International Journal of English Language and Literature Studies
ISSN(e): 2306-0646
ISSN(p): 2306-9910
DOI: 10.55493/5019.v14i3.5632
Vol. 14, No. 3, 390-408.

URL: www.aessweb.com

English consonant pronunciation accuracy among Arabic speakers: Impact of learning onset age and educational context

Aeshah Sultan Alkhaldi[1]
Hissah Mohammed Alruwaili[2+]

[1,2]*English Language Department, College of Arts, Jouf University, Sakaka, Saudi Arabia.*
[1]*Email: aesha1417@gmail.com*
[2]*Email: hissah@ju.edu.sa*

(+ Corresponding author)

Article History
Received: 10 July 2025
Revised: 9 September 2025
Accepted: 15 September 2025
Published: 25 September 2025

Keywords
Age of onset
Arabic-English
Educational context
Phonological acquisition
Critical period hypothesis
L2 phonology
Second language acquisition
Phonemic transfer.

ABSTRACT

Second language phonological acquisition among Arabic speakers remains underexplored despite Arabic's significance and English's status as the primary foreign language in Arabic-speaking countries. This study examines English consonant production accuracy among Arabic speakers from educational contexts: early starters began English at age 7 in private schools (n = 10), versus late starters who began English at age 14 in government schools (n = 10), matched for total exposure duration. Participants completed pronunciation tasks targeting six English consonants absent from Arabic: /p/, /v/, /ɹ/, /ŋ/, /ʧ/, and /ʒ/. The quasi-experimental design employed a three-stage assessment adapted from the Test of Spoken English, with audio recordings analyzed for error identification and statistical comparison. Results revealed substantial differences favoring the early starter/private school group (8.9% vs. 30.8% error rates, d = 3.31), with phoneme-specific analysis revealing distinct patterns: /p/, /ʧ/, and /ʒ/ showed dramatic advantages; /ɹ/ and /ŋ/ proved challenging for both groups despite differences; and /v/ presented minimal difficulties regardless of context. The systematic relationship between age of onset and educational context provides insights into how maturational and instructional factors interact in naturalistic learning environments. Differential phoneme responsiveness challenges traditional contrastive analysis predictions. This suggests pronunciation curricula should acknowledge variation in L2 phonological acquisition hierarchies, with implications for Arabic-English instruction policy and practice.

Contribution/ Originality: This work represents an original empirical investigation of patterns in Arabic-English phonological acquisition in relation to educational contexts and has not been previously published or submitted elsewhere. Our study contributes novel evidence regarding the interaction between the age of onset of learning and the educational context in the development of L2 pronunciation, thereby addressing a significant gap in the literature on Arabic-English acquisition.

1. INTRODUCTION

The critical period hypothesis (CPH) is one of the most influential yet contentious theories in second language (L2) acquisition. The CPH proposes that the ability to learn a language decreases significantly after a biologically determined critical period, typically coinciding with puberty (Johnson & Newport, 1991; Lenneberg, 1967), while substantial evidence supports age effects in phonological acquisition. Empirical studies testing the predictions of CPH in Arabic-English contexts remain surprisingly limited, despite Arabic being spoken by over 400 million people worldwide (Eberhard, Simons, & Fennig, 2023) and English serving as the primary foreign language in most Arabic-speaking countries (Kharma & Hajjaj, 1997).

In the present study, we present a preliminary investigation comparing the accuracy of the production of English consonants between Arabic speakers who began to learn English at different ages. Specifically, we examined the pronunciation patterns of learners who began English instruction at the age of seven versus age of 14, representing a seven-year difference that spans the theoretically proposed critical period boundary. This age difference, coupled with distinct educational contexts (private versus government schools), provided an opportunity to examine how maturational and instructional factors interacted in authentic learning environments. This design choice prioritized ecological validity by enabling the examination of age-educational context interactions as they occur naturally, rather than through artificial experiments in isolation.

The research design focuses on phonemes that are categorically absent from Arabic, thus ensuring clear contrastive analysis predictions while minimizing individual variations in first language (L1) phonological competence. Our investigation contributes to the growing body of research on Arabic-English phonological transfer patterns while providing preliminary evidence that is relevant to the broader theoretical debate surrounding the CPH. We addressed the following primary research question in this study:

To what extent do Arabic speakers who began learning English at the age of seven differ in their accuracy in the production of English consonants compared to those who began learning English at the age of 14?

In this exploratory investigation, we employed statistical analysis to examine pronunciation patterns in a sample of Arabic-English learners within their authentic educational contexts while acknowledging methodological considerations, including constraints on the sample size.

Our study provides initial evidence pertaining to age-related differences in the acquisition of Arabic-English phonology, thus offering preliminary insights that may inform future research directions.

The characteristics of the study's design necessitate a measured interpretation of the findings, with conclusions about broader theoretical implications and educational policy requiring empirical validation through complementary research approaches that include larger, more controlled samples.

2. LITERATURE REVIEW

2.1. The CPH in L2 Acquisition

The Critical Period Hypothesis (CPH), initially proposed by Lenneberg (1967) for L1 acquisition, has generated extensive debate in research on L2 acquisition. Lenneberg established that the ability to learn a language is significantly affected after puberty due to the completion of lateralization processes in the brain. This biological foundation was later extended to L2 learning, with researchers arguing that achieving native-like proficiency becomes increasingly challenging with a later onset of acquisition.

Johnson and Newport (1991) provided seminal evidence for the Critical Period Hypothesis in L2 contexts by demonstrating a clear correlation between the age of arrival and the ultimate attainment in English grammatical competence among Korean and Chinese immigrants. The authors' findings revealed a steady decline in the performances of learners who arrived after the age of seven, with particularly marked decreases after puberty. Similarly, Patkowski (1980) found that learners under the age of 15 achieved superior syntactic competency compared to those who were exposed to the target language after the age of 15, and suggested that there were maturational constraints on various linguistic domains.

The theoretical landscape surrounding age effects has shifted considerably in recent years. Singleton and Leśniewska (2024) challenged long-standing assumptions about critical periods and argued that the plasticity of adult brains for language learning may be far greater than was traditionally assumed. Their findings are particularly relevant to Arabic-English contexts, in which intensive intervention studies have shown that learners can achieve significant improvements in pronunciation well beyond puberty (Altamimi, 2015). This emerging evidence suggests that the biological determinism underlying traditional CPH models may be overstated.

However, the picture is more complex than simple age effects. Hartshorne, Tenenbaum, and Pinker (2018) dataset of 670,000 English learners revealed multiple sensitive periods rather than a single critical cut-off point. Their findings aligned with what we observed in Arabic-English acquisition: pronunciation difficulties varied dramatically across different sounds, with some phonemes responding well to late intervention and others remaining persistently challenging regardless of age. This variation challenges researchers to move beyond broad age comparisons towards more nuanced, phoneme-specific investigations.

The methodological concerns raised by scholars such as Vanhove (2013) and DeKeyser (2012) are particularly relevant in Arabic-speaking countries, where the age of onset for English is typically correlated with the educational context. For example, in Saudi Arabia, early English instruction mainly occurs in well-resourced private schools, while later instruction characterizes government institutions that have different teacher-training and pedagogical approaches. DeKeyser's emphasis on learning opportunities rather than biological constraints resonates strongly with this educational reality.

Despite these complexities, research on pronunciation continues to show the most consistent age effects across linguistic domains (Krashen, 1973). For Arabic speakers, this phonological vulnerability to age effects has particular significance given the substantial differences between Arabic and English consonantal systems. These systematic transfer challenges may interact with maturational constraints in ways that remain incompletely understood, making Arabic-English pronunciation an ideal testing ground for refined theories of age effects in L2 acquisition.

2.2. Critical Period Effects in Phonological Acquisition

Pronunciation has been found to be the most age-sensitive aspect of L2 learning. Snow and Hoefnagel-Höhle (1977) longitudinal study of English speakers in Holland revealed an interesting developmental pattern: older learners initially excelled in vocabulary and grammar, but younger participants gradually overtook them in pronunciation accuracy over the course of a year. This crossover effect suggests that phonological learning operates differently from other language skills early cognitive advantages eventually give way to what appears to be enhanced perceptual sensitivity in younger learners.

Auditory perception occupies a significant position in age-related L2 research. As Singleton and Leśniewska (2024) noted, "there are some grounds for believing that there may be some benefit to an early start in terms of listening skills, even though findings on this issue remain contradictory." This potential listening advantage is important because research has consistently shown "minimal empirical evidence supporting linguistic advantages relative to early second language (L2) instruction, with the exception of a possible benefit in terms of enhanced auditory skills, particularly speech perception, pronunciation, and listening comprehension" (2024, p. 564).

Ji (2021) study of Chinese learners in English-speaking environments provided a more nuanced perspective on age effects. While early starters achieved better pronunciation outcomes with less conscious effort, Ji emphasised that age functioned as a sensitive rather than a strictly critical factor. This distinction is important for Arabic-English acquisition, as the substantial phonological differences between the languages create systematic challenges that may interact with a learner's age in unpredictable ways.

The picture becomes more complex when we consider individual variations. Late learners sometimes achieve remarkable success in terms of pronunciation, particularly when they have strong motivation, extensive exposure, and access to quality instruction. These exceptional cases highlight how optimal learning conditions might compensate for a later age of onset, although such combinations of favourable factors remain relatively uncommon in typical language-learning contexts.

Case studies by Abu-Rabia and Kehat (2004) and by Bongaerts, van Summeren, Planken, and Schils (1997) documented individuals who achieved native-like accents despite starting to learn the target language after puberty. However, these exceptional cases involved highly motivated learners with extensive exposure and formal training in

pronunciation, which suggests that, while postcritical period success is possible, it requires significantly greater effort and optimal conditions.

2.3. Arabic-English Phonological Transfer

The phonological distance between Arabic and English creates systematic transfer effects that have been documented extensively. Arabic has a rich consonantal system with 28 consonants but lacks several English phonemes, thus creating predictable pronunciation difficulties for Arabic-speaking learners (Kharma & Hajjaj, 1997).

Numerous studies have identified certain English consonants that are consistently problematic for Arabic speakers. Ahmad (2011) found that Arabic learners experienced significant difficulties with /p/ (45% error rate), /v/ (52% error rate), /ʒ/ (95% error rate), /ʧ/ (35% error rate), and /ŋ/ (38% error rate).

Al-Badawi (2012) documented the predictable substitution patterns that Arabic speakers used when encountering unfamiliar English sounds in his analysis of 20 Saudi university students. His findings revealed substantial error rates, as 60% of the students substituted /b/ for /p/ in the initial position (45% in the medial position), while 40% replaced /v/ with /f/ initially and 60% medially. As Al-Badawi explained, "such misarticulation can be attributed to the absence of /v/ and /p/ in the Arabic phonological system, leading L2 learners to substitute Arabic equivalents for them that are imperfect in that they have a different manner of articulation" (p. 537).

In recent years, data-driven approaches have enabled a more precise examination and quantification of transfer patterns. Rehman, Silpachai, Levis, Zhao, and Gutierrez-Osuna (2020) applied corpus analysis methods to identify pronunciation errors by examining 599 phonetically balanced sentences read by four advanced Arabic speakers; the sentences were obtained from the L2 Arctic database. The authors used a systematic annotation approach to reveal specific error patterns: /dʒ/ had the highest substitution rate (40.1%), followed by /ɹ/ (35.1% distortion rate), and /p/ (27.8% substitution rate). Most importantly, their data-driven methodology revealed discrepancies with expert predictions while /v/ showed 100% error rates (always pronounced as /f/) in their sample, the bidirectional /p/-/b/ confusion suggested more complex transfer patterns than had traditionally been assumed. This empirical approach demonstrated how a corpus-based analysis could reveal pronunciation patterns that diverged from linguistically informed predictions, particularly with regard to the frequency and directionality of specific phonemic substitutions in advanced learners' speech. In addition to difficulties with individual phonemes, Arabic speakers encounter systematic challenges with English sound sequences that violate Arabic phonological constraints.

Pronunciation difficulties for Arabic speakers who are learning English extend beyond isolated phoneme substitutions. Hago and Khan (2015) empirical investigation of Saudi secondary school learners revealed that the production of consonant clusters posed particular challenges, with participants inserting epenthetic vowels in 71.65% of initial two-consonant clusters and in 57.45% of initial three-consonant clusters. The authors documented systematic errors across multiple consonant sounds most notably /p/, /ʒ/, and /ŋ/ with error rates exceeding 80% for certain phonemes in specific word positions. These findings point to deeper structural incompatibilities between Arabic and English phonotactic systems, as the learners' tendencies regarding vowel insertion reflected Arabic phonological constraints that prohibit the consonant sequences common in English.

The theoretical foundation for these transfer effects is the contrastive analysis hypothesis (CAH), which predicts that differences between L1 and L2 phonological systems will result in learning difficulties (Lado, 1957). Although the CAH has limitations in predicting all learning difficulties, it is particularly successful in predicting phonological transfer patterns, especially for sounds that are categorically absent from the learner's L1.

2.4. Pedagogical Interventions and the Educational Context

Recent intervention studies have demonstrated that targeted pedagogical approaches can significantly improve the pronunciation of Arabic speakers when speaking in English, even for phonemes that exhibit persistent L1 interference. Altamimi (2015) conducted action research with 13 intermediate Arabic students; the author

implemented training in minimal pairs over four weeks and targeted five problematic consonants. The intervention yielded substantial improvements: /p/ difficulties decreased from 76.92% to 7.69%, /ŋ/ difficulties from 92.30% to 23.07%, /ʒ/ difficulties from 84.61% to 30.76%, /v/ difficulties from 53.84% to 15.38%, and /ʃ/ difficulties from 23.07% to 0%. These findings demonstrated that an intensive focus on specific phonological contrasts could overcome predicted difficulties via explicit instruction.

However, recent research has highlighted the complexity of language exposure and its relationship with proficiency outcomes. Almashy (2025) conducted a qualitative study with 12 Saudi technical college students and revealed that, despite increased English exposure through English-medium instruction (EMI), passive exposure often led to limited communicative competence. The author identified sociopsychological barriers, including language anxiety and reduced intrinsic motivation, as key factors that decreased the potential benefits of increased exposure, thereby supporting Krashen, (1985) affective filter hypothesis.

Additional intervention studies have supported the effectiveness of targeted approaches. Mirza (2015) found that teaching the International Phonetic Alphabet (IPA) improved Saudi students' phonological awareness, leading to enhanced pronunciation skills, while Al-Jasser (2008) reported improvements in English word recognition following explicit instruction in English phonotactics.

An essential but understudied factor in research on age effects is the educational context. In many developing countries, including Arabic-speaking nations, early English instruction is often associated with attending private schools, while later instruction typically occurs in government schools (Al-Seghayer, 2014). These educational contexts differ systematically in terms of teacher training, class sizes, instructional materials, and pedagogical approaches, thus creating potential confounds in studies that compare early versus late learners.

Research on educational quality and L2 outcomes has suggested that the instructional context can moderate age effects significantly (Muñoz, 2006). High-quality instruction involving trained teachers, appropriate materials, and communicative approaches may compensate for a later age of onset of learning, while poor instructional contexts may fail to capitalize on the advantages of early learning.

2.5. Research Gap and Study Rationale

While the literature provides substantial evidence for age-related differences in the acquisition of L2 phonology, several important gaps remain. Firstly, most of the CPH research has focused on immigrants in naturalistic learning environments, with limited attention paid to instructed learning contexts in which the age of onset may be confounded by educational factors such as the type of school, teaching quality, and instructional methods.

Secondly, few studies have systematically examined how variables in the educational context interact with the effects of age in the acquisition of phonology. The distinction between private and government schools in terms of English instruction in many Arabic-speaking countries provides a natural context for investigating these interactions, although such educational differences necessarily complicate simple age-based interpretations.

Thirdly, while Arabic-English phonological transfer patterns have been well documented, limited research has examined how these patterns might vary as a function of the age of onset of learning, the educational context, or the interaction of these two factors. Understanding these relationships has both theoretical importance for refining our understanding of the effects of age and practical significance for the outcomes of English instruction in Arabic-speaking contexts.

A critical gap concerns the interaction between the age of acquisition and the strength of L1 interference. While numerous studies have documented CPH effects and Arabic-English transfer patterns separately, few have systematically examined whether early exposure can mitigate the effects of L1 interference. Theoretical predictions suggest that early learners should be less susceptible to L1 interference due to greater neural plasticity and less entrenched L1 phonological patterns; however, empirical evidence regarding this interaction remains limited, particularly in Arabic-English contexts.

In the present study, we address these gaps by comparing the accuracy of the pronunciation of problematic English consonants by Arabic learners who began English instruction at different ages and in different educational contexts. While this design does not permit definitive conclusions about critical period effects per se, it does provide valuable preliminary evidence regarding the interaction of age, educational context, and the outcomes of phonological acquisition.

3. METHODS

3.1. Research Design

We employed a quasi-experimental design to compare the accuracy of pronunciation between two groups of L1 Arabic learners of English as a foreign language (EFL) who had different ages of English onset and different educational backgrounds.

The between-subjects design was applied to examine naturally occurring variations in the timing of English instruction within educational systems, in which early English instruction (at the age of seven) typically occurs in private schools, while later instruction (at the age of 14) typically occurs in government schools.

This design capitalized on the systematic relationship between the age of onset and the educational context that characterizes naturalistic learning environments by examining how these factors interacted in real-world educational practices.

In contrast to studying by artificially isolating variables through experimental manipulation, our approach provided ecologically valid insights into the combined effects of maturational and instructional factors as they naturally occur in Arabic-speaking educational contexts.

3.2. Participants

Twenty university students from the northern region of Saudi Arabia participated in this study; they were strategically divided into two groups based on their age of onset of learning English:

1. Group 1 (Early starters): Ten participants (five males and five females) who began English instruction at the age of seven (first grade) at private schools.
2. Group 2 (Late starters): Ten participants (five males and five females) who began English instruction at the age of 14 (seventh grade) at government schools.

All the participants were aged 22-23 years at the time of testing and were enrolled in third- or fourth-year programs at the College of Business and Management at Jouf University. Business students were selected because they represented a homogeneous population in terms of educational backgrounds and current levels of English exposure, while also having had diverse early experiences of learning English that aligned with the comparative design of this study.

This age matching ensured an equivalent duration of exposure to English instruction (approximately eight to nine years) while controlling for the effects of cognitive maturity (Birdsong & Molis, 2001). The gender-balanced design enhanced external validity and enabled the examination of potential gender effects in age-related patterns of phonological acquisition (Asher & García, 1969).

3.2.1. Selection Criteria

Strict inclusion criteria were established to ensure group homogeneity and to control for potentially confounding variables following established protocols for research on L2 phonological acquisition (Flege, 1995; Johnson & Newport, 1991):

1. Geographical origin: The same geographical region was selected to control for dialectal variation.
2. Linguistic background: All the participants spoke Najdi Arabic as their native dialect to ensure a consistent L1 phonological baseline.

3. Monolingual background: Both of each participant's parents were of Arab origin, and Arabic was their family language.
4. Limited international exposure: None of the participants had resided outside of Arabic-speaking countries for more than one month.
5. Sensory integrity: None of the participants reported having hearing or speech difficulties.
6. Educational continuity: All the participants had received consistent English instruction within the same type of educational system.

3.2.2. Recruitment Procedure

The participants were recruited in September 2023 through a comprehensive survey distributed to approximately 311 students at the College of Business and Management.

After receiving 85 responses, participants were selected based on strict adherence to the inclusion criteria. The final sample consisted of 20 participants, representing all individuals who met the selection requirements within the recruitment timeframe.

3.3. Target Phonemes

Six English consonant phonemes were selected for analysis based on their systematic absence from the Arabic phonological inventory, as well as due to their documented status as being problematic for Arabic-speaking learners (Table 1).

Table 1. Target phonemes with phonetic descriptions.

Phoneme	Description	Example word	Arabic substitute
/p/	Voiceless bilabial plosive	'paper'	/b/
/v/	Voiced labiodental fricative	'voice'	/f/
/ɹ/	Alveolar approximant	'right'	/ɾ/
/ŋ/	Velar nasal	'thing'	/n/
/ʧ/	Voiceless postalveolar affricate	'chair'	/ʃ/
/ʒ/	Voiced postalveolar fricative	'measure'	/dʒ/

These phonemes were selected because they are categorically absent from the Najdi Arabic phonological inventory, thus providing clear predictions for a contrastive analysis while minimizing individual variations in L1 competence. The selection aligned with previous research that identified these specific sounds as being particularly problematic for Arabic-speaking learners (Ahmad, 2011; Al-Badawi, 2012; Hago & Khan, 2015). Recent corpus-based research by Rehman et al. (2020) provided empirical validation for this selection, as it demonstrated significant error rates for these phonemes amongst advanced Arabic speakers.

3.4. Data Collection Instrument

3.4.1. Adaptation of the Test of Spoken English (TSE)

The pronunciation assessment was based on the TSE, a standardized measure of speaking proficiency for non-native English speakers validated for university-level assessments (Powers & Stansfield, 1983). Three components (tasks II, III, and VI) from Powers and Stansfield's original seven-part TSE were selected and adapted to design a three-stage pronunciation assessment framework (Figure 1).

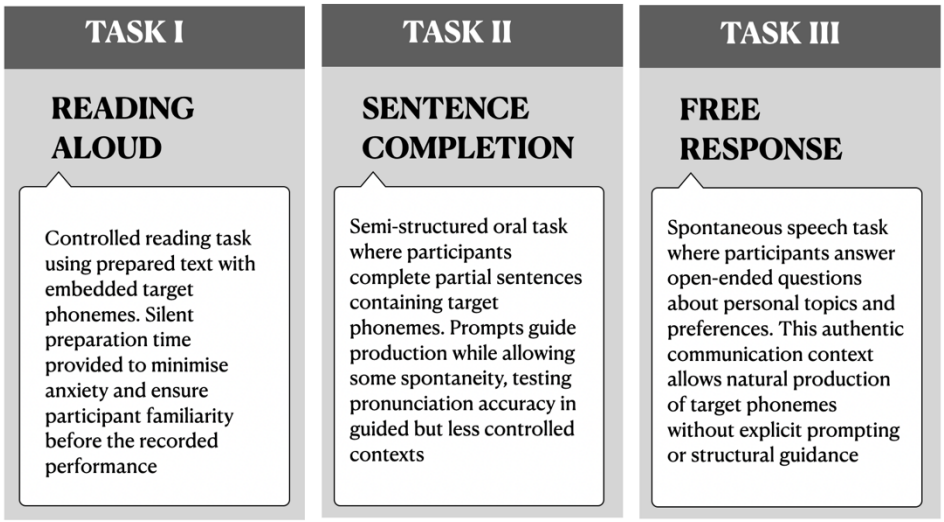


Figure 1. Three-stage Pronunciation assessment framework.

Source: Powers and Stansfield (1983).

We applied the principles of contrastive analysis to each task in our three-stage framework design for pronunciation assessment. Thus, each task covered each of the target phonemes (/p/, /v/, /ɹ/, /ŋ/, /ʧ/, and /ʒ/) across different phonological contexts (in isolation, in clusters, and in connected speech), and in word positions such as initial, middle, and final positions (Appendix A).

3.4.2. Pilot Testing and Validation

We conducted a small pilot study with two participants to check the testing procedures and to determine whether in-person or remote administration was a better option. As there were no significant differences in either the audio quality or the participants' performances between the two conditions, according to the results, we decided to use remote testing for all participants, as it was more convenient and eliminated any scheduling difficulties. This choice was also supported by previous research, which showed that the results of speech assessments did not differ between online and face-to-face testing (Waite, Theodoros, Russell, & Cahill, 2010).

3.5. Data Collection Procedures

We conducted all the testing sessions remotely via Zoom, with individual appointments lasting approximately 30–45 minutes each. After obtaining informed consent and confirming that the participants met our selection criteria, we verified the audio quality and initiated recordings with the participants' permission. The session then proceeded with the presentation of the three pronunciation tasks in sequence; each participant was allowed to complete each component before moving on to the next. We concluded each session with a brief debriefing period, during which the participants could ask questions about the research or clarify any aspects of their participation. This remote format proved to be practical for our participant population while maintaining the necessary audio quality for subsequent pronunciation analysis.

3.6. Data Analysis

3.6.1. Error Identification and Coding

The audio recordings were analyzed to identify pronunciation errors in the target phonemes. We counted the total instances of the target phonemes across all three tasks for each participant. The pronunciation errors were then tallied and converted to percentage error rates for each phoneme using the following formula:

$$Error\ rate = (number\ of\ mispronunciations \div total\ target\ phoneme\ instances) \times 100$$

3.6.2. Reliability Analysis

To ensure the reliability of the coding, a subset of recordings was coded independently by a second researcher who was familiar with patterns in Arabic-English pronunciation. The interrater agreement for identifying phoneme errors exceeded 95%, as calculated using Cohen's kappa for categorical judgments.

3.6.3. Statistical Analysis

A statistical analysis using SPSS 28.0 was conducted following the systematic framework outlined in Figure 2. The normality of the data was verified via Shapiro-Wilk tests and a Q-Q plot inspection before proceeding with the parametric analyses.

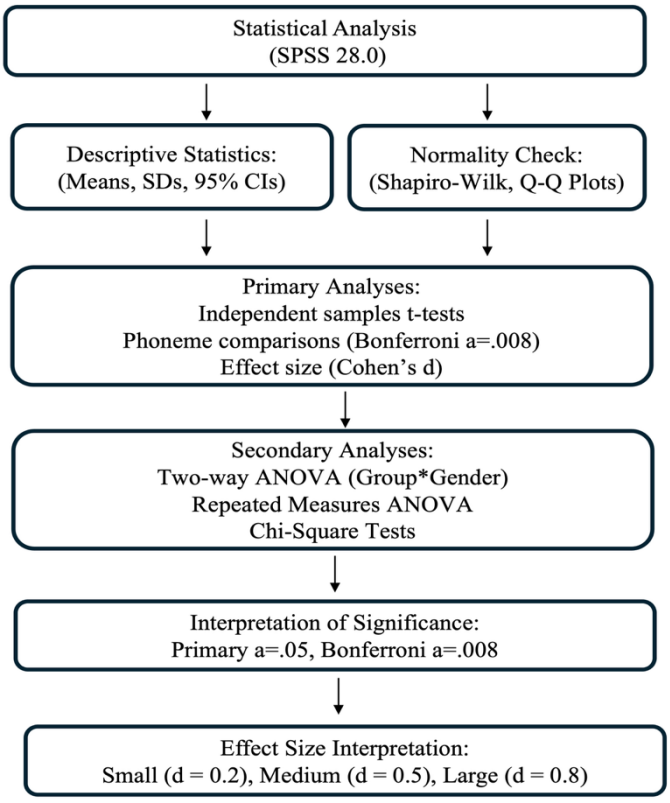


Figure 2. Statistical analysis framework showing the systematic analytical approach from the initial data processing to the final interpretation.
Note: The asterisk in "Group*Gender" denotes the interaction term between the factors Group and Gender in the two-way ANOVA.

In the primary analyses, we employed independent samples *t*-tests with the Bonferroni correction ($\alpha = .008$) for multiple phoneme comparisons, while the secondary analyses explored the interaction effects and categorical patterns. Effect sizes were calculated with 95% confidence intervals and were interpreted using Cohen's conventions.

3.7. Methodological Limitations and the Scope of the Study

3.7.1. Justification of the Sample Size

The sample size of 20 participants ($n = 10$ per group) was determined based on the practical constraints of the recruitment context and preliminary estimates of effect size. Based on pilot data suggesting large effect sizes for pronunciation differences between educational contexts, this sample provided adequate power (>80%) for detecting large group differences ($d > 1.2$) at $\alpha = .05$ using independent samples *t*-tests.

We acknowledge that this sample size is insufficient for several types of analyses required to draw definitive theoretical conclusions. Critical period breakpoint detection requires samples of 1000+ participants (Hartshorne et

al., 2018), while differentiating between age effects and the effects of educational contexts would require either experimental manipulation or much larger observational samples with sophisticated statistical controls. Our sample size is therefore appropriate for the study's actual scope, which was to provide preliminary evidence regarding pronunciation patterns in this specific educational context and to contribute to the descriptive literature on Arabic-English phonological transfer.

3.7.2. Design Limitations and Analytical Constraints

Following Vanhove (2013)'s methodological critique of CPH research, we acknowledge that our group-comparison design, while revealing substantial differences between early and late starters, cannot distinguish between the continuous age-related decline and the discontinuous function predicted by the strong form of the CPH. The dichotomization of the learners into early (age seven) versus late (age 14) starters, while theoretically motivated, overlooks information about continuous age-proficiency relationships and cannot test nuanced predictions about the form of age effects.

Future research should employ piecewise regression models to test for breakpoints in age-proficiency functions and should compare linear versus discontinuous models using likelihood ratio tests and information criteria. Our findings should be interpreted as evidence for substantial age-related differences in phonological acquisition that are consistent with, but do not definitively establish, critical period effects.

3.8. Ethical Considerations

This study received ethical approval from the relevant institutional review board. All participants provided informed consent and were assured of data confidentiality. The participants were informed that the recordings would be used solely for research purposes and that individual performances would not be disclosed outside of the research context.

4. RESULTS

4.1. Overall Performance Comparison

The pronunciation assessment revealed substantial differences between the two groups across the six target phonemes. Group 1 (early starters who attended private schools) demonstrated markedly superior pronunciation accuracy, with an overall error rate of 8.9%, while Group 2 (late starters who attended government schools) exhibited an error rate of 30.8%. This represents a 21.9 percentage point difference favoring the early starter/private school group.

The statistical analysis confirmed the significance of this difference: $t(18) = 7.42$, $p < .001$, $d = 3.31$, 95% CI [2.14, 4.48]. The very large effect size indicated a substantial and practically meaningful distinction between the groups. The observed effect size was notably larger compared to typical findings in research on L2 phonological acquisition, in which effect sizes generally range from small to medium ($d = 0.2$ to 0.8), probably reflecting the combined impact of age differences, variations in educational quality, and the specific phonemes selected for this study.

4.2. Phoneme-Specific Patterns

Individual phoneme comparisons revealed three distinct patterns of group differences, as presented in Table 2 below. Complete individual participant data for both groups are provided in Appendices B and C for a detailed examination and a potential reanalysis.

Table 2. Statistical comparison between groups by Phoneme.

Phoneme	Group 1 mean (SD)	Group 2 mean (*SD*)	Difference	*t*(18)	*p*-value	Cohen's *d*	95% CI
/p/	10.2%	28.4%	18.2%	6.84	<.001	3.05	[1.71, 4.39]
/v/	0.2%	1.2%	1.0%	1.26	.224	0.56	[-0.07, 1.19]
/ɹ/	25.5%	51.0%	25.5%	8.92	<.001	3.98	[2.51, 5.45]
/ŋ/	16.0%	45.4%	29.4%	7.15	<.001	3.19	[1.83, 4.55]
/ʧ/	1.1%	25.0%	23.9%	9.44	<.001	4.21	[2.69, 5.73]
/ʒ/	0.6%	31.4%	30.8%	11.76	<.001	5.24	[3.51, 6.97]
Overall	8.9%	30.8%	21.9%	7.42	<.001	3.31	[2.14, 4.48]

All the phoneme comparisons except for /v/ reached statistical significance after the Bonferroni correction (α = .008) was applied. The effect sizes ranged from very large ($d > 3.0$) for most of the comparisons to medium ($d = 0.56$) for /v/.

4.2.1. Highly Differentiated Phonemes

Three phonemes showed dramatic differences between the groups with very large effect sizes:

1. /p/ (voiceless bilabial plosive): An 18.2 percentage point difference (d = 3.05). Group 2 participants consistently substituted /b/ for /p/, reflecting the absence of /p/ in Arabic phonology.
2. /ʧ/ (voiceless postalveolar affricate): A 23.9 percentage point difference (d = 4.21). Group 2 showed frequent /ʃ/ substitutions, suggesting difficulties with affricate production.
3. /ʒ/ (voiced postalveolar fricative): A 30.8 percentage point difference (d = 5.24), representing the largest group difference observed. Group 2 participants predominantly substituted /dʒ/, which is consistent with Arabic phonological patterns.

4.2.2. Moderately Differentiated Phonemes

Two phonemes proved challenging for both groups despite significant group differences:

1. /ɹ/ (alveolar approximant): A 25.5 percentage point difference (d = 3.98). Even Group 1 showed substantial difficulties (25.5% error rate), with both groups frequently substituting the Arabic alveolar tap /ɾ/.
2. /ŋ/ (velar nasal): A 29.4 percentage point difference (d = 3.19). Group 1 maintained a 16.0% error rate, suggesting inherent acquisition difficulties. Both groups frequently substituted /n/, reflecting Arabic phonotactic constraints.

4.2.3. Minimally Differentiated Phonemes

/v/ (voiced labiodental fricative): There was only a 1.0 percentage point difference (d = 0.56, p = .224). Contrary to predictions based on a contrastive analysis and previous research findings, /v/ presented minimal challenges for both groups, with Group 1 achieving near-perfect accuracy (a 0.2% error rate) and Group 2 showing minimal difficulties (a 1.2% error rate).

4.3. Analysis of Error Patterns

Systematic phonemic substitution patterns emerged from the qualitative analysis, reflecting predictable L1 interference that was consistent with Arabic phonological constraints.

1. /ɹ/ → /ɾ/: The Arabic alveolar tap was consistently substituted for the English approximant.
2. /ŋ/ → /n/: The Arabic alveolar nasal replaced the English velar nasal.
3. /p/ → /b/: The voiced equivalent was substituted due to the absence of /p/ in Arabic.
4. /ʧ/ → /ʃ/: A fricative was substituted for an affricate.
5. /ʒ/ → /dʒ/: An Arabic affricate was substituted for an English fricative.

These substitution patterns were observed consistently across the participants within each group, with Group 2 participants showing more frequent and persistent use of L1-based substitutions across all the contexts that were tested.

4.4. Individual Variation Patterns

4.4.1. Group 1 (Early Starters)

The individual participant analysis revealed remarkably consistent patterns within Group 1, with most of the participants experiencing primary difficulties with /ɹ/ and /ŋ/ while achieving near-perfect accuracy for /p/, /ʧ/, and /ʒ/. The individual variation within the group was relatively small, with the highest individual /p/ error rate (18%) still being substantially lower than it was for any participant in Group 2.

4.4.2. Group 2 (Late Starters)

All the Group 2 participants demonstrated substantial difficulties across multiple phonemes, with consistently elevated error rates for /p/ (range: 25-32%), /ɹ/ (range: 48-55%), /ŋ/ (range: 42-49%), /ʧ/ (range: 22-28%), and /ʒ/ (range: 28-35%). The group showed remarkable homogeneity in error patterns, with relatively small standard deviations across phonemes, suggesting systematic effects rather than individual variations. The complete individual error rates for all the participants are provided in Appendices B and C for detailed examination and potential reanalysis. Appendix B presents individual pronunciation error rates for each phoneme across all ten participants in Group 1 (early starters), while Appendix C displays the corresponding data for Group 2 (late starters).

4.5. Secondary Analysis: Gender and Phoneme Interactions

A two-way analysis of variance (ANOVA) examining Group × Gender interactions for overall error rates revealed no significant main effect of gender ($F(1,16) = 0.73$, $p = .405$) or interaction ($F(1,16) = 1.24$, $p = .282$), indicating that the observed group differences were consistent across male and female participants.

A repeated measures ANOVA with Phoneme as the within-subjects factor and Group as a between-subjects factor confirmed a significant main effect of Group ($F(1,18) = 55.03$, $p < .001$, $\eta p^2 = .754$) and a significant Group × Phoneme interaction ($F(5,90) = 12.47$, $p < .001$, $\eta p^2 = .409$), thus supporting the phoneme-specific patterns identified in the individual analyses.

5. DISCUSSION

5.1. Educational Context and Phonological Acquisition: A Complex Interaction

The substantial performance advantage observed for early starter/private school participants (21.9 percentage points, $d = 3.31$) represents one of the most pronounced group differences documented in research on Arabic-English pronunciation. While previous studies have typically reported small to medium effect sizes for age-related differences (Johnson and Newport (1991) $d = 1.2$), our findings reveal effects that are dramatically larger, suggesting that the combination of the age of onset of learning and the educational context may create synergistic rather than simply additive benefits for phonological acquisition.

This finding challenges simplistic interpretations of either pure age effects or pure educational effects. Instead, our data suggest that optimal learning conditions characterized by early exposure, smaller classes, qualified teachers, and communicative pedagogical approaches may become increasingly important as the theoretical critical period window narrows. This interpretation aligns with Muñoz (2006) framework, which proposed that high-quality instruction could moderate age effects, but extends this concept by demonstrating that such moderation may be most pronounced when multiple favorable factors converge.

The magnitude of our observed effects also sheds light on why previous studies that examined isolated variables often produced inconsistent results. The confounded nature of our design, rather than simply representing a

limitation, may provide insights into how multiple factors interact in real-world educational contexts. DeKeyser (2012) critique of studies that failed to control for educational variables becomes particularly relevant here our findings suggest that attempting to separate age effects and the educational context may obscure rather than clarify the mechanisms underlying L2 phonological development.

5.2. Phoneme-Specific Responsiveness to Learning Conditions

Perhaps the most theoretically intriguing aspect of our findings was the discrepancy in the responsiveness of the phonemes to the combined advantages of early onset and enhanced educational context. The emergence of three distinct patterns highly differentiated, moderately differentiated, and minimally differentiated phonemes suggests that phonological contrasts vary systematically in their susceptibility to the effects of learning conditions.

5.2.1. Highly Responsive Phonemes: The Role of Explicit Instruction

The dramatic group differences observed for /p/, /ʧ/, and /ʒ/ (effect sizes 3.05-5.24) aligned remarkably well with Altamimi (2015) intervention findings, in which these same phonemes showed the most substantial improvements following explicit training in minimal pairs. This convergence suggests that certain phonological contrasts may be particularly amenable to focused pedagogical attention, which is precisely the type of instruction more likely to occur in well-resourced educational contexts.

Group 1's near-perfect performance for these phonemes (error rates <2%) demonstrates that Arabic speakers can achieve native-like accuracy for sounds that are absent in their L1, thus contradicting deficit-based perspectives on L1 interference. Instead, our findings support the predictions of Flege (1995) Speech Learning Model, which posited that challenging L2 sounds could be mastered when the learning conditions optimized both the perceptual input quality and the opportunities for articulatory feedback.

5.2.2. Persistent Challenges: When Optimal Conditions Meet Inherent Difficulty

Both groups struggled significantly with /ɹ/ and /ŋ/, with Group 1 showing error rates above 15% for these sounds. This suggests that some phonemes are inherently difficult to acquire, regardless of educational advantages. These findings align with Rehman et al. (2020) corpus analysis, which showed /ɹ/ distortion rates of 35.1%, even among advanced Arabic speakers. This indicates that certain sounds remain challenging for Arabic learners regardless of their proficiency levels or learning conditions.

In the case of /ŋ/, its persistent difficulties may reflect deeper phonotactic constraints rather than simple segmental absence. While Arabic lacks /ŋ/ as a phoneme, Arabic speakers systematically insert consonants after velar nasals in coda. This indicates difficulty with English patterns that allow /ŋ/ at the end of words and shows how L1 distributional constraints extend beyond basic segmental differences. We argue that this extends predictions of traditional contrastive analyses (Lado, 1957) to include phonotactic factors in L2 acquisition, thus challenging assumptions that segmental differences alone predict pronunciation difficulties.

5.2.3. Unexpected /v/ Patterns

The absence of meaningful group differences for /v/ directly contradicted both the predictions of traditional contrastive analyses and previous empirical findings (Ahmad (2011); 52% error rate). This unexpected result requires theoretical reconsideration and highlights the limitations of relying solely on comparisons of phonemic inventories for predicting L2 difficulties.

Several explanations merit consideration. Firstly, the increasing global exposure to English via digital media may have provided Arabic speakers with substantial incidental exposure to /v/ prior to receiving formal instruction, potentially neutralizing the advantages of the educational context. Secondly, the articulatory similarity between /v/ and /f/ may facilitate rapid acquisition once learners become phonetically aware of the distinction. Thirdly, modern

teaching materials and methods may have evolved to address /v/ production more effectively than has historically been documented.

This finding emphasizes the dynamic nature of L2 phonological challenges and suggests that empirical validation should be updated regularly to reflect theoretical predictions, particularly in our rapidly changing linguistic landscape.

5.3. Theoretical Implications for Critical Period Research

Our findings contribute to the evolving theoretical discourse surrounding critical period effects while highlighting the complexity of differentiating between maturational and experiential factors. The patterns that we observed are consistent with multiple theoretical frameworks, each suggesting different mechanistic explanations for the same empirical phenomena.

From a strong CPH perspective (Lenneberg, 1967) the substantial group differences support predictions that early L2 exposure capitalizes on heightened neural plasticity before lateralization is complete. The different responsiveness of the phonemes could reflect varying degrees of neural entrenchment, with some sounds becoming more resistant to modification as the critical period closes.

However, our findings are equally compatible with usage-based theories emphasising the cumulative effects of exposure (Hartshorne et al., 2018). The seven-year exposure difference between the groups, combined with qualitatively different instructional approaches, could account for the observed differences in terms of frequency and the quality of linguistic input, rather than maturational constraints per se.

Most intriguingly, our data suggest that these theoretical perspectives may be complementary rather than competing: early exposure may create neurobiological advantages that are only actualized when combined with optimal learning conditions. This interpretation aligns with Singleton and Leśniewska (2024) emphasis on perceptual foundations in the acquisition of phonology; that is, early exposure may establish superior perceptual categories that subsequently benefit from high-quality instructional support.

5.4. Implications for Educational Policy: Evidence and Caution

While our confounded design prevents definitive policy recommendations, the substantial group differences that we observed warrant the serious consideration of educational planners in Arabic-speaking contexts. The data suggest that decisions about the timing of English instruction cannot be divorced from considerations regarding instructional quality and resource allocation.

The superior outcomes achieved by early starter/private school participants indicate that optimal phonological acquisition outcomes are achievable for Arabic speakers when favorable conditions converge. However, the crucial question for policymakers concerns cost-effectiveness: Would investing resources in improving the quality of instruction at government schools at later ages yield comparable benefits to earlier instruction in better-resourced contexts?

Our findings regarding phoneme-specific patterns offer more immediate practical guidance. The persistent challenges associated with /ɹ/ and /ŋ/ across both groups suggest that these sounds should receive sustained pedagogical attention regardless of the learning context, while the responsiveness of /p/, /ʧ/, and /ʒ/ to optimal conditions indicates that focused intervention can yield dramatic improvements for certain contrasts.

The findings for /v/ provide a reason for optimism, as they suggest that some traditionally problematic sounds may become more tractable as global English exposure increases and teaching methods evolve. This highlights the importance of updating pedagogical approaches regularly based on current empirical evidence rather than on historical assumptions.

5.5. Methodological Contributions and Future Research Imperatives

Our study contributes methodologically by demonstrating both the opportunities and the constraints inherent in quasi-experimental designs that examine naturally occurring educational variations. While the confounded nature of our design limits causal attribution, it provides ecologically valid insights into how multiple factors interact in real-world contexts insights that may be obscured by experimental designs that artificially isolate variables.

The phoneme-specific patterns that we identified suggest several promising research directions. Future investigations should employ longitudinal designs that track learners across different educational contexts to establish their developmental trajectories. Cross-linguistic research examining whether similar phoneme hierarchies emerge in other L1-English combinations would test the generalizability of our findings beyond Arabic-English contexts. Most importantly, future research must develop innovative designs that can differentiate between the effects of age and the educational context. Natural experiments arising from educational policy changes, within-school-type age comparisons, and sophisticated statistical modeling of educational quality factors represent promising approaches for advancing theoretical understanding while maintaining ecological validity. The persistent challenges observed for certain phonemes also warrant focused investigation of instructional interventions. Building on Altamimi (2015)'s success with minimal pairs, future research should systematically evaluate pedagogical approaches for the most resistant contrasts, potentially informing evidence-based pronunciation curricula.

5.6. Study Scope and Interpretive Constraints

Our findings provide valuable preliminary evidence regarding the effects of the educational context on Arabic-English phonological acquisition while acknowledging significant interpretive constraints. The small sample size (N = 20), the single-site design, and the specific demographic focus limit the generalizability of the findings across Arabic-speaking populations and educational systems. In addition, the confounded design prevents the attribution of the observed differences to specific causal mechanisms. These limitations position our study as contributing to descriptive knowledge and the generation of hypotheses rather than providing definitive theoretical conclusions. The substantial effect sizes that we observed, while impressive, require replication across diverse contexts before broader generalizations about age-educational context interactions can be supported. Our study's scope aligned with our exploratory aims of documenting pronunciation patterns in a specific educational context, quantifying effect sizes for future power analyses, and contributing to the growing body of evidence pertaining to Arabic-English phonological acquisition.

6. CONCLUSIONS

Our comparison of Arabic speakers from different educational backgrounds revealed substantial pronunciation differences that challenge conventional thinking about the effects of age and learning contexts. The 21.9 percentage point advantage for early starters represents one of the largest effect sizes (d = 3.31) documented in research on pronunciation, suggesting that age and educational quality may work together in ways that are not yet fully understood. The key finding concerns how different sounds respond to optimal learning conditions. Arabic speakers achieved near-perfect accuracy for /p/, /ʧ/, and /ʒ/ when they started learning English at an early age and in well-resourced schools, contradicting the common assumption that sounds absent in the L1 will remain permanently problematic. However, both groups struggled with /ɹ/ and /ŋ/, indicating that some pronunciation challenges will persist regardless of the learning context. Perhaps most surprising was the ease with which both groups mastered /v/, a finding that contradicts decades of research on Arabic-English transfer patterns and suggests that global English exposure may be reshaping traditional pronunciation hierarchies. Our research design, while preventing simple causal claims, offers something equally valuable: insights into how multiple factors interact in real educational settings. The confounding of age and school type in Saudi Arabia is not just a methodological limitation it reflects how language learning actually takes place in most multilingual contexts. Attempting to separate these factors

artificially in laboratory studies might overlook the very interactions that are most important for educational outcomes. These patterns suggest that researchers may need to rethink the relationship between critical period effects and educational factors. Rather than viewing them as competing explanations, our data indicate that they may be fundamentally interconnected. Early exposure might create biological advantages that only emerge under optimal instructional conditions, while high-quality teaching might be most effective during sensitive developmental periods.

For educators and policymakers in Arabic-speaking countries, these findings highlight both challenges and opportunities. While we cannot make direct policy recommendations based on our small-scale study, the results suggest that decisions about the timing of English instruction cannot be divorced from questions regarding instructional quality. The dramatic improvements that were possible for certain sounds indicate that focused intervention could yield impressive results, while the persistent difficulties with others indicate where sustained effort is most needed. Most importantly, the /v/ findings remind us that pronunciation challenges are not fixed they evolve as learners' exposure to English changes through digital media and global connectivity.

Looking forward, our research has established a foundation for more sophisticated investigations that can differentiate between the effects of age and educational contexts through innovative designs that maintain ecological validity. The substantial effect sizes that we documented provide essential benchmarks for future studies, while the phoneme-specific patterns suggest productive directions for both theoretical developments and for research on pedagogical interventions. Ultimately, our study contributes to an emerging understanding that L2 phonological acquisition is the result of dynamic interactions among multiple factors rather than single-cause mechanisms. By embracing this complexity rather than attempting to eliminate it, future research may develop more comprehensive theories that enhance both scientific understanding and educational practice in our increasingly multilingual world.

Funding: This study received no specific financial support.
Institutional Review Board Statement: This study was approved by the Jouf University, Saudi Arabia under protocol number (IRB No.: JU REC/2/45/2314), dated (Approval date March 17, 2023). Informed consent was obtained from all participants, and all data were anonymised to protect participant confidentiality.
Transparency: The authors state that the manuscript is honest, truthful, and transparent, that no key aspects of the investigation have been omitted, and that any differences from the study as planned have been clarified. This study followed all writing ethics.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: Both authors contributed equally to the conception and design of the study. Both authors have read and agreed to the published version of the manuscript.
Disclosure of AI Use: The author used OpenAI's ChatGPT (GPT-4) to edit and refine the wording of the Introduction and Literature Review. All outputs were thoroughly reviewed and verified by the author.

REFERENCES

Abu-Rabia, S., & Kehat, S. (2004). The critical period for second language pronunciation: Is there such a thing? *Educational Psychology*, *24*(1), 77-97. https://doi.org/10.1080/0144341032000146467

Ahmad, J. (2011). Pronunciation problems among Saudi learners: A case study at the Preparatory Year Program, Najran University, Saudi Arabia. *Language in India*, *11*(7), 274–285.

Al-Badawi, K. (2012). An analysis of phonetic, morphological and syntactic errors in English: A case study of Saudi BA students at King Khalid University. *International Journal of Social Science and Humanity*, *2*(6), 536-538. https://doi.org/10.7763/IJSSH.2012.V2.165

Al-Jasser, F. (2008). The effect of teaching English phonotactics on the lexical segmentation of English as a foreign language. *System*, *36*(1), 94-106. https://doi.org/10.1016/j.system.2007.12.002

Al-Seghayer, K. (2014). The four most common constraints affecting English teaching in Saudi Arabia. *International Journal of English Linguistics*, *4*(5), 17-26. https://doi.org/10.5539/ijel.v4n5p17

Almashy, A. (2025). Language exposure and language proficiency: A qualitative study of technical colleges in Saudi Arabia. *Revista Romaneasca Pentru Educatie Multidimensionala*, *17*(2), 38-65. https://doi.org/10.18662/rrem/17.2/973

Altamimi, A. K. (2015). Improving English pronunciation among Arabic EFL school-Age student using minimal pairs. Unpublished M.A. Thesis: State University of New York.

Asher, J. J., & García, R. (1969). The optimal age to learn a foreign language. *The Modern Language Journal*, *53*(5), 334-341. https://doi.org/10.2307/323026

Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language*, *44*(2), 235-249. https://doi.org/10.1006/jmla.2000.2750

Bongaerts, T., van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, *19*(4), 447-465.

DeKeyser, R. (2012). Interactions between individual differences, treatments, and structures in SLA. *Language Learning*, *62*(s2), 189-200. https://doi.org/10.1111/j.1467-9922.2012.00712.x

Eberhard, D. M., Simons, G. F., & Fennig, C. D. (2023). *Ethnologue: Languages of the world* (26th ed.). Dallas, TX: SIL International.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), Speech perception and linguistic experience: Issues in cross-language research. In (pp. 233–277). Timonium, MD: New York Press.

Hago, O. E., & Khan, W. A. (2015). The pronunciation problems faced by Saudi EFL learners at secondary schools. *Education and Linguistics Research*, *1*(2), 85-99. https://doi.org/10.5296/elr.v1i2.7783

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

Ji, F. (2021). Investigating the influences of starting age on pronunciation: A comparative study of Chinese learners of English as a second language. *Journal of Higher Education Research*, *2*(3), 108-123. https://doi.org/10.32629/jher.v2i3.324

Johnson, J. S., & Newport, E. L. (1991). Critical period effects on universal properties of language: The status of subjacency in the acquisition of a second language. *Cognition*, *39*(3), 215-258. https://doi.org/10.1016/0010-0277(91)90054-8

Kharma, N., & Hajjaj, A. (1997). *Errors in English among Arabic speakers: Analysis and remedy*. Beirut: Librairie Du Liban.

Krashen, S. D. (1973). Lateralization, language learning, and the critical period: Some new evidence. *Language Learning*, *23*(1), 63-74. https://doi.org/10.1111/j.1467-1770.1973.tb00097.x

Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. New York: Longman.

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor, MI: University of Michigan Press.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: John Wiley & Sons.

Mirza, H. S. (2015). ESL and EFL learners improve differently in pronunciation: The case of Lebanon. *Procedia-Social and Behavioral Sciences*, *199*, 486-495. https://doi.org/10.1016/j.sbspro.2015.07.536

Muñoz, C. (2006). *Age and the rate of foreign language learning*. Clevedon, England: Multilingual Matters.

Patkowski, M. S. (1980). The sensitive period for the acquisition of syntax in a second language 1. *Language Learning*, *30*(2), 449-468. https://doi.org/10.1111/j.1467-1770.1980.tb00328.x

Powers, D. E., & Stansfield, C. W. (1983). *The test of spoken English as a measure of communicative ability in the health professions: Validation and standard setting*. ETS Research Report Series, RR-83-27. Educational Testing Service.

Rehman, I., Silpachai, A., Levis, J., Zhao, G., & Gutierrez-Osuna, R. (2020). The English pronunciation of Arabic speakers: A data-driven approach to segmental error identification. *Language Teaching Research*, *26*(6), 1055-1081. https://doi.org/10.1177/1362168820931888

Singleton, D., & Leśniewska, J. (2024). The role of age in second language development. *Language Teaching Research Quarterly*, *39*, 359-371. https://doi.org/10.32038/ltrq.2024.39.22

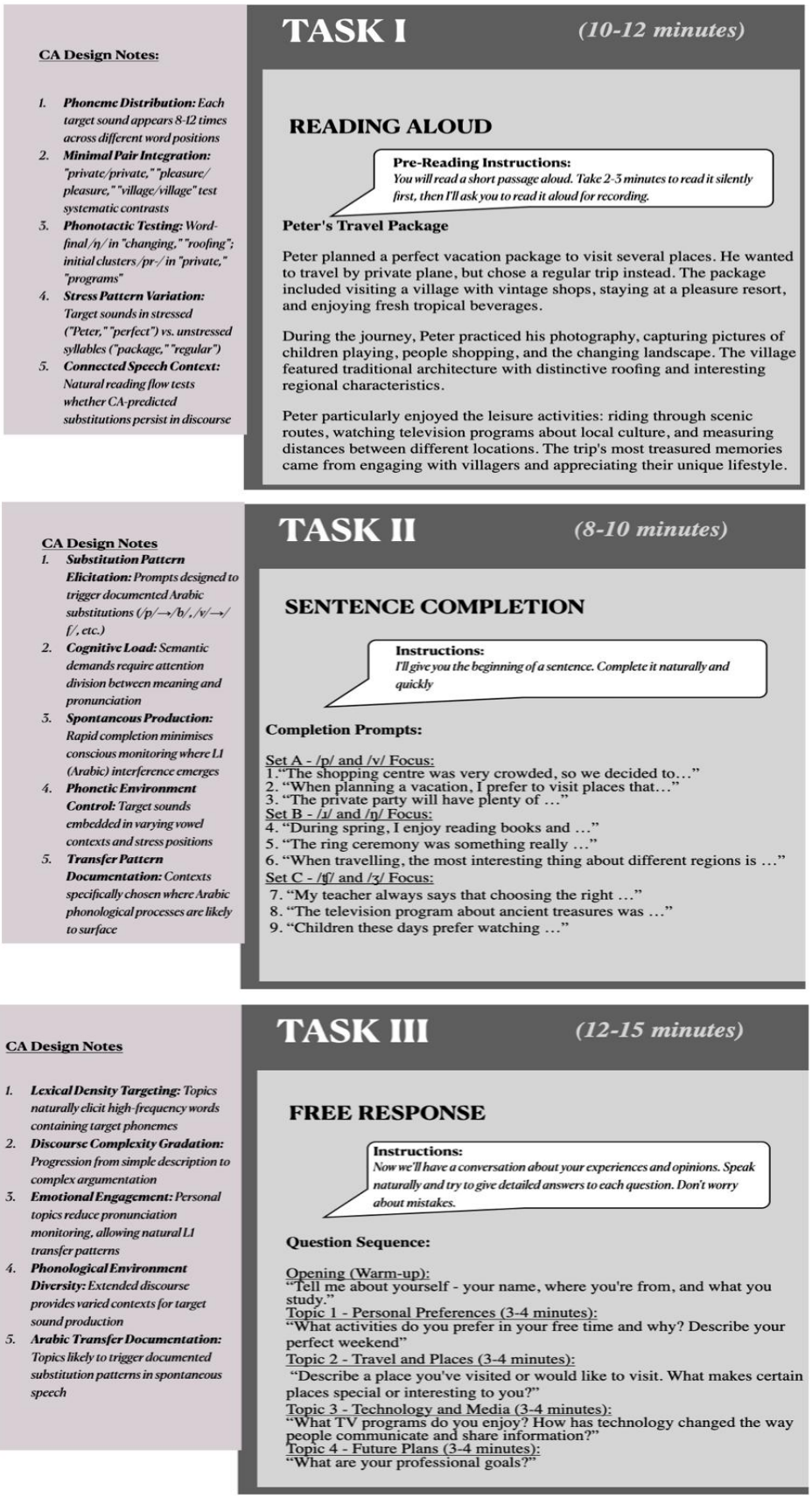
Snow, C. E., & Hoefnagel-Höhle, M. (1977). Age differences in the pronunciation of foreign sounds. *Language and Speech*, *20*(4), 357-365. https://doi.org/10.1177/002383097702000407

Vanhove, J. (2013). The critical period hypothesis in second language acquisition: A statistical critique and a reanalysis. *PloS One*, *8*(7), e69172. https://doi.org/10.1371/journal.pone.0069172

Waite, M. C., Theodoros, D. G., Russell, T. G., & Cahill, L. M. (2010). Internet-based telehealth assessment of language using the CELF–4. *Language, Speech, and Hearing Services in Schools*, *41*(4), 445-458.

Appendices

Appendix A.



Appendix B.

Table. Pronunciation Error Rates - Group 1 (Early Starters).

Phoneme	P1 (F)	P2 (F)	P3 (F)	P4 (F)	P5 (F)	P6 (M)	P7 (M)	P8 (M)	P9 (M)	P10 (M)	Mean (SD)
/p/	8%	9%	6%	7%	10%	11%	13%	9%	18%	11%	10.2% (3.6)
/v/	0%	0%	1%	0%	0%	0%	1%	0%	0%	1%	0.2% (0.4)
/ɹ/	29%	26%	23%	24%	28%	22%	25%	25%	29%	24%	25.5% (2.4)
/ŋ/	15%	19%	19%	18%	20%	11%	15%	14%	12%	17%	16.0% (3.2)
/ʧ/	1%	0%	2%	2%	1%	0%	2%	0%	3%	0%	1.1% (1.1)
/ʒ/	1%	0%	0%	0%	1%	0%	3%	0%	0%	1%	0.6% (1.0)
Overall Average											8.9% (1.8)

Appendix C.

Table. Pronunciation Error Rates - Group 2 (Late Starters).

Phoneme	P11 (F)	P12 (F)	P13 (F)	P14 (F)	P15 (F)	P16 (M)	P17 (M)	P18 (M)	P19 (M)	P20 (M)	Mean (SD)
/p/	28%	32%	25%	30%	27%	26%	31%	29%	28%	28%	28.4% (2.7)
/v/	2%	1%	0%	2%	1%	1%	2%	0%	1%	2%	1.2% (0.8)
/ɹ/	52%	48%	55%	49%	51%	50%	53%	49%	52%	51%	51.0% (2.8)
/ŋ/	44%	49%	42%	47%	45%	43%	48%	44%	46%	47%	45.4% (2.9)
/ʧ/	25%	28%	22%	26%	24%	23%	27%	25%	26%	24%	25.0% (2.4)
/ʒ/	31%	35%	28%	33%	30%	29%	34%	31%	32%	33%	31.4% (2.7)
Overall Average											30.8% (1.9)