given an answer sheet to record their own responses.

### 2.3. Examining FDH's Arabic Production

The 22 FDH participants who were recruited for the previous task participated in the picture-naming task.

#### 2.3.1. Stimuli for the Production Task

A picture-naming task was used to elicit single words from the participants. A list of Arabic words that contained the same target Arabic consonants that were used in the AX discrimination was created for this task. Another list that included words with onset and coda consonant clusters was also created. The words selected represented home objects and hence were assumed to be familiar to most FDH. Thirty-six pictures that represented the stimuli words were used to elicit productions from the FDH. The pictures were compiled in a Powerpoint file, one picture per slide.

#### 2.3.2. Procedure for the Production Task

The FDH named the objects in the pictures presented to them using the slideshow function of Powerpoint on a MacBook laptop controlled by the first author. The same high-quality recorder and microphone used in task 1 were used again here to record the participants' productions. When a participant struggled to name an object, a delayed repetition technique was used (Ratner 2000; Guion et al. 2000): if the participant could not name the object in the picture, the first author produced the target word and then asked the participant what the prompt was again. The delay between the prompt and the participant's repetition was approximately 4 s in order to minimize the effect of direct mimicry. The productions were analysed for their target-like accuracy; additionally, clusters were examined for simplification patterns such as epenthesis or deletion of one of the consonants.

### 2.4. Examining FDH's Foreign Accent Rating

#### 2.4.1. Listeners

The listeners for this task were 10 L1 Omani speakers (5 males and 5 females). Their ages ranged between 31 and 40 years at the time they carried out the task. They were all born in Oman and spoke Omani Arabic. None had experience or training in a linguistic-related field. None reported having a history of speech, language or hearing problems.

#### 2.4.2. Material for the Accent Rating Task

The stimuli for this task were taken from production data collected from the picture-naming task, with a focus on words where the target sound of interest was in a word-initial position. Ten words were selected for inclusion in this experiment (/χass/'lettuce', /ħabll/'rope', /qalam/'pen', /ð<sup>s</sup> arf/'envelope', /s<sup>s</sup> aabu:n/'washing liquid', /t<sup>s</sup> a:wleh/'table', /<sup>ʁ</sup> arʃeh/'bottle', /ʕ<sup>s</sup> alam/'flag', /θo:m/'garlic', /ðurah/'corn'). The stimuli were extracted from the sound files of the picture-naming task in Praat as whole words. The defined stimuli were windowed by a parabolic function and normalized to 50 dB. Two L1 Omani speakers aged 33 and 35 and who spoke Omani Arabic also produced the same words. The total number of stimuli included in the rating experiment was 240. The stimuli were randomized and distributed in three blocks (each including 80 stimuli).

#### 2.4.3. Procedure for the Accent Rating Task

The stimuli were presented to the participants in Praat on a Macbook Pro laptop. The listeners were asked to rate the stimuli they heard on a scale from 1 (not at all native-like) to 9 (completely native-like). They knew that they were going to hear Arabic words produced by L1 and L2 speakers. They could replay the stimuli as many times as they wished before making their choices and moving on to the next stimulus by pressing the 'next' button. They were offered a break after every 80 stimuli. The listeners spent approximately 30–40 min on this experiment.

## 2.5. Statistical Analysis

Data from the AX discrimination task was coded and tabulated manually in Excel, and subsequently imported into R (R Development Core Team 2012). To measure listeners' accuracy and correct for biases, we used some variants from Signal Detection Theory (Macmillan and Douglas 1991). We generated the *d prime* value (sensitivity index) for each individual listener separately. The *d prime* models the difference between the 'true positive' responses and 'false positive' responses in standard units, as in the following formula:  $(d \text{ prime}) = Z(\text{True Positive Rate}) - Z(\text{False Positive Rate})$ . We then used a linear model to statistically measure the difference in the *d prime* mean of NS and FDH groups using the lme4 package in R (Bates et al. 2015). To examine which factors contribute to any differences in *d prime* values among FDH listeners, we used a linear model, with *d prime* as the dependent variable and LoR (continuous), L1 schooling (continuous) and L2 literacy (categorical) as the independent variables (predictors). The model used was diagnosed for the presence of multicollinearity using the variance inflation factor, *vif()*, that measures the influence of collinearity among the predictors in a regression model (Midi et al. 2010). The VIF scores obtained were low (equal to 1), which indicated that it was safe to accurately assess the contribution of the predictors to the model.

To analyse data from the picture-naming task, the productions were first phonetically transcribed in Praat following the labelling of target segments. A Praat script designed by the first author was then used to extract all target word productions and their relevant information from Praat and transfer these to Excel files. Target consonants were assigned a value of 0 or 1 depending on whether the production of the sound was target-like. Descriptive statistics based on the percentages of target-like productions of each consonant averaged across speakers were then provided. To examine which factors play a role in the accuracy of productions, a GLMM was used with LoR, L1 schooling and L2 literacy as predictors and random intercept of speakers as random effects. The model was detected for multicollinearity using VIF. The scores obtained from the VIF were low, indicating the low collinearity of the predictors.

With regard to consonant cluster productions, a GLMM was used to examine the difference in modification between onset and coda consonant clusters. First, consonant clusters were assigned a value of 0 or 1 depending on whether the cluster was in the onset or coda position. Next, the productions of the consonant clusters were assigned a value of 0 or 1 depending on whether the production involved modification<sup>1</sup>. The GLMM had *modification* (yes or no) as a dependent variable and the type of consonant cluster (onset or coda) as predictor. For random effects, speaker was used as random intercept. To examine the effect of psycho-social factors on the pattern of cluster production, onset and coda clusters were analysed differently. For onset clusters, qualitative analysis was carried out because L1 inventories of some of the FDH's contained onset clusters while others did not (see Table 2). This indicates that the L1s of FDH's are incomparable and will affect their production patterns differently. Therefore, using statistical analysis with factors such as L1, LoR, L1 schooling and L2 literacy might not generate meaningful results, and one factor could cancel the significance of other factors (collinearity effects). Therefore, findings from a qualitative analysis will be more reliable and can be indicative of patterns that we can test more thoroughly in future research. As for coda clusters, a GLMM was used to test the effect of LoR, L1 schooling and L2 literacy on the modification of consonant clusters. All FDH's L1 inventories lacked coda clusters, and thus FDH's were considered comparable in relation to their L1s. All three factors were used as predictors, with speaker as random intercept for random effects.

For the foreign accent rating task, raters' responses were first tabulated in an Excel file which was then imported into R. Descriptive statistics were generated in R and included mean, median, SD and variance. In order to determine whether there is a difference in rating scores among the two groups (NS and FDHs), we used cumulative link mixed models (CLMM), using the *ordinal* package in R (Christensen 2019). For the dependent variable, we used an ordered factor, the rating response. For the independent variable, we



used group. For random effects, we used rater and item as intercepts. This was the optimal random effect structure that suited these data.

### 3. Results

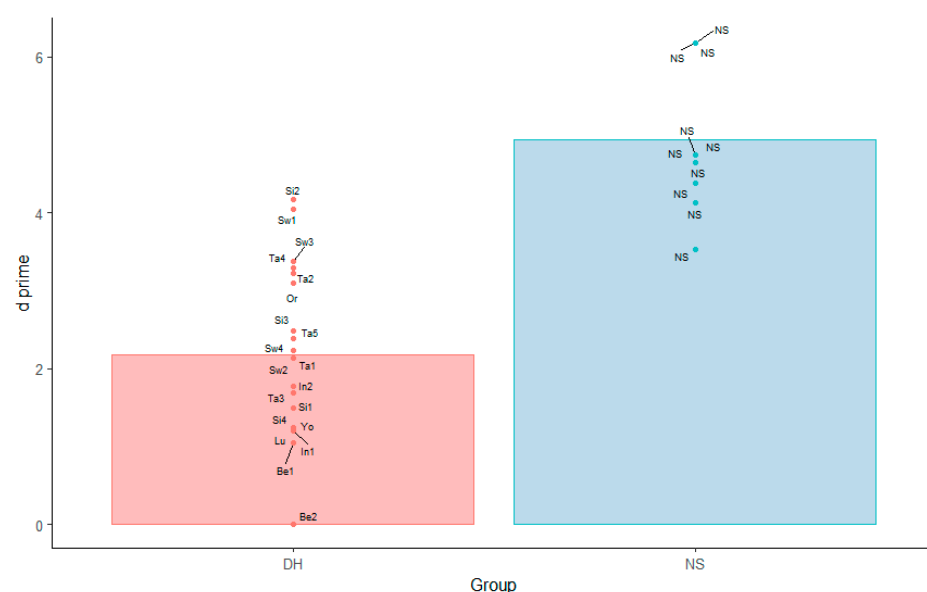
#### 3.1. Discrimination of Arabic Consonant Contrasts

Table 3 shows descriptive statistics that were obtained from a number of operations. The accuracy of the FDH was 0.72, while that of NSs was 0.96. The error rate of FDH (0.27) was higher than that of NSs (0.03). The true positive rate reflects the rate at which listeners responded ‘different’ when the stimulus was ‘different’. On the other hand, the false positive rate reflects the rate at which listeners responded ‘different’ when the stimulus was ‘same’. The sensitivity (which reflects the rate at which listeners responded ‘same’ when the stimulus was ‘same’) of both groups was higher than 50%. Precision reflects the rate at which listeners responded correctly when the stimulus was ‘same’. This was also higher than 70% for both groups. These statistics show that regardless of the difference between FDH and NS groups with regard to accuracy and sensitivity rates, the FDH group has successfully obtained high accuracy and sensitivity results (above 50%).

**Table 3.** Descriptive statistics obtained from the AX discrimination responses.

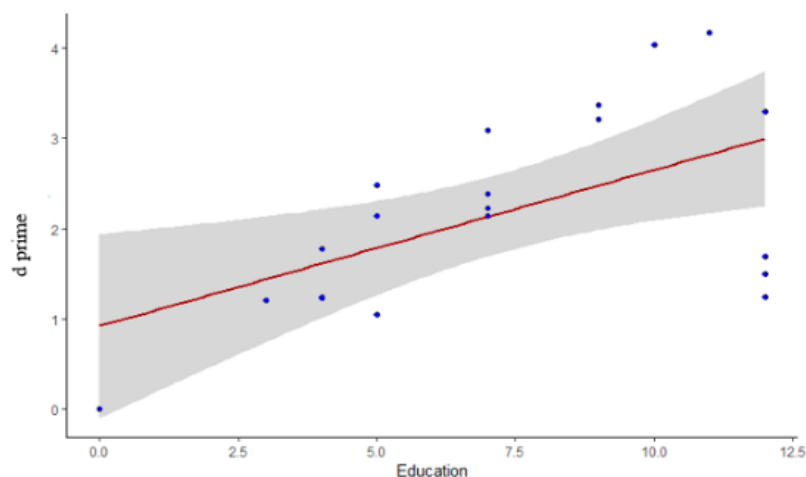
Group	Accuracy	Error	True Positive Rate (Specificity)	False Positive Rate	True Negative Rate (Sensitivity)	Precision
FDH	0.72	0.27	0.64	0.13	0.86	0.89
NS	0.96	0.03	0.98	0.06	0.93	0.96

When examining the difference between FDH and NSs’ discriminability of consonant contrasts using *d prime* values, a linear model revealed that the mean *d prime* value of the FDH group (mean = 2.17) was significantly lower than that of the NS group (mean = 4.94,  $\beta = -2.76$ ,  $SD = 0.404$ ,  $p < 0.01$ ). Unsurprisingly, this indicates that the NS group outperformed the FDH group in the AX discrimination task. However, the FDH *d prime* values in Figure 1 reveal a great variation among FDH’s performance (Figure 1), suggesting that some FDH performed as well as some NSs while others performed very poorly.



**Figure 1.** Mean *d prime* scores of the foreign domestic helper listeners (DH) and the native speaker (NS) control group.

Further analyses to examine the factors that affected FDH's variable performance in the AX discrimination task revealed that FDH's L1 schooling played a significant role in their *d prime* scores ( $\beta = 0.18$ ,  $SE = 0.06$ ,  $p = 0.01$ ). The more years the FDH had spent at school in her first language, the greater her discriminability of the contrastive sounds in Arabic was (Figure 2). LoR had no significant effect on FDH's discriminability of consonantal contrasts ( $\beta = 0.02$ ,  $SE = 0.03$ ,  $p > 0.05$ ). Thus, no matter how long the FDH had spent in the Arabic-speaking world, her discriminability of consonant contrasts had not changed or improved. Similarly, L2 literacy did not have a significant effect on FDH's *d prime* scores ( $\beta = 0.25$ ,  $SD = 0.44$ ,  $p > 0.05$ ).



**Figure 2.** The relationship between years of L1 schooling and *d prime* values relevant to the FDH group.

### 3.2. The Production of Arabic Consonants and Consonant Clusters

As expected, complex consonants that are specific to the Arabic inventory posed the greatest challenge for the participants (Figure 3). Common realizations included stopping for fricatives, e.g., [t] for /θ/, [d] for /ð/, [g] or [k] for /ʁ/, [q] for /χ/, [ʔ] for /ʕ/, and [d<sup>ʕ</sup>] for /ð<sup>ʕ</sup>/; de-emphasis, e.g., [ð] for /ð<sup>ʕ</sup>/, [t] for /t<sup>ʕ</sup>/ and [s] or [ʃ] for /s<sup>ʕ</sup>/; (de-)voicing (along with other processes), e.g., [q] for /ʁ/, [g] for /χ/; fronting/backing (along with other processes), e.g., [k] or [g] for /q/ or /χ/, [t] for /q/, [z] or [s] for /ð/ or [s] for /θ/; and weakening, e.g., [h] for /ʕ/ or /ʁ/ deletion. L2 literacy had a significant effect on FDH's accurate productions of the target sounds ( $\beta = -0.53$ ,  $SE = 0.22$ ,  $p = 0.01$ ). Figure 4 shows that speakers who were literate in Arabic had more target-like productions (50%) of the target consonants than those who were non-literate (37.8%). This suggests that FDH who learned Arabic via recitation of the Qur'an performed better than those who did not. On the other hand, LoR did not play any significant role in the target-like production of Arabic consonants by FDH ( $\beta = 0.02$ ,  $SE = 0.01$ ,  $p > 0.05$ ). In fact, the visual examination of the results showed that speakers with the longest LoRs appeared to be slightly less accurate than those with shorter LoRs. Similarly, L1 schooling did not play a considerable role in FDH's accuracy ( $\beta = 0.03$ ,  $SE = 0.03$ ,  $p > 0.05$ ).

Results showed a high proportion of modified consonant clusters in the production of FDH in the onset and coda position (Table 4). The main strategy used to modify clusters was vowel epenthesis. This indicates that FDH have not acquired complex clusters despite their exposure to them in the target input (Al-Kendi 2021). The tendency to modify onset consonant clusters (89.4%) was more frequent than for coda clusters (48.9%) (Figure 5). A GLMM demonstrated that this difference was significant ( $\beta = -2.51$ ,  $SE = 0.41$ ,  $p < 0.01$ ).

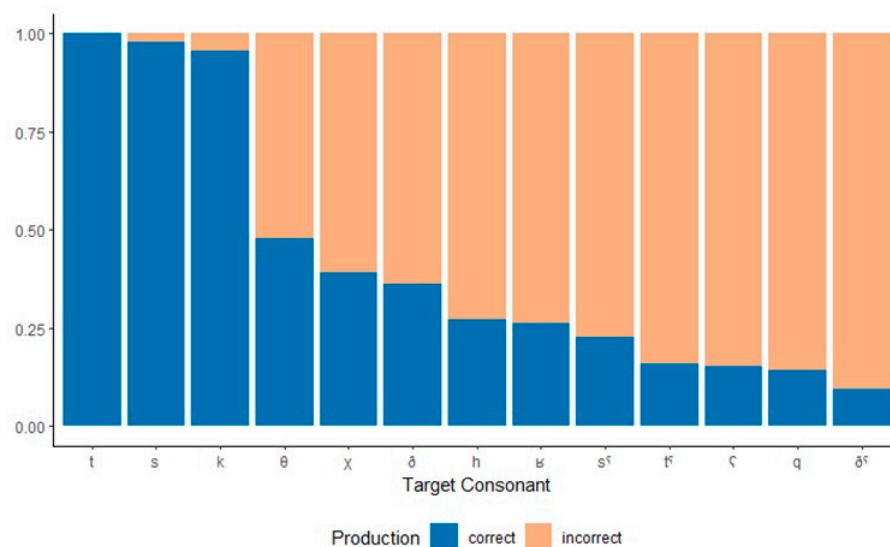


Figure 3. Accuracy in the productions of the target Arabic consonants by FDH.

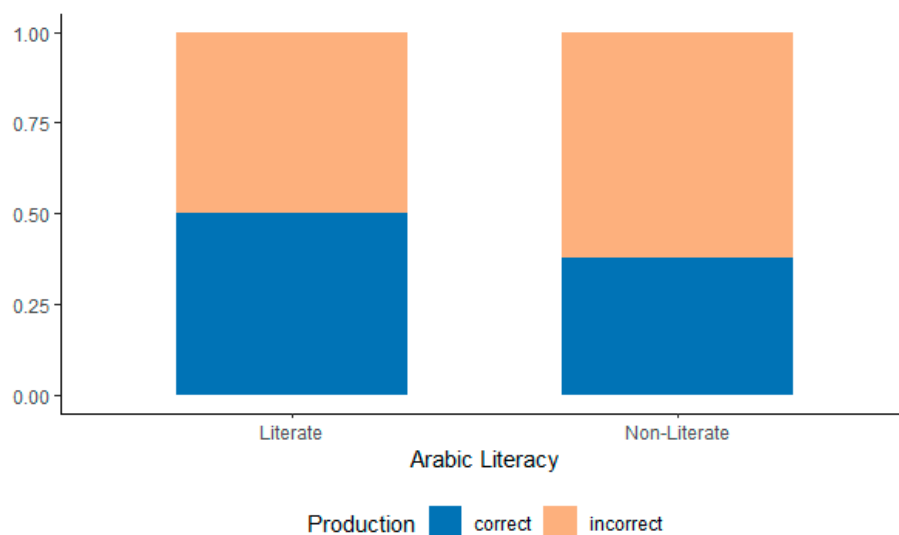


Figure 4. Percent accuracy in the productions of the target Arabic consonants by literate and non-literate FDHs.

Table 4. Percentage of modification in onset and coda consonant clusters (CC) in FDH’s productions and the modification strategy used.

Syllable Type	Syllable Modification		Modification Strategy	
	Yes	No	Consonant Deletion	Vowel Epenthesis
Onset CC	(118) 89.4%	(14) 10.6%	(2) 1.69%	(116) 98.305%
Coda CC	(43) 48.9%	(45) 51.13%	0	(43) 100%

In terms of factors that could have affected FDH production of onset consonant clusters, descriptive statistics and visual inspection showed that, generally, the tendency to produce less marked onset consonant clusters was evident in all FDH’s productions regardless of the L1 (Figure 6). Nevertheless, FDH with Oromo, Sinhala and Telugu L1 backgrounds had the highest rate of onset modification, producing most of the target CCs with epenthetic vowels. FDH with Indonesian, Tagalog, Bengali, Yoruba, Swahili and Luganda L1 backgrounds showed more variation in their production of onsets clusters,

sometimes maintaining the CC realization. To illustrate, these FDH sometimes produced the target word /kta:b/ as [kta:b] and others as

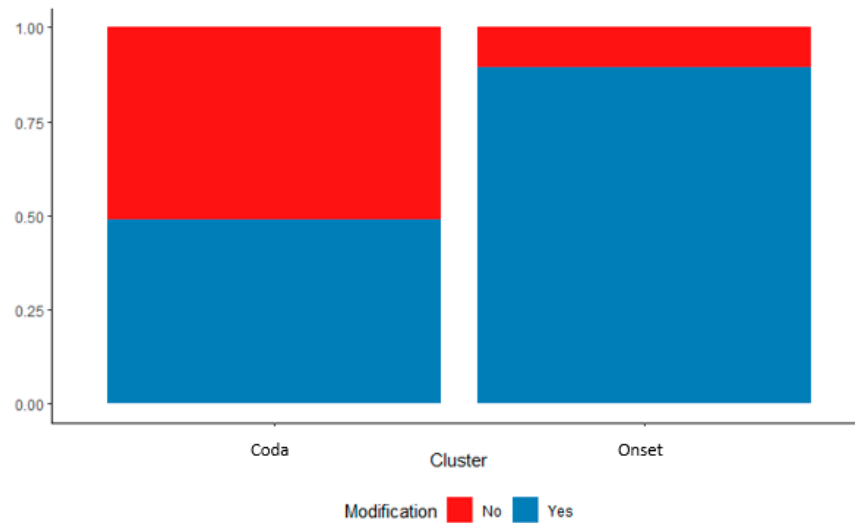


Figure 5. Percentage of onset and coda consonant clusters modification.

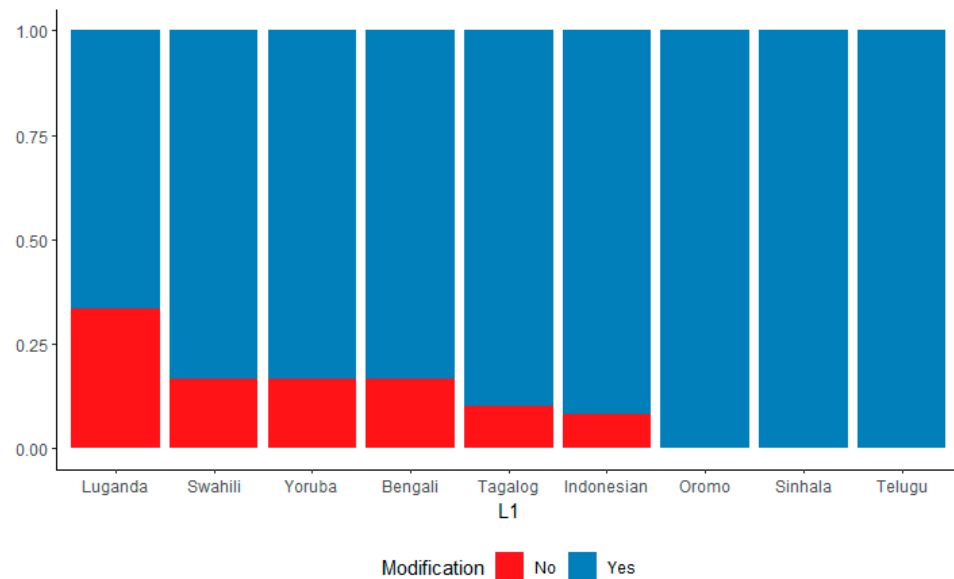


Figure 6. Production accuracy of onset CCs as a function of the L1 of the FDH participants.

Figure 7a shows patterns of onset cluster production by FDH as a function of LoR. Surprisingly, fewer rates of onset simplification are evident in productions of FDH who had the shortest LoR. The trend then shows a stable pattern for FDH with 5 to 15 years of LoR and is then highest for the speakers with the highest LOR. From this, we can conclude that LoR alone does not play a role in the successful production of onset consonant clusters in FDH’s productions. Equally, the modification of onset consonant clusters does not appear to change considerably as a function of years of formal schooling. Figure 7b shows a stable trend of onset cluster simplification regardless of the years of schooling.

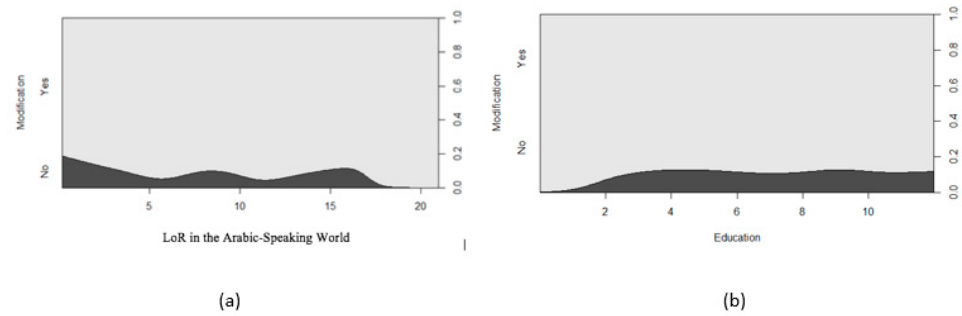


Figure 7. The effects of LoR (a) and L1 schooling (b) on FDH’s production accuracy of onset CCs.

Figure 8 illustrates that FDH who were literate in Arabic exhibited fewer modifications of onset consonant clusters (7.77%) than those who were non-literate in Arabic (16.66%).

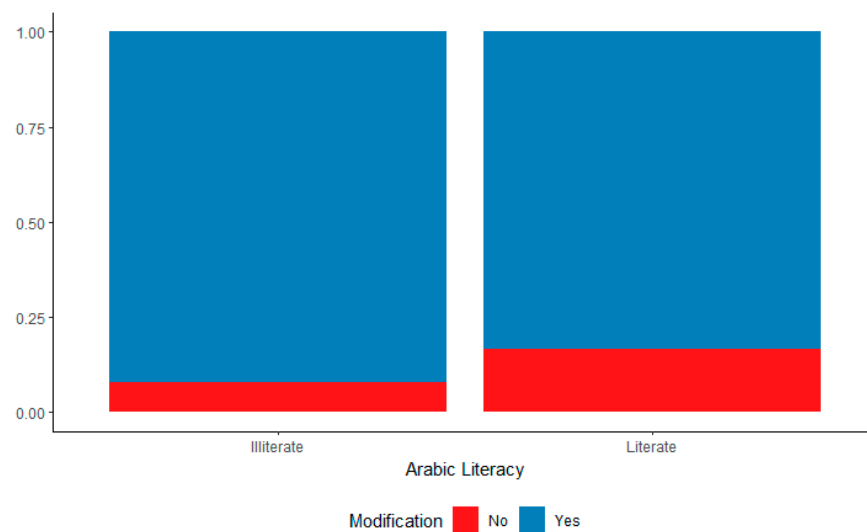


Figure 8. The percentages of target-like productions of Arabic onset consonant clusters by literate and non-literate FDHs.

Moving on to factors that may have played a role in FDH’s production of coda consonant clusters, the results once again revealed that LoR had a significantly negative effect on target-like consonant cluster production ( $\beta = -0.102, SE = 0.04, p = 0.03$ ). FDH with the shortest LoR produced clusters in the coda position more frequently than those with a longer LoR (Figure 9a). The trend also shows fluctuation, which implies individual differences in the target-like realization of coda clusters. L1 formal education had a significant effect on the modification of coda consonant clusters ( $\beta = 0.13, SE = 0.06, p = 0.04$ ). The more educated a foreign domestic helper was, the more successful she was at producing a target-like syllable structure (Figure 9b). This is in line with the significant role L1 schooling played in FDH’s performance in the discrimination task.

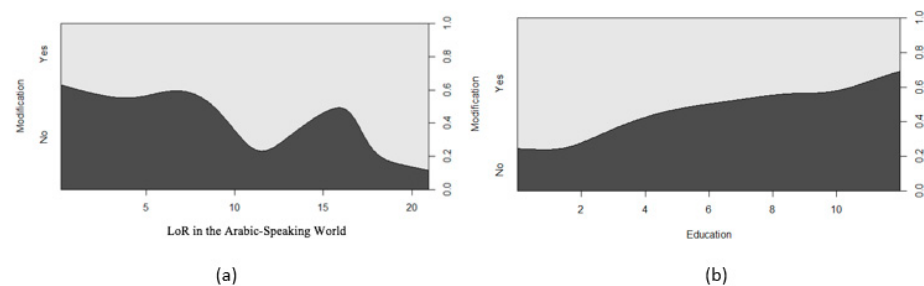


Figure 9. The effects of LoR (a) and L1 (b) schooling on FDH’s production accuracy of coda CCs.

L2 literacy, on the other hand, did not play any significant role in the pattern of coda consonant cluster production ( $\beta = 0.85, SE = 0.55, p > 0.05$ ), though there was a trend for more target-like production by FDH who reported to be literate in Arabic than those who reported to be non-literate (Figure 10). Literate FDH produced coda clusters in 64.28% of the instances, while non-literate FDH produced them in 45% of the instances. Hence, literacy in Arabic—or, more specifically, knowledge of the Arabic script—seems to aid the L2 learners’ acquisition of target-like oral forms, even though this trend was not significant.

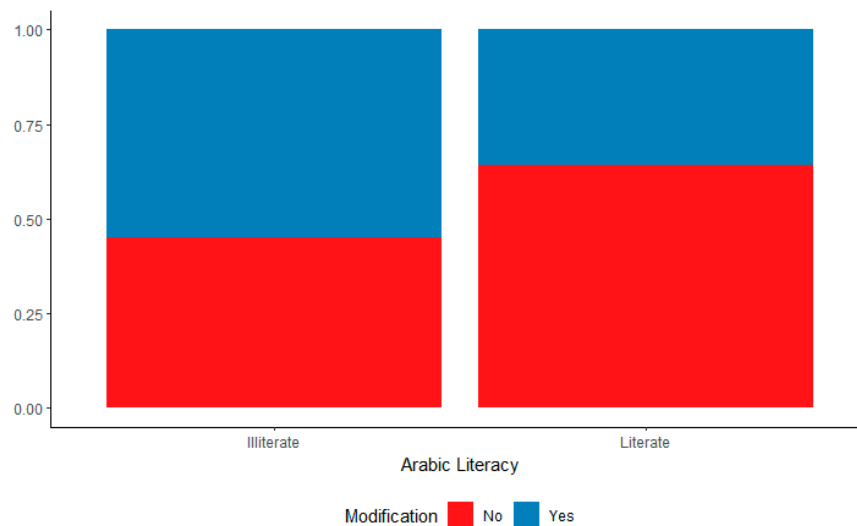


Figure 10. Production accuracy of Arabic coda consonant clusters by literate and non-literate FDHs.

### 3.3. Foreign Accent Rating of FDHs

Figure 11 illustrates the rating scale of the foreign accent task and the percentage of total scores given to words produced by the two groups (NS and FDHs). It shows that around 80% of the ratings given to NSs fell into the ‘completely native-like’ category, that is number 9 on the scale. However, the highest percentage of rating in the FDH group went to the ‘not at all native-like’ ranking on the scale, that is number 1.

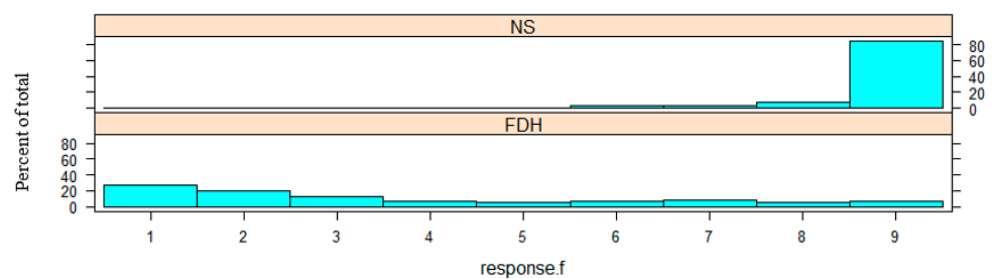
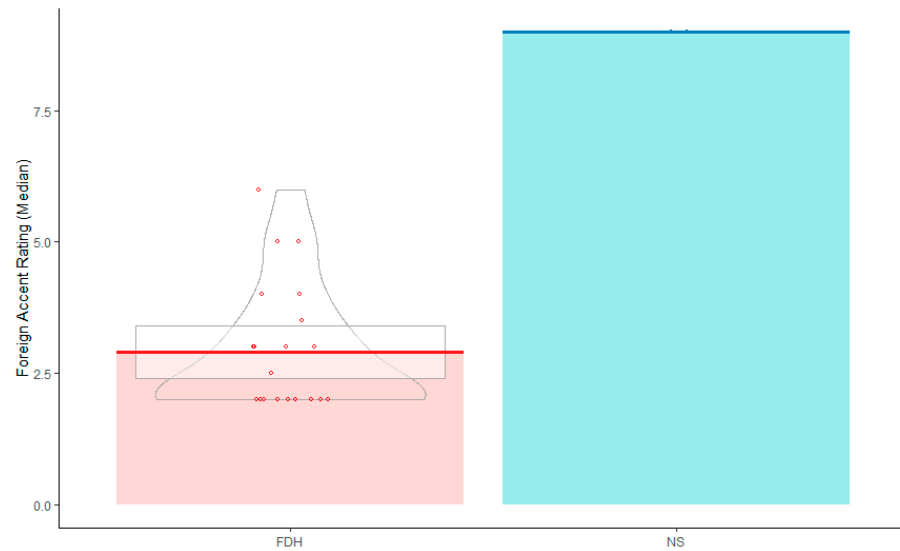


Figure 11. Percentages of total responses for each rate in the foreign accent scale for both FDH and NS productions.

Descriptive statistics of the foreign accent rating task indicate that the FDH were rated very low on the foreign accent rating scale (median = 3) compared to the NS group (median = 9), as shown in Table 5. There was, however, considerable variation in the rating scores given to the FDH production (variance = 6.86) compared to those given to the NSs (variance = 1.49). Further analyses of these results revealed that the difference between NSs and FDH’s foreign accent rating was significant ( $\beta = 5.31, SE = 0.22, p < 0.01$ ). Figure 12 illustrates the difference in the foreign accent rating median for the NS and the FDH groups.

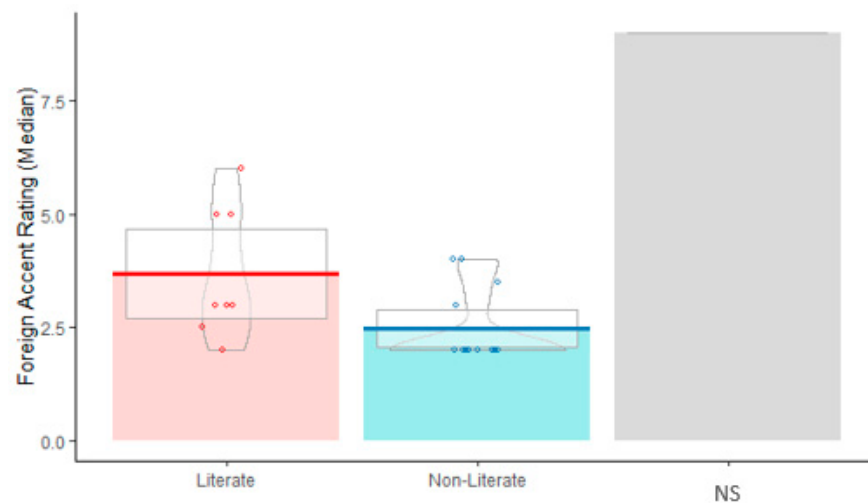
**Table 5.** Descriptive statistics obtained from the foreign accent rating task for both the FDH and the NS groups.

Group	Mean	Median	SD	Variance
FDH	3.64	3	2.62	6.86
NS	8.61	9	1.22	1.49



**Figure 12.** Median and distribution of foreign accent rating scores given to NS and the FDH groups.

When examining the factors that may have affected foreign accent rating, the results revealed that FDH’s L2 literacy played a significant role in their foreign accent rating (Figure 13). Non-literate FDH’s were rated as more foreign-accented than literate FDH ( $\beta = -1.24, SE = 0.49, p = 0.02$ ). However, LoR did not play any significant effect on FDH’s accent rating ( $\beta = -0.03, SE = 0.04, p > 0.05$ ). Likewise, L1 schooling was not found to significantly affect FDH’s foreign accent rating ( $\beta = -0.03, SE = 0.06, p > 0.05$ ). Interestingly, these results are similar to those obtained from the examination of FDH’s production of Arabic consonants.



**Figure 13.** Median and distribution of foreign accent rating scores given to NS and the FDH groups, the latter split according to L1 literacy.

#### 4. Discussion

Despite a constant exposure to Arabic from L1 speakers, the length of residence that FDH spent in the Arab world appeared to play no role in their L2 perception of consonant contrasts or in their production of Arabic consonants or consonant clusters. In some cases, FDH's scores correlated negatively with LoR, as in the production of coda consonant clusters. These findings do not support the assumption that the more L2 input one receives the better the opportunities to master the L2 (McAllister 2001). There have been other studies over the years that have challenged LoR effects on learners' performance (e.g., Flege 1988; Oyama 1976), but their methodologies have rarely included a total immersion in the L2 with input from L1 speakers, as in the current study. Below we reflect on the potential reasons for these surprising results.

It is hard to ignore the different ways in which a lack of effect of LoR has been interpreted in the literature. For example, Flege and Liu (2001) suggest three possible interpretations: (1) the amount of L2 input is not a crucial predictor of L2 performance, (2) L2 performance is constrained by a critical or sensitive period, (3) LoR provides a meaningful index of L2 input for some individuals but not others. The age effect has been dealt with extensively in the literature, but it is the third point that we focus on here. On the one hand, one can interpret these differences in terms of differential access to input from L1 speakers. For instance, Flege (2002) found that LoR can play a role in adults' L2 performance only if they are exposed to a considerable amount of L1 speaker input. In this study, Chinese students with longer LoR in the United States were significantly better in their L2 performance compared to students who had shorter LoR. LoR, however, did not predict the performance of non-students. Flege (2002) concluded that because students had more opportunities for receiving NS input compared to the non-students, their performance improved noticeably over time. However, Moyer (2004, 2009) argues that, in order for late learners to gain sufficient input, they need to engage in the L2 environment in different ways. Situations favourable to such attempts vary across individuals, depending not only on age but also on educational, social and ethnic background. Moyer (2004) further suggests that for LoR to reliably index the L2 experience, an integrated approach that takes into account cognitive, psychological and social factors needs to be carried out. Psychologically, LoR correlates with a sense of overall fluency and satisfaction with L2 attainment as well as motivation to learning the L2. Socially, LoR correlates with the frequency of contact with NSs and the intention for permanent residency in the L2 target community. Cognitively, LoR correlates with L2 instruction and communicative use of the L2 rather than just focusing on form, as well as the amount of feedback on pronunciation and the kind of phonological training.

In light of this integrated model, it is not surprising that LoR was not found to play a role in FDH's phonological performance. Despite the high number of years many FDH spend in the Arab world, their intention is not for permanent residence in the L2 country, but rather to return home once they have saved enough to support their families (Bizri 2014). Their aim of L2 attainment may be restricted to the ability to interact with their employers rather than any motivation to achieve native-like fluency. Furthermore, FDH's language contact with their employers or with other family members is often restricted to conversations around home chores. Due to the task-oriented nature of these interactions, it is unlikely that FDH receive any feedback on their pronunciation. Equally, they do not receive any training on L2 phonology or other linguistic aspects. In addition, the input they receive can be variable within the constricted context in which they work: they attend to children who have not yet fully developed their phonology and hear accented Arabic from other FDH of various nationalities when running errands and during their day off. This is akin to Flege and Liu's (2001) description of immigrants to North America, who are likely to use their L2 English with other NNSs as a lingua franca, though in their study those interactions were happening in the workplace, whereas our participants' main workplace is their NS employers' home.

When the above factors are considered, FDH's experience does not provide an optimal environment for L2 attainment, despite the near total immersion in input from L1 speakers.



Opportunities for meaningful input, contact with other L1 speakers and L2 instruction are very limited in this context, also highlighting the potential role of the lack of variability in the input. This supports the observation that LoR is not a reliable index of L2 experience (Flege and Liu 2001; Moyer 2004, 2009). In order to reliably examine the extent to which input and LoR modulate L2 phonological performance, methodologies always need to take into account cognitive and socio-psychological factors that may shape the L2 speakers' experience.

While LoR was not a significant predictor of FDH's phonological performance, L1 schooling and L2 literacy each played a role in some of the variables examined. L1 schooling correlated significantly with FDH's perceptual sensitivity scores and rate of final consonant cluster productions. However, no effect of L1 schooling was found in the production of Arabic consonants or initial consonant clusters. L2 literacy generally correlated positively with sensitivity scores and production of Arabic consonants as well as initial and final consonant clusters. However, the results were only significant with regard to the production of L2 consonants. The differential roles of L1 formal schooling and L2 literacy in FDH's performance are discussed below.

In terms of the positive role of L1 schooling on the perception task, a first justification for this finding is that the AX discrimination task was an experimental paradigm that required the listeners to understand task instruction and procedure as well as sit for an actual test. This protocol might be more familiar to adults who had attended school and experienced such a situation compared to adults who had little to no experience with carrying out cognitively demanding tasks due to not attending school. Another positive effect of L1 literacy on L2 perception may be due to the higher level of phonological awareness that comes with learning an alphabetic script (e.g., Morais et al. 1979; Adrián et al. 1995; Tarone and Bigelow 2005), and which may have equipped the participants with metalinguistic skills that they could subsequently use in their L2.

An emerging body of work has recently highlighted the role of L1 orthography in L2 production (rather than perception), with results suggesting that L1 orthography leads to a convergence between the L1 and the L2 production patterns (e.g., Bassetti and Atkinson 2015; Escudero et al. 2014; Nimz and Khattab 2020). Our results do not demonstrate strong effects of L1 literacy on L2 production, with the only difference seen in the greater target-like production of final consonant clusters by participants who were literate in the L1. One reason for this may be due to the higher prevalence of CC realizations of consonant clusters in the final than in the initial position in the L2 input that the FDH receive (Al-Kendi 2021), rendering these structures more salient. The production of consonant clusters requires attention to structures in the target input and adjusting L1 phonology accordingly. Schmidt's Noticing Hypothesis (Schmidt 1990) claims that a conscious awareness (i.e., noticing) of the input plays a substantial role in the process of language acquisition. Several researchers have provided support to this hypothesis and confirmed the importance of noticing for language learning (e.g., Jeremy 2002; Lynch 2001; Skehan 1998). Among the factors that Schmidt claims to affect noticing is frequency, and hence the likelihood of FDH's noticing coda more than onset clusters. L1 schooling may have helped FDH develop a language learning aptitude and conscious phonological processing, improving the literate learners' 'noticing' skills that are required for L2 learning (Granena and Long 2013).

It is with L2 literacy that we see a stronger effect on L2 production. There has been a notable increase in attention to the role of orthography in L2 speech learning (e.g., Bassetti et al. 2015; Escudero et al. 2014; Nimz and Khattab 2020), but results are typically mixed, signalling both facilitatory and inhibitory effects in terms of learning L2 phonological categories. Here it is worth focusing on the unusual way in which Arabic literacy is taught for religious purposes to speakers of other languages, like the FDH in this study. While the script is of course key, there is a strong focus on recitation and rote learning in such contexts (Binte Faizal 2019; Supriyadi and Julia 2019), emphasizing the role of production practice in this process. This is likely to have helped FDH who had experience with L2 literacy,

supporting their production of L2 consonants through more advanced motor control along with the establishment of categories for new sounds (e.g., [Guenther 1994](#); [Flege 1995](#)).

Note, however, that the L2 literacy effect was mainly seen in single consonant production, with clusters requiring much more motor control and experience with the language before target-like realization. Here it is worth noting that the epenthesis of CC clusters is more common in the onset than in the coda position in Omani Arabic ([Al-Kendi 2021](#)), and this is reflected in the patterns found for CC realizations by FDH in those two contexts, albeit with a much higher occurrence of epenthesis in the FDH's production. This supports the expectation that learners will acquire less complex L2 structures (e.g., CV) before more complex ones ([Anderson 1987](#); [Eckman 1985](#); [Rice 2007](#); [Zec 2007](#)). This was also reflected in the productions of the speakers whose L2 systems lacked onset consonant contrasts compared with those whose L2 had such contrasts.

## 5. Conclusions

The results from the present study shed light on the perception and production ability of a group of uninstructed low-educated foreigners acquiring the language in a naturalistic setting. The status quo in SLA research has been to investigate the cross-linguistic performance of highly educated adults and to focus on LoR and AoA as primary factors affecting these learners' performance. The results from this study highlight the importance of looking at low-literacy learners and investigating the role of other non-linguistic factors, such as the nature of daily social interactions in an L2 context, the long-term aims of the learners and the power relations between them and their main interlocutors. Despite being totally immersed in Omani Arabic for a number of years, the FDH in this study still struggled with the phonology of Omani Arabic and had a pronounced foreign accent as judged by L1 listeners. Their perception scores were not found to be influenced by LoR or L2 literacy, but rather by the amount of their L1 schooling; this could in itself be a proxy for learning to perform tasks and follow instructions, but L1 literacy may have also increased these learners' metalinguistic awareness. Their production patterns did not show any LoR effect either, and only a modest influence from L2 instruction. The study demonstrates how difficult it is to control for external factors that are beyond the focus of a given study in SLA research. For instance, while the focus of the current study was to investigate what looked like an optimal case of LoR with guaranteed input in order to address previous criticisms of LoR, low literacy and the fact that input and interaction are dominated by the employer and their family may have attenuated any LoR effects, showing how multi-faceted the L2 speech learning experience is.

**Author Contributions:** Conceptualization, A.A.-K. and G.K.; methodology, A.A.-K. and G.K.; formal analysis, A.A.-K.; resources, A.A.-K.; data curation, A.A.-K.; writing—original draft preparation, A.A.-K.; writing—review and editing, G.K.; visualization, A.A.-K.; supervision, G.K.; project administration, A.A.-K.; funding acquisition, A.A.-K. Both authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by grants from Sultan Qaboos University.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of the School of English Literature, Language and Linguistics at Newcastle University (Faculty of Humanities and Social Sciences Ethics Committee, approved on the 9th of June 2017).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding authors. The data are not publicly available due to ethical restrictions.

**Acknowledgments:** We thank Martha Young-Scholten for comments on an earlier version.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## Notes

- <sup>1</sup> This refers to whether the consonant cluster was maintained or modified by, for example, by epenthesis of a vowel to break it up or omitting one of the consonants.

## References

- Adrián, José Antonio, Jesús Alegria, and José Morais. 1995. Metaphonological abilities of Spanish illiterate adults. *International Journal of Psychology* 30: 329–51. [CrossRef]
- Akahane-Yamada, Reiko. 1995. Age and acquisition of second language speech sounds perception of American English /r/ and /l/ by native speakers of Japanese. In *Speech Perception and linguistic Experience*. Baltimore: York Press.
- Alhawari, Mohammed, ed. 2018. *The Routledge Handbook of Second Language Acquisition*. New York: Routledge.
- Al-Kendi, Azza. 2021. Foreigner-Directed Speech and L2 Speech Learning in an Understudied Interactional Setting: The Case of Foreign-Domestic Helpers in Oman. Ph.D. thesis, Newcastle University, Newcastle upon Tyne, UK.
- Anderson, Janet. 1987. The markedness differential hypothesis and syllable structure difficulty. In *Interlanguage Phonology: The Acquisition of a Second Language Sound System*. Edited by Georgette Ioup and Weinberger Steven. Cambridge: Newbury House, pp. 279–91.
- Aoyama, Katsura, James Emil Flege, Susan G. Guion, Reiko Akahane-Yamada, and Tsuneo Yamada. 2004. Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics* 32: 233–50. [CrossRef]
- Bassetti, Bene, and Nathan Atkinson. 2015. Effects of orthographic forms on pronunciation in experienced instructed second language learners. *Applied Psycholinguistics* 36: 67–91. [CrossRef]
- Bassetti, Bene, Paola Escudero, and Rachel Hayes-Harb. 2015. Second language phonology at the interface between acoustic and orthographic input. *Applied Psycholinguistics* 36: 1–6. [CrossRef]
- Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1–48. [CrossRef]
- Beddor, Patrice, and Terry Gottfried. 1995. Methodological issues in adult cross-language research. In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Edited by Winifred Strange. Baltimore: York Press, pp. 207–32.
- Best, Catherine T., and Winifred Strange. 1992. Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics* 20: 305–30. [CrossRef]
- Best, Catherine T., Gerald W. McRoberts, and Nomathemba M. Sithole. 1988. Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14: 345. [CrossRef]
- Bigelow, Martha, and Elaine Tarone. 2004. The role of literacy level in second language acquisition: Doesn't who we study determine what we know? *TESOL Quarterly* 38: 689–700. [CrossRef]
- Binte Faizal, Siti Syuhada. 2019. Visual Word Processing of Non-Arabic-speaking Qur'anic Memorisers. Ph.D. thesis, Newcastle University, Newcastle upon Tyne, UK.
- Bizri, Fida. 2014. Maids' Talk. *Colonization and Domestic Service: Historical and Contemporary Perspectives* 14: 113.
- Boersma, Paul, and David Weenink. 2009. Praat: Doing Phonetics by Computer. Available online: <http://www.fon.hum.uva.nl/praat/> (accessed on 30 September 2016).
- Christensen, Rune Haubo Bojesen. 2019. *Regression models for ordinal data [R package ordinal version 12–10]*.
- Craats, Ineke van de, Jeanne Kurvers, and Martha Young-Scholten. 2006. Research on low-educated second language and literacy acquisition. *LOT Occasional Series* 6: 7–23.
- DeKeyser, Robert M. 2000. The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22: 499–533. [CrossRef]
- Eckman, Fred R. 1985. Some Theoretical and Pedagogical Implications of the Markedness Differential Hypothesis. *Studies in Second Language Acquisition* 7: 289–307. [CrossRef]
- Escudero, Paola, and Paul Boersma. 2004. Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition* 26: 551–85. [CrossRef]
- Escudero, Paola, Ellen Simon, and Karen E. Mulak. 2014. Learning words in a new language: Orthography doesn't always help. *Bilingualism-Language and Cognition* 17: 384–95. [CrossRef]
- Escudero, Paola. 2006. The phonological and phonetic development of new vowel contrasts in Spanish learners of English. *English with a Latin Beat: Studies in Portuguese/Spanish-English Interphonology* 31: 41.
- Flege, James Emil, and Kathryn L. Fletcher. 1992. Talker and listener effects on degree of perceived foreign accent. *The Journal of the Acoustical Society of America* 91: 370–89. [CrossRef]
- Flege, James Emil, and Ocke-Schwen Bohn. 2021. The Revised Speech Learning Model (SLM-r). In *Second Language Speech Learning: Theoretical and Empirical Progress*. Edited by Wayland Rastree. Cambridge: Cambridge University Press, pp. 3–83. [CrossRef]
- Flege, James Emil, and Serena Liu. 2001. The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition* 23: 527–52. [CrossRef]
- Flege, James Emil, Murray J. Munro, and Ian R. A. MacKay. 1995. Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America* 97: 3125–34. [CrossRef] [PubMed]

- Flege, James Emil, Ocke-Schwen Bohn, and Sunyoung Jang. 1997. Effects of Experience on Non-native Speakers' Production and Perception of English Vowels. *Journal of Phonetics* 25: 437–70. [CrossRef]
- Flege, James Emil. 1981. The phonological basis of foreign accent: A hypothesis. *TESOL Quarterly* 15: 443–55. [CrossRef]
- Flege, James Emil. 1987. A critical period for L2 learning to pronounce foreign languages? *Applied Linguistics* 8: 162–77. [CrossRef]
- Flege, James Emil. 1988. Factors affecting degree of perceived foreign accent in English sentences. *The Journal of the Acoustical Society of America* 84: 70–79. [CrossRef]
- Flege, James Emil. 1995. Second language speech learning: Theory, findings, and problems. In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Edited by Winifred Strange. Timonium: York Press, pp. 233–77.
- Flege, James Emil. 2002. Interactions between the native and second-language phonetic systems. In *An Integrated View of Language Development: Papers in Honor of Henning Wode*. Edited by Burmeister Petra, Piske Thorsten and Rohde Andreas. Germany: Wissenschaftlicher Verlag Traer, pp. 217–44.
- Flege, James Emil. 2009. Give input a chance. In *Input Matters in SLA*. Edited by T. Piske and M. Young-Scholten. Bristol: Multilingual Matters, pp. 175–90.
- Flege, James Emil. 2018. It's input that matters most, not age. *Bilingualism Language and Cognition* 21: 919–20. [CrossRef]
- Gass, Susan M. 2013. *Second Language Acquisition: An Introductory Course*. New York: Routledge.
- Granena, Gisela, and Michael H. Long. 2013. Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research* 29: 311–43. [CrossRef]
- Guenther, Frank H. 1994. A neural network model of speech acquisition and motor equivalent speech production. *Biological Cybernetics* 72: 43–53. [CrossRef] [PubMed]
- Guion, Susan G., James E. Flege, Serena H. Liu, and Grace H. Yeni-Komshian. 2000. Age of learning effects on the duration of sentences produced in a second language. *Applied Psycholinguistics* 21: 205–28. [CrossRef]
- Haznedar, Belma, Joy Kreeft Peyton, and Martha Young-Scholten. 2018. Teaching adult migrants. *Critical Multilingualism Studies* 6: 155–83.
- Henrich, Joseph, Steven J. Heine, and Ara Norenzayan. 2010. Most people are not WEIRD. *Nature* 466: 29. [CrossRef] [PubMed]
- Højen, Anders. 2019. Improvement in Young Adults' Second-language Pronunciation after Short-term Immersion. In *A Sound Approach to Language Matters—In Honor of Ocke-Schwen Bohn*. Edited by Anne Mette, Hejná Michaela, Højen Anders, Jespersen Anna and Sørensen Mette. Aarhus: Department of English, School of Communication & Culture, Aarhus University, pp. 543–59.
- Holliday, Adrian. 2018. Native-Speakerism. In *The TESOL Encyclopedia of English Language Teaching*. Edited by J. I. Liontas. Hoboken: John Wiley & Sons, pp. 1–7. [CrossRef]
- Ioup, Georgette, Elizabeth Boustagui, Manal El Tigi, and Martha Moselle. 1994. Reexamining the critical period hypothesis: A case study of successful adult SLA in a naturalistic environment. *Studies in Second Language Acquisition* 16: 73–98. [CrossRef]
- Iverson, Paul, and Patricia K. Kuhl. 1996. Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/. *The Journal of the Acoustical Society of America* 99: 1130–40. [CrossRef]
- Jeremy, Cross. 2002. Noticing in SLA: Is it a valid concept. *TESL-EJ* 6: 1–9.
- Klein, Wolfgang, and Clive Perdue. 1997. The Basic Variety (or: Couldn't natural languages be much simpler?). *Second Language Research* 13: 301–47. [CrossRef]
- Kuhl, Patricia K. 1991. Human adults and human infants show a 'perceptual magnetic effect' for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics* 50: 93–107.
- Kurvers, Jeanne, Ton Vallen, and Roeland van Hout. 2006. Discovering features of language: Metalinguistic awareness of adult illiterates. *LOT Occasional Series* 6: 69–88.
- Lenneberg, Eric H. 1967. The biological foundations of language. *Hospital Practice* 2: 59–67. [CrossRef]
- Lombardi, Linda. 2003. Second language data and constraints on manner: Explaining substitutions for the English interdentals. *Second Language Research* 19: 225–50. [CrossRef]
- Lynch, Tony. 2001. Seeing what they meant: Transcribing as a route to noticing. *ELT Journal* 55: 124–32. [CrossRef]
- Macmillan, Neil A., and Creelman C. Douglas. 1991. *Detection Theory: A User's Guide*. Cambridge: Cambridge University Press.
- McAllister, Robert. 2001. Experience as a factor in L2 phonological acquisition. *Working Papers-Lund University Department of Linguistics* 49: 116–19.
- Midi, Habshah, Saroje Kumar Sarkar, and Sohel Rana. 2010. Collinearity diagnostics of binary logistic regression model. *Journal of Interdisciplinary Mathematics* 13: 253–67. [CrossRef]
- Morais, José, Luz Cary, Jesús Alegria, and Paul Bertelson. 1979. Does awareness of speech as a sequence of phones arise spontaneously? *Cognition* 7: 323–31. [CrossRef]
- Moyer, Alene. 2004. *Age, Accent, and Experience in Second Language Acquisition: An Integrated Approach to Critical Period Inquiry*. Buffalo: Multilingual Matters, vol. 7.
- Moyer, Alene. 2009. Input as a critical means to an end: Quantity and quality of experience in L2 phonological attainment. In *Input Matters in SLA*. Edited by Thorsten Piske and Martha Young-Scholten. Buffalo: Multilingual Matters, pp. 159–74.
- Moyer, Alene. 2014. What's age got to do with it? Accounting for individual factors in second language accent. *Studies in Second Language Learning and Teaching* 4: 443–64. [CrossRef]
- Nimz, Katharina, and Ghada Khattab. 2020. On the role of orthography in L2 vowel production: The case of Polish learners of German. *Second Language Research* 36: 623–52. [CrossRef]

- Ordem, Eser, and Erdogan Bada. 2017. A Variety of Turkish in Mid-South Turkey: Arabic-Turkish Inter-Language. *Journal of Education and Training Studies* 5: 12–19. [CrossRef]
- Oyama, Susan. 1976. A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research* 5: 261–83. [CrossRef]
- Pfenninger, Simone E., and David Singleton. 2017. *Beyond Age Effects in Instructional L2 Learning: Revisiting the Age Factor*. Bristol: Multilingual Matters.
- Piske, Thorsten, Ian R. A. MacKay, and James E. Flege. 2001. Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics* 29: 191–215. [CrossRef]
- R Development Core Team. 2012. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, Available online: <http://www.R-project.org/> (accessed on 18 April 2017).
- Ratner, Nan Bernstein. 2000. Elicited imitation and other methods for the analysis of trade-offs between speech and language skills in children. In *Methods for Studying Language Production*. New York: Psychology Press, pp. 291–312.
- Rice, Keren. 2007. Markedness in phonology. In *The Cambridge Handbook of Phonology*. Edited by Paul de Lacy. Cambridge: Cambridge University Press, pp. 79–97.
- Schmidt, Richard W. 1990. The role of consciousness in second language learning. *Applied Linguistics* 11: 129–58. [CrossRef]
- Scovel, Thomas. 1988. *A Time to Speak: A Psycholinguistic Inquiry into the Critical Period for Human Speech*. Rowley: Newbury House Publishers.
- Seliger, Herbert W. 1975. Maturation constraints in the acquisition of second language accent. *Language Sciences* 36: 20–22.
- Skehan, Peter. 1998. *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
- Strange, Winifred, and Valerie L. Shafer. 2008. Speech perception in second language learners. In *Phonology and Second Language Acquisition*. Edited by Jette Hansen and Mary Zampini. Philadelphia: John Benjamins, pp. 153–91.
- Supriyadi, Tedi, and J. Julia. 2019. The problem of students in reading the Quran: A reflective-critical treatment through Action research. *International Journal of Instruction* 12: 311–26. [CrossRef]
- Tarone, Elaine, and Martha Bigelow. 2005. Impact of literacy on oral language processing: Implications for second language acquisition research. *Annual Review of Applied Linguistics* 25: 77–97. [CrossRef]
- Werker, Janet F., and John S. Logan. 1985. Cross-language evidence for three factors in speech perception. *Attention, Perception & Psychophysics* 37: 35–44.
- Young-Scholten, Martha, and Nancy Strom. 2006. First-time L2 readers: Is there a critical period? *LOT Occasional Series* 6: 45–68.
- Young-Scholten, Martha. 2013. Low-educated immigrants and the social relevance of second language acquisition research. *Second Language Research* 29: 441–54. [CrossRef]
- Zec, Draga. 2007. The syllable. In *The Cambridge Handbook of Phonology*. Edited by Paul de Lacy. Cambridge: Cambridge University Press, pp. 161–94.