Universitas Muhammadiyah Malang, East Java, Indonesia
Izdihar : Journal of Arabic Language Teaching, Linguistics, and Literature
p-ISSN: 2622-738X, e-ISSN: 2622-7371 // Vol.9 No.2 Agustus 2026, pp. 149-178
https://doi.org/10.22219/jiz.v8i2.45039 http://ejournal.umm.ac.id/index.php/izdihar/index izdihar.jurnalpba@umm.ac.id

Audio-Laboratory-Based Evaluative Orthography as a Diagnostic Approach for Errors in Writing Arabic Texts

Ibnu Rawandhy N. Hula [a,1], Nur Hasaniyah [b,2], Wala Fahad bin Subait [c,3], Hairuddin [a,4], Ana Mariana [d,5]

[a]IAIN Sultan Amai Gorontalo, Indonesia
[b]UIN Maulana Malik Ibrahim Malang, Indonesia
[c]Princess Nourah bint Abdulrahman University Riyadh, Saudi Arabia
[d]Universitas Muhammadiyah Gorontalo, Indonesia

[1]ibnurawandi@iaingorontalo.ac.id*, [2]hasaniyah@bsa.uin-malang.ac.id, [3]wfalsobait@pnu.edu.sa, [4]hairuddin@iaingorontalo.ac.id, [5]anamariana@umgo.ac.id

ARTICLE INFO

Article History:
Received: 28/05/2026
Revised: 14/07/2026
Accepted:30/08/2026
Published:31/08/2026

***Corresponding Author:**
Name: Ibnu Rawandhy N. Hula
Email: ibnurawandi@iaingorontalo.ac.id

ABSTRACT

This study aimed to identify orthographic errors based on the *Qawāʿid Khamsah*, determine the rules most frequently associated with errors, and develop the OEA-Lab Model (Audio-Laboratory-Based Evaluative Orthography Model). A pre-experimental, one-group pretest–posttest design was employed with 25 Arabic Language Education students, who completed 337 lexical items presented through 29 audio files covering Hamzah, Alif, Ḥażf, *Ziyādah*, and *Waṣl-Faṣl*. The mean score increased from 83.28 to 87.03 on the posttest. Because the difference scores deviated significantly from normality (Shapiro–Wilk $W = 0.904$, $p = .023$), the Wilcoxon signed-rank test was used to examine the difference between pretest and posttest scores, and indicated a statistically significant improvement, $Z = -4.382$, $p < 0.001$, with a large effect size, $r = 0.619$, and all 25 participants demonstrated improvement. Internal consistency of the 337 dichotomously scored items was acceptable, with KR-20 = 0.853 (pretest) and 0.841 (posttest). The highest pretest error rate occurred in *Ziyādah* (Hāʾ Saktah), at 31.45%, followed by Hamzah, at 14.14%. However, in terms of the proportion of all errors, *Hamzah* accounted for the largest share (35.13%). The findings suggest that diagnostically structured, audio-based dictation, rather than text copying, can produce measurable improvements in orthographic accuracy, positioning imlāʾ as a diagnostic-remedial instrument.

Keyword
Arabic orthography; Arabic writing errors; audio-laboratory; diagnostic assessment; evaluative dictation.

مستخلص البحث

يُقدِّم هذا البحث الإملاءَ الاختباريَّ القائمَ على المختبر الصوتي حلًّا تشخيصيًّا لأخطاء الطلبة في كتابة النصوص العربية. وتهدف إلى تحديد الأخطاء الإملائية في ضوء القواعد الخمس، وتحديد أكثر القواعد تسببًا فيها، وصياغة نموذج OEA-Lab (نموذج الإملاء التقويمي القائم على المختبر الصوتي). واعتمدت الدراسة تصميمًا ما قبل تجريبي من نوع المجموعة الواحدة ذات الاختبارين القبلي والبعدي، بمشاركة ٢٥ طالبًا من طلبة تعليم اللغة العربية، أنجزوا ٣٣٧ مفردةً معجميةً عبر ٢٩ ملفًّا صوتيًّا شملت الهمزة، والألف، والحذف، والزيادة، والوصل والفصل. وارتفع متوسط الدرجات من ٨٣,٢٨ إلى ٨٧,٠٣. ولعدم اعتدالية فروق الدرجات وفق اختبار Shapiro-Wilk، $W = 0.904$، $p = 0.023$، جرى تقويم الفرق بين الاختبارين باختبار Wilcoxon للرتب الموقعة، فأظهر تحسنًا دالًّا إحصائيًّا، $Z = 4.382$، $p < 0.001$، بحجم أثر كبير، $r = 0.619$، مع تحسن جميع الطلبة الخمسة والعشرين دون استثناء. وبلغ معامل الثبات KR-20 للبنود الثنائية البالغ عددها ٣٣٧ بندًا ٠,٨٥٣ في الاختبار القبلي و٠,٨٤١ في الاختبار البعدي. وتركزت الأخطاء في الزيادة، ولا سيما هاء السكت، بنسبة ٣١,٤٥٪ في الاختبار القبلي، مقابل ١٤,١٤٪ في الهمزة؛ غير أن الهمزة شكّلت أكبر مصدر منفرد من مجموع الأخطاء بنسبة ٣٥,١٣٪. وتؤكد النتائج أن الإملاء الصوتي المنظم تشخيصيًّا، لا نسخ النصوص، يحقق تحسنًا إملائيًّا قابلًا للقياس، مما يجعل الإملاء أداةً تشخيصيةً علاجيةً.

كلمات أساسية الإملاء العربي؛ الإملاء التقويمي؛ المختبر الصوتي؛ التقويم التشخيصي؛ أخطاء الكتابة العربية

Please cite this article as Hula, I.R.N., Hasaniyah, N., Subait, W.F., Hairuddin, & Mariana, A. (2026). Audio-Laboratory-Based Evaluative Orthography as a Diagnostic Approach for Errors in Writing Arabic Texts. *Izdihar : Journal of Arabic Language Teaching, Linguistics, and Literature*, 9(2), 149-178. DOI: https://doi.org/10.22219/jiz.v8i2.45039

INTRODUCTION

Proficient writing of Arabic texts constitutes a core competency in the acquisition of the Arabic language (Khanabadi et al., 2026). This competency involves more than the mechanical arrangement of letters; it requires the integration of sound recognition, word form comprehension, and the application of Arabic orthographic rules to auditory material (Hansen, 2020). Therefore, writing Arabic necessitates the coordinated engagement of hearing, vision, and motor skills, rather than functioning solely as a copying activity. (Mohsen et al., 2026; Saiegh-Haddad & Schiff, 2025; Taha, 2013). This process is supported by the relationship between graphemes and phonemes as learners transcribe spoken language into written form (Bouti, 2025). Auditory input does not always correspond directly to written forms; thus, active cognitive engagement is essential for accurately writing each word according to the rules of *imlāʾ* (Zarytovskaya & Al-Rahbi, 2026).

Within Arabic orthography instruction, three primary approaches to dictation can be distinguished: copied, observed, and evaluative orthography. These approaches correspond to the classical Arabic typology of dictation: imlāʾ manqūl, in which learners copy a text that remains visible; *imlāʾ manẓūr*, in which learners examine a text that is then withdrawn before reproduction; and *imlāʾ ikhtibārī*, in which learners write a text they have not previously seen, relying on auditory reception and internalized orthographic rules (Al-Gani, 2012; Harun, 2005; Ibrahim, 1975). The OEA-Lab Model adopts the technique of *imlāʾ ikhtibārī* but reorients its function toward *imlāʾ taqwīmī*, in which dictation serves evaluative and diagnostic purposes rather than merely for examination. The written product is therefore treated not only as a score but as evidence to be analyzed. Three features distinguish OEA-Lab from conventional *imlāʾ ikhtibārī*: (1) the stimulus is pre-recorded audio delivered through individual laboratory devices, standardizing the source and maximum exposure; (2) items are constructed against the sub-rules of the *Qawāʿid Khamsah* so each response is traceable to a specific rule; and (3) responses are coded into an error map that informs remediation rather than being reduced to a single total score. In the present study, this reorientation is conceptualized as evaluative orthography.

Previous studies have established that Arabic orthography is a complex writing system involving short vowels, diacritics, letter forms, and grammatical context. Arabic writing can be represented in vocalized and unvocalized forms, with unvocalized texts requiring readers to rely more heavily on contextual, grammatical, syntactic, and lexical knowledge (Mohamed & Sayyed, 2019). Arabic is a consonantal alphabetic system written from right to left, with orthographic characteristics that differ from those of alphabetic languages such

as English (Abu-Rabia, 2021; Haddad-Najjar & Abu-Rabia, 2025). In terms of instructional media, Al-Razgan and Alotaibi (2022) found that mobile games can enhance Arabic orthographic competence, engagement, motivation, and student interaction. Nevertheless, these studies have focused primarily on the effectiveness of instructional media and have not addressed the systematic evaluation of Arabic orthography assessment instruments.

Despite the breadth of existing research on Arabic orthography instruction, the gap operates at two levels. From a theoretical standpoint, prior studies have mapped individual sub-rules such as *Hamzah waṣl, Hamzah qaṭʿ,* and *Alif Layyinah* primarily within print-based or cognitively oriented frameworks, leaving the integration of these sub-rules into audio-based laboratory assessment instruments largely unexplored. From an empirical standpoint, the pedagogical context of Islamic higher education in Indonesia particularly in Arabic Language Education (Pendidikan Bahasa Arab/PBA) programmes has received limited attention in the orthographic assessment literature. Evidence from a diagnostic pretest administered to 25 second-semester PBA students at the Language Laboratory (LAB UPA Bahasa), IAIN Sultan Amai Gorontalo, further indicates the need for such an assessment framework: across 337 orthographic items in 29 audio files, the highest error concentration was observed in the rule of *Ziyādah* (31.45%) and the lowest in *Hamzah* (14.14%). These figures indicate that orthographic difficulties are not uniformly distributed across the *Qawāʿid Khamsah*, yet existing instruments have not been designed to diagnose such rule-specific error patterns systematically.

In response to these gaps, this study proposes *Qawāʿid Khamsah*-based evaluative orthography as an assessment framework that links listening skills with competence in writing Arabic texts. In this framework, the five rules function not merely as instructional content but as primary indicators of students' orthographic accuracy (Abandah et al., 2022; Aloufi, 2022). Arabic spelling research shows that orthographic errors can reflect interacting phonological and morphological demands rather than a single mechanical weakness (Abu-Rabia, 2021; Issa, 2023). Operationally, the solution is implemented as an audio-based evaluation of Arabic orthography through a computer laboratory. The 337 scored units are lexical items; in some audio files, target items are embedded in short phrases or sentences to provide auditory context, but the unit of scoring remains the individual lexical item. Students write the heard material while attending to the *Qawāʿid Khamsah*. The written output is evaluated using an evaluative orthography rubric to map error types, accuracy levels, and rule aspects that have not yet been mastered (Zeghli et al., 2022). In this way, Arabic orthography learning shifts from mere dictation practice into a diagnostic process linking auditory input to accurate Arabic writing.

Please cite this article as Hula, I.R.N., Hasaniyah, N., Subait, W.F., Hairuddin, & Mariana, A. (2026). Audio-Laboratory-Based Evaluative Orthography as a Diagnostic Approach for Errors in Writing Arabic Texts. *Izdihar : Journal of Arabic Language Teaching, Linguistics, and Literature*, 9(2), 149-178. DOI: https://doi.org/10.22219/jiz.v8i2.45039

The closest existing work stops short of this combination. Al-Razgan and Alotaibi (2022) improved orthographic competence through mobile games, but their instrument measures engagement and general accuracy rather than rule-specific mastery. Umri et al. (2026) built an AI-driven Arabic listening e-module that raised comprehension scores substantially, yet its output is a listening score and not an orthographic error profile. Hakim and Wahid (2026) evaluated an *imlā'* e-module with a one-group pretest–posttest design and reported a significant gain, but the module delivers written material rather than controlled auditory input, and the outcome is a single accuracy score reported without an effect size or a mapping of errors to specific orthographic rules. None of these integrates a controlled audio-laboratory stimulus, a *Qawā'id Khamsah* diagnostic rubric, and competency measurement within a single instrument, which is the gap the OEA-Lab Model is designed to fill.

To the best of the authors' knowledge, therefore, no prior study has proposed or empirically validated an audio-laboratory-based evaluative orthography model, termed here the OEA-Lab Model, that simultaneously serves as a dictation-based writing instrument, a *Qawā'id Khamsah* diagnostic rubric, and a competency-measurement tool integrating receptive listening with productive Arabic writing in a controlled laboratory setting. The present study addresses this gap through a pre-experimental one-group pretest–posttest design (N = 25), in which the mean score rose from 83.28 to 87.03, a statistically significant improvement with a large effect size. The present study addresses this gap through a pre-experimental one-group pretest–posttest design involving 25 students.

Based on this framework, the present study aims to (1) identify students' orthographic errors according to the *Qawā'id Khamsah*; (2) map the orthographic rules that most frequently serve as sources of error; and (3) develop an audio-based Arabic orthography learning model as an evaluative instrument that is diagnostic, measurable, and consistent with Arabic orthographic rules.

The contribution of this study is twofold. Theoretically, it reconceptualizes Arabic orthography from a technical writing skill to an evaluative instrument that integrates the receptive skill of listening and the productive skill of writing within a unified diagnostic framework. Practically, the integration of audio-based evaluative orthography, computer laboratories, and a *Qawā'id Khamsah*-based rubric enables lecturers to identify students' error patterns, determine orthographic aspects requiring reinforcement, and design more targeted, effective, and evidence-based instructional interventions.

METHOD

This study employed a pre-experimental one-group pretest-posttest design (O_1 X O_2), with O_1 as the pretest, X as the treatment, and O_2 as the posttest. Because an equivalent control group was not feasible in a single intact class, the findings are interpreted cautiously as indications of changes in students' orthographic competence following the treatment rather than causal claims free from external variables.

The experiment was conducted at the UPA Language Laboratory of IAIN Sultan Amai Gorontalo, equipped with computers, headsets, Arabic keyboards, and digital storage for individual listening and typing. The target population was Arabic Language Education program (PBA) students who had completed the *Qawāʿid al-Imlāʾ* course. Through purposive sampling, 25 second-semester students were selected based on three criteria: (1) completion of the course, (2) prior Arabic Rapid Typing (ART) experience, and (3) willingness to complete the full sequence. All participants completed every stage resulting in a 100% completion rate. No a priori power analysis was conducted; the sample of 25, reflecting one intact class, is acknowledged as a limitation.

The experimental materials comprised 337 scored Arabic lexical items organized into 29 audio files, developed based on three references on Arabic orthography: *al-Kafi fi Qawāʿid al-Imlāʾ wa al-Kitabah* (Ayman Amin ʿAbd al-Ghani), *Qawāʿid al-Imlāʾ wa ʿAlamat al-Tarqim* (ʿAbd al-Salam Muhammad Harun), and *Qawāʿid al-Kitabah wa al-Imlāʾ* (Wali, 2015). Items were structured according to the *Qawāʿid Khamsah* framework of five orthographic rules: Hamzah (140 items), Alif (68 items), Ḥażf (62 items), *Ziyādah* (22 items), and *Waṣl-Faṣl* (45 items). The detailed distribution of indicators, subcategories, and item counts per rule is reported in the Results section.

Short phrases or sentences were used only as auditory context in selected files; they were not counted as separate scoring units. The individual lexical item served as the unit of scoring. A further set of 29 items on hāʾ and tāʾ was administered in the same session but was excluded from the present analysis because the study was framed exclusively around the *Qawāʿid Khamsah*; those items are reserved for separate reporting.

Treatment was delivered through audio-based evaluative Arabic orthography procedure: students listened through individual laboratory devices and retyped the heard material into a provided Evaluative Orthography Word document. Participants could replay each file at their discretion up to a maximum of three times; replay frequency was therefore not fixed, although the same maximum exposure was applied to all participants. During the task, students could not alter the worksheet format, use dictionaries, books, the internet,

translation applications, or autocorrect, nor consult classmates. Completed files were saved using each student's assigned name and identification number and uploaded to a designated Google Drive folder. The experiment was conducted over across four 60-minute meetings (Table 1).

Table 1. Experimental Procedure

Meeting	Duration	Activity	Output
Meeting 1	60 minutes	Technical orientation, explanation of the procedure, device checking, and pretest	Initial data on students' orthographic competence
Meeting 2	60 minutes	Treatment through audio-based evaluative Arabic orthography	Students' written answers
Meeting 3	60 minutes	Continuation of the treatment through audio-based evaluative Arabic orthography	Students' written answers
Meeting 4	60 minutes	Posttest, final submission, and checking of answer documents	Final data on students' orthographic competence

Instrument development proceeded through five stages: (1) formulating indicators, (2) developing 337 items across 29 audio files, (3) preparing worksheets/answer keys/rubrics, (4) expert validation, and (5) revision. Three validators with distinct roles examined the instruments: an Arabic orthography expert evaluated the accuracy of the *Qawā'id Khamsah* materials, answer keys, and classifications, an Arabic learning/maharah kitabah expert evaluated their suitability for second-semester students, and an educational evaluation expert evaluated the rubric, test format, scoring procedures, and feasibility. Validators were distinct from the raters. Validation used qualitative expert judgment; the instrument was declared feasible after revision. Quantitative content-validity estimates (CVI, Aiken's V) are recommended for future development.

Internal consistency of the 337 dichotomous items was estimated with the Kuder-Richardson formula 20 and was high in both administrations (pretest KR-20 = 0.853; and posttest KR-20 = 0.841). Subscale reliabilities were uneven, however, and were unacceptably low for the shortest subscale, *Ziyādah* (22 items), which is acknowledged as the rule-specific reliability estimates.

The study obtained institutional approval from the study program and the language laboratory. Participants received an explanation of the objectives, procedures, duration, data types, and academic use of the data; participation was voluntary, and participants were given an opportunity to ask questions before the study began. Participants' identities were kept confidential**.** As no ethics-committee approval number had yet been obtained, this is stated as an administrative limitation.

Data analysis comprised error identification, classification, scoring, percentage calculation, diagnostic interpretation, and statistical testing. Each student answer was compared with the validated answer key; every word not conforming to Arabic orthographic rules was classified into one of the five *Qawāʿid Khamsah* categories. Where one word contained multiple error types, all were recorded for diagnostic purposes but the word was counted once as incorrect in the final score. The diagnostic framework linking each rule category to its interpretation is presented in the Results section.

Scoring used a dichotomous system (correct = 1, incorrect = 0), the final score calculated as follows: final score = (Number of Correct Words / 337) × 100. The percentage of errors per category was computed as Errors in Category / Words in Category × 100 to identify dominant error sources. Statistical testing used IBM SPSS Statistics version 25 at α = 0.05. Normality was tested with Shapiro-Wilk given the small sample (n = 25). Because the difference scores violated the normality assumption ($p<0.05$), pretest–posttest differences were evaluated using the non-parametric Wilcoxon signed-rank test. The effect size was reported using r, computed as $r = |Z| / \sqrt{N}$, where N denotes the total number of observations rather than the number of pairs, that is, 2 × 25 = 50 (Field, 2013). Values were interpreted as 0.10 = small, 0.30 = medium, and ≥ 0.50 = large.

Student products were scored by the researcher using the established orthography rubric across the five *Qawāʿid Khamsah* categories. Because a single rater conducted the primary assessment, inter-rater reliability was not calculated; to mitigate this, the same rubric, answer key, and reference texts were applied to all data, and each correction was manually rechecked by the researcher for consistency. Scoring consistency was thus maintained procedurally rather than through coefficients such as Cohen's kappa. The inclusion of a second independent rater and the calculation of inter-rater reliability are recommended for future research. The analysis techniques and decision criteria are summarized in Table 2.

Table 2. Data Analysis Techniques and Decision Criteria

Statistical Component	Test/Technique	Purpose	Decision Criteria
Normality test	Shapiro-Wilk	Testing the normality of the posttest-minus-pretest difference scores	Sig. > 0.05 = normal; Sig. < 0.05 = non-normal
Difference test	Wilcoxon signed-rank test	Comparing pretest and posttest scores when difference scores violate normality	Sig. < 0.05 = significant difference

Please cite this article as Hula, I.R.N., Hasaniyah, N., Subait, W.F., Hairuddin, & Mariana, A. (2026). Audio-Laboratory-Based Evaluative Orthography as a Diagnostic Approach for Errors in Writing Arabic Texts. *Izdihar : Journal of Arabic Language Teaching, Linguistics, and Literature*, 9(2), 149-178. DOI: https://doi.org/10.22219/jiz.v8i2.45039

Statistical Component	Test/Technique	Purpose	Decision Criteria
Effect size	Effect size r $r = \frac{\lvert Z \rvert}{\sqrt{N}}$	Measuring the magnitude of the treatment effect for non-parametric paired data	0.10 = small; 0.30 = medium; >= 0.50 = large

The study assumed adequate basic Arabic literacy and technical readiness among participants, audio meeting clarity and tempo standards, and a validated instrument representing the *Qawāʿid Khamsah* indicators. As the scope covered only listening and orthographic writing among 25 second-semester PBA students (337 items, 29 audio files), and the one-group design carries threats of testing effect, maturation, and regression to the mean alongside the small purposive sample, the results are read as indications of effectiveness within a pre-experimental context rather than causal evidence; future research should add a control group, enlarge the sample, and strengthen inter-rater reliability. Assessment consistency was maintained procedurally: all pretest and posttest data were corrected with the same rubric, answer key, and reference texts, then re-verified by the researcher.

RESULTS & DISCUSSION

Identification of Orthographic Error Forms Based on the *Qawāʿid Khamsah*

Before presenting the data, the implementation process is outlined to provide procedural context. The experiment ran over four 60-minute meetings in the UPA Language Laboratory, where each student worked individually with a computer, headset, and Arabic keyboard. The first meeting comprised technical orientation and the pretest, in which students transcribed audio-recorded Arabic lexical units into the Evaluative Orthography worksheet while applying the *Qawāʿid Khamsah*. This procedure established their baseline competence prior to the intervention. The second and third meetings delivered the treatment: students listened to each audio file, played up to three times, and rewrote the material without dictionaries, the internet, translation applications, or autocorrect, after which their output was evaluated against the validated answer key. The fourth meeting administered the posttest using an instrument identical to the pretest, following a two-week interval intended to minimize potential practice effects. The resulting scores, both aggregate and disaggregated by rule, are presented in Table 3.

Table 3. Descriptive Statistics of Test Scores

Test Type		N	Min	Max	Mean	SD	Median
Pretest Score		25	71.81	90.50	83.28	4.82	83.98
Rule	*Hamzah*		75.00	96.43	85.86	6.25	85.71
	Alif		76.47	94.12	85.71	5.62	86.76
	Ḥażf		72.58	93.55	85.23	6.56	87.10
	Ziyādah		50.00	81.82	68.55	8.70	68.18
	Waṣl-Faṣl		62.22	91.11	76.09	6.26	77.78
Posttest Score		25	78.34	92.58	87.03	4.11	86.94
Rule	*Hamzah*		77.86	96.43	87.60	5.57	88.57
	Alif		77.94	100.00	90.00	4.73	89.71
	Ḥażf		77.42	96.77	88.97	5.56	88.71
	Ziyādah		63.64	90.91	80.18	6.28	81.82
	Waṣl-Faṣl		64.44	93.33	81.42	7.93	80.00

Table 3 presents descriptive statistics for the pretest and posttest scores, both in aggregate and disaggregated across the five categories of *Qawāʿid Khamsah*. Overall, the mean score increased from 83.28 in the pretest to 87.03 in the posttest, representing an increase of 6.53 points, indicating a higher minimum level of performance at posttest. In addition, the relatively close mean and median values in both measurements indicate that the central tendencies of the distributions were broadly similar.

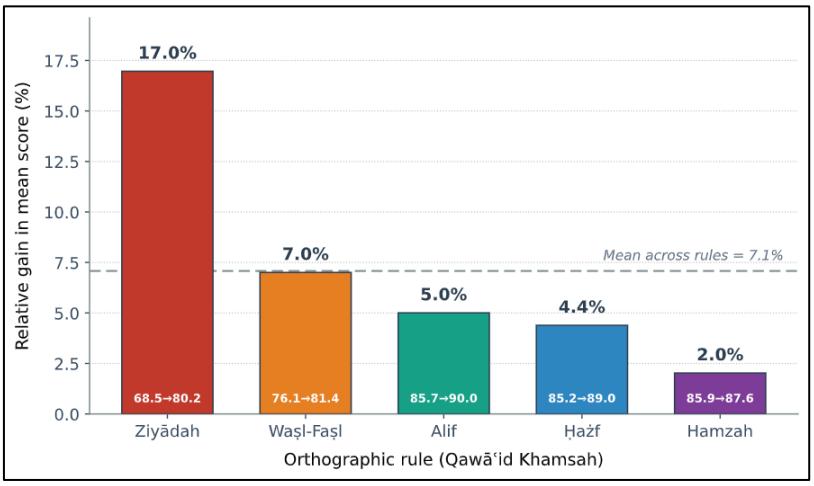

Figure 1. Interpretation of the rule-specific gains

Figure 1 shows that disaggregation by rule provides a more informative pattern. In the pretest, Hamzah (85.86), Alif (85.71), and Ḥażf (85.23) showed relatively high mean scores, whereas Ziyādah (68.55) and Waṣl-Faṣl (76.09) showed comparatively lower mean scores. The pattern of improvement was broadly inverse to the initial performance levels: Ziyādah improved most, by 17.0%, followed by Waṣl-Faṣl at 7.0%, while the three rules that were already high rose only between 2.0% and 5.0%, giving an unweighted mean of 7.1% across the five rules. The two weakest rules therefore produced the two largest

gains. This pattern indicates the diagnostic-remedial effect of audio-laboratory-based evaluative orthography and suggests that repeated auditory-graphemic exposure closes mastery gaps precisely where they are widest, particularly in the differentiation of final word forms and the rules of letter addition, rather than merely sharpening competencies that were already established.

Table 4. Results of the Shapiro-Wilk Normality Test

Tests of Normality						
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Difference of Posttest - Pretest	.236	25	.001	.904	25	.023
a. Lilliefors Significance Correction						

Normality was assessed using the Shapiro–Wilk test on the posttest–pretest difference scores, as the paired-samples t test assumes approximate normality of these scores. As shown in Table 4, the Shapiro-Wilk statistic for the difference scores was W = 0.904, p = 0.023, indicating a mild departure from normality. Consequently, the non-parametric Wilcoxon signed-rank test was conducted as the primary analysis instead of the paired-samples t-test.

Table 5. Wilcoxon Signed-Rank

Test Statistics[a]	
	Pretest – Posttest
N	25
Z	-4.382[b]
R	.619
Asymp. Sig. (2-tailed)	< .001
a. Wilcoxon Signed Ranks Test	
b. Based on positive ranks.	

Note. N = 25 pairs. The effect size r is calculated as |Z| / √N with N = 50, where N represents the number of paired observations

As presented in Table 5, the Wilcoxon signed-rank test revealed a statistically significant difference between pretest and posttest scores, Z = −4.382, p < .001. The effect size was large (r = .619), indicating a substantial practical difference between the two measurements.

As the correction process was conducted by a single rater, inter-rater reliability was not assessed. Consistency in assessment was ensured by applying a uniform correction rubric to all pretest and posttest data, encompassing the *Qawāʿid Khamsah* categories (*Hamzah, Alif, Ḥażf, Ziyādah,* and *Waṣl-Faṣl*). Final decisions were verified by the researcher using reference texts and the rubric.

Therefore, while the correction results can be considered procedurally standardized, they do not equate to inter-rater reliability.

Overall, audio-laboratory-based evaluative orthography improved students' scores across all five Qawāʿid Khamsah categories. The Wilcoxon signed-rank test confirmed a statistically significant pretest–posttest difference ($Z = -4.382$, $p < .001$), with a large effect size ($r = .619$).

The reduction in errors was closely related to the characteristics of the intervention: students did not merely see word forms, but listened to audio input, processed the sounds, and rewrote them in Arabic script. The pathway from listening to writing was followed by *Qawāʿid Khamsah*-based error evaluation, which strengthened sensitivity to sound-graphic relations, especially in *Alif*, *Ḥażf*, *Ziyādah*, and *Waṣl-Faṣl* categories that require decisions about whether a sound element should be written, omitted, added, joined, or separated. However, because the degree of reduction differed across categories, Arabic orthographic competence is not a single uniform ability; each rule has its own difficulty structure and instructional needs. The following subsection maps this structure in detail.

Mapping the Orthographic Rules That Most Frequently Generated Errors

Based on phonological-orthographic analysis of all erroneous items, the dominance of errors in the *Qawāʿid Khamsah* can be explained through four interacting factors.

Factor 1. Complexity of Linguistic Rules. Arabic orthographic rules generally require the simultaneous integration multiple levels of phonological, morphological, syntactic, semantic, and/or pragmatic knowledge (Halpern, 2008). Cherifi and Guerti (2021) emphasize that the complexity of this multilayered mapping is one of the most persistent sources of orthographic-processing deficits in Arabic language learning.

Factor 2. Lexical Distribution. Rules with broad morphological distribution, which appear in many productive patterns, and rules with very low lexical exposure in regular instructional input may both become frequent sources of error, although through different mechanisms (Masrai, 2016; Saiegh-Haddad et al., 2023).

Factor 3. Auditory Perception. Because this study is audio-laboratory-based, auditory perception is crucial. Arabic has a relatively deep orthographic system, in which grapheme-phoneme relations are not one-to-one (Al Qunayeer, 2020). Three modes of auditory difficulty appeared in the data: (a) acoustic similarity between two sounds mapped onto different graphemes; (b) phonologically similar or identical sounds represented by different graphemes; and (c) silent

letters or missing sounds (*idgham*, *Ḥażf*) that require mental reconstruction of the original word form. Haluti et al. (2024) show that Arabic audiovisual media effectively improve auditory discrimination, while Balshe and Masterson, (2026) add that purely sound-based differentiation is insufficient and must be complemented by explicit visual cues.

Factor 4. Visual Differentiation. Arabic script involves positional allographs, letter forms that change according to position, ligaturing, non-linearity through diacritics above or below the main line, and visual minimal pairs such as (ﺩ) versus (ﺫ) (Milo & Martínez, 2019). AlJassmi dan Perea (2024) confirm that multiple letter forms and visual similarity among graphemes are persistent sources of orthographic difficulty across learner levels.

Hamzah: Complexity of Harakat and Acoustic Similarity of Vowels

Table 6. Mapping of Orthographic Errors in the Hamzah Rule by Sub-Rule

No.	Orthographic Sub-Rule	Σ Tested Items	Σ Incorrect Items	Error Share (%)	Item-Level Error Rate (%)
1	*Waṣl Asmāʾ Thamāniyah*	8	1	3.23%	12.50%
2	*Waṣl Amar Thulāthī*	4	0	0.00%	0.00%
3	*Waṣl Māḍī Khumāsī*	3	0	0.00%	0.00%
4	*Waṣl Amar Khumāsī*	3	3	9.68%	100.00%
5	*Waṣl Maṣdar Khumāsī*	4	0	0.00%	0.00%
6	*Waṣl Māḍī, Amar, and Maṣdar Sudāsī*	2	0	0.00%	0.00%
7	*Qaṭʿ Awal Ism ʿAlam*	4	0	0.00%	0.00%
8	*Qaṭʿ Awal Ḍamīr*	3	0	0.00%	0.00%
9	*Qaṭʿ Awal Thulāthī*	4	0	0.00%	0.00%
10	*Qaṭʿ Awal Rubāʿī*	4	0	0.00%	0.00%
11	*Qaṭʿ Awal Amar Rubāʿī*	4	0	0.00%	0.00%
12	*Qaṭʿ Awal Maṣdar Rubāʿī*	3	0	0.00%	0.00%
13	*Qaṭʿ Awal Muḍāraʿah*	3	0	0.00%	0.00%
14	*Qaṭʿ Awal Ḥurūf Maʿānī*	7	0	0.00%	0.00%
15	*Qaṭʿ Waṣaṭ ʿalā Alif*	10	2	6.45%	20.00%
16	*Qaṭʿ Waṣaṭ ʿalā Wāw*	21	13	41.94%	61.90%
17	*Qaṭʿ Waṣaṭ ʿalā Yāʾ*	16	7	22.58%	43.75%
18	*Munfaridah Waṣaṭ al-Kalimah*	8	2	6.45%	25.00%
19	*Qaṭʿ Akhir*	18	3	9.68%	16.67%
20	*Qaṭʿ Akhir fī Naṣb*	11	0	0.00%	0.00%
	Total	**140**	**31**	**100.00%**	**22.14%**

Table 6 shows four levels of dominance. High dominance consists of *Qaṭʿ Waṣaṭ ʿalā Wāw* (13 errors, 41.94%) and *Qaṭʿ Waṣaṭ ʿalā Yāʾ* (7 errors, 22.58%), together contributing 64.52% of all recorded errors. Moderate dominance includes *Waṣl Amar Khumāsī* (اِنْتَظِرْ، اِجْتَهِدْ) and *Qaṭʿ Akhir* (3 errors each, 9.68%). Low dominance includes *Qaṭʿ Waṣaṭ ʿalā Alif* (سَأَلَ، رَأْس) and *Munfaridah Waṣaṭ al-Kalimah* (مُرُوءَة، تَسَاءَلَ) at 2 errors each, and *Waṣl Asmāʾ Thamāniyah* (اِسْم، اِبْن) with 1 error.

A fourth level, absent from the distribution altogether, comprises thirteen sub-rules covering 56 items in which no error occurred, including all eight *Qaṭʿ Awal* sub-rules (أَكَلَ، إِبْرَاهِيم، أُسْتَاذ). The distribution of errors also varies according to word position: word-initial *Hamzah* (56 items) accounted for 4 errors (7.14%), *Waṣaṭ Hamzah* (55 items) accounted for 24 errors (43.64%), and *akhir Hamzah* (29 items) accounted 3 errors (10.34%). Item-level rates point the same way, with *Qaṭʿ Waṣaṭ ʿalā Wāw* at 61.90% and *Qaṭʿ Waṣaṭ ʿalā Yāʾ* at 43.75%; the sole higher value, *Waṣl Amar Khumāsī* at 100.00%, was based on only three tested items and therefore represents a narrow pocket of difficulty rather than a dominant source of error.

Two factors converge on medial position. Factor 1 is rule complexity: the carrier follows the *al-aqwā fa al-aqwā* principle (*kasrah* > *ḍammah* > *fatḥah* > *sukūn*), which requires comparing the vowel of the hamzah with that of the preceding letter. In (تَفَاؤُل) the hamzah carries ḍammah after a fatḥah, so the stronger vowel determines wāw as the carrier; in (سُؤَال) the *hamzah* carries fatḥah after a ḍammah, and the same principle again selects wāw. Factor 3 is auditory: ḍammah (/u/) is acoustically close to *wāw madd* (/uː/), and *kasrah* (/i/) to *yāʾ madd* (/iː/), so in مَشِيئَة and مَلِيئَة learners write the *yāʾ nabrah* inconsistently, possibly reflecting interference from first-language phonological categories. Factors 2 and 4 reinforce this: hamzah on wāw recurs across productive patterns such as (رُؤُوس) and (تَفَاؤُل), and demands visual precision in placing the small hamzah above its carrier (Yassin et al., 2020; Taha & Khateeb, 2018). Faizah et al. (2024) report parallel difficulties in maṣdar nouns of khumāsī and sudāsī verbs.

The error-free sub-rules confirm this reading from the opposite direction. Qaṭʿ Awal requires no comparison of adjacent vowels, since the hamzah is written on alif regardless of its own vowel (أَكَلَ، إِبْرَاهِيم، أُسْتَاذ), and Waṣl Asmāʾ Thamāniyah is limited to eight fixed high-frequency words with a stable written form (اِسْم، اِبْن، اِبْنَة، اِمْرَأَة). The absence of errors in these sub-rules is consistent with the view that reduced decision-making demands may facilitate accurate

orthographic production. However, this pattern should not be attributed exclusively to cognitive load, because lexical frequency, auditory characteristics, and visual form may also contribute to performance.

Alif: Etymological Complexity and Homophonous Orthographic Forms

Table 7. Alif Sub-Rule Error Mapping

No.	Orthographic Sub-Rule	Σ Tested Items	Σ Incorrect Items	Error Share (%)	Item-Level Error Rate (%)
1	*Yābisah*	6	0	0.00%	0.00%
2	*Layyinah*	14	4	26.67%	28.57%
3	*Layyinah fī al-Afʿāl*	16	5	33.33%	31.25%
4	*Layyinah fī al-Ḥurūf*	10	1	6.67%	10.00%
5	*Maqṣūrah*	14	4	26.67%	28.57%
6	*Mamdūdah*	8	1	6.67%	12.50%
	Total	**68**	**15**	**100.00%**	**22.06%**

Table 7 shows a highly uneven distribution: (a) the highest concentration of errors occurred in *Alif Layyinah fi al-Afʿal* (5 errors, 33.33%), *Alif Layyinah* (4 errors, 26.67%), and *Alif Maqṣurah* (4 errors, 26.67%), totalling 86.67%; (b) low dominance in *Alif Layyinah fi al-Huruf* and *Alif Mamdudah* (1 error each, 6.67%); and (c) complete mastery in *Alif Yabisah* (0 errors).

The Alif category shows a different error pattern from that of Hamzah, with the distribution of errors potentially related to etymological–morphological complexity and the limited informativeness of auditory input for distinguishing homophonous orthographic forms. In Factor 1, Regarding etymological–morphological complexity Alif Layyinah depends on the letter's origin (*aṣl al-kalimah*): the *Maqṣurah* form (ى) is generally used when the underlying final radical is yāʾ, whereas the Alif Mamdūdah form (ا) is used when the underlying final radical is wāw. In triliteral forms, the orthographic choice requires consideration of the origin of the final radical, whereas derived verb forms may follow different orthographic conventions and exceptions. Thus, (قَسَا) (with a wāw-based origin represented by Alif (ا)) versus (يَرْقَى) (with a yāʾ-based origin represented by* Alif Maqṣūrah (ى)) requires etymological analysis, including identifying the root structure, determining the morphological origin of the final radical, and selecting the corresponding orthographic form (Khresat & Ahmad, 2025).

What makes Alif distinct from Hamzah in this respect, as factor 3, is the different role of auditory information. In *Hamzah*, auditory difficulty arises from acoustic similarity between two different sounds; in *Alif*, difficulty arises from one

long /a:/ sound that is identical for two different orthographic forms (*Maqṣurah* versus *mamdudah*). The words (قَسَا) and (يَرْقَى) both end in /a:/, as do (العَصَا) and (كِسْرَى). Consequently, auditory input alone may not provide sufficient information to distinguish between these orthographic alternatives, suggesting that accurate production of Alif requires attention to both orthographic form and morphological structure. Auditory differentiation being insufficient, the audio-laboratory approach must be complemented by explicit visual cues and structured etymological analysis (Brierley et al., 2016; Cuevas, 2016). Factor 4 may also contribute to the observed pattern: *Alif* has at least four visual variants (ا, ى, آ, أ/إ), and *Maqṣurah* is visually identical to undotted final *yā'*, causing students to write it incorrectly as ordinary *yā'* (Retnawati et al., 2020).

Ḥażf: Missing Sounds and Visual Multiplicity

Table 8. Ḥażf Sub-Rule Error Mapping

No.	Orthographic Sub-Rule	Σ Tested Items	Σ Incorrect Items	Error Share (%)	Item-Level Error Rate (%)
1	*Ḥażf Alif Awal Kalimah*	11	2	14.29%	18.18%
2	*Ḥażf Alif Waṣaṭ Kalimah*	13	1	7.14%	7.69%
3	*Ḥażf Alif Akhir Kalimah*	7	2	14.29%	28.57%
4	*Ḥażf Lām*	4	1	7.14%	25.00%
5	*Ḥażf Wāw*	9	1	7.14%	11.11%
6	*Ḥażf Yā'*	13	4	28.57%	30.77%
7	*Ḥażf Nūn*	5	3	21.43%	60.00%
	Total	**62**	**14**	**100.00%**	**22.58%**

Table 8 shows that 14 incorrect responses were recorded among the 62 tested items, corresponding to an overall item-level error rate of 22.58%. These errors were distributed across all seven sub-rules, with the highest error shares occurring in *Ḥażf Yā'* (4 errors, 28.57%) and *Ḥażf Nūn* (3 errors, 21.43%), together accounting for 50.00% of the total errors. The remaining errors were distributed across *Ḥażf Alif* Awal Kalimah and *Ḥażf Alif* Akhir Kalimah (2 errors each, 14.29%), *Ḥażf Lām* (1 error, 7.14%), *Ḥażf Wāw* (1 error, 7.14%), and *Ḥażf Alif Waṣaṭ Kalimah* (1 error, 7.14%).

The *Ḥażf* category illustrates why the two percentages should not be conflated. *Ḥażf Yā'* contributes the largest share of distinct incorrect items within the rule (4 of 14; 28.57%), whereas *Ḥażf Nūn* has the highest item-level error rate (3 of 5 tested items; 60.00%). Thus, error share identifies where the error

Please cite this article as Hula, I.R.N., Hasaniyah, N., Subait, W.F., Hairuddin, & Mariana, A. (2026). Audio-Laboratory-Based Evaluative Orthography as a Diagnostic Approach for Errors in Writing Arabic Texts. *Izdihar : Journal of Arabic Language Teaching, Linguistics, and Literature*, 9(2), 149-178. DOI: https://doi.org/10.22219/jiz.v8i2.45039

pool is concentrated, while item-level error rate identifies the sub-rule with the greatest difficulty relative to its test coverage.

The *Ḥażf* category has another distinctive signature of difficulty: Factor 1 (syntactic-morphological complexity) and Factor 3 in the mode of silent letters or missing sounds. In Factor 1, letter deletion depends on syntactic-morphological context rather than on the word form itself. *Ḥażf Yā'* occurs in indefinite *manquṣ* nouns, final weak verbs in jussive form, *iḍafah* in sound masculine plural, and vocative forms attached to yā' al-mutakallim. *Ḥażf Nūn* occurs in *idgham* and in defective verbs. Accurate application of these deletion rules therefore requires the integration of morphological, syntactic, and orthographic knowledge (Cheragui & Hiri, 2020).

In Factor 3, the difficulty of *Ḥażf* arises from orthographic deletion in which a letter is absent from the written form despite the grammatical or morphological information associated with the underlying form be detected. This contrast with *Hamzah*, whose source of difficulty is sound similarity. In الداع (from الداعي), the word is still read with *kasrah* without a clear yā' sound; in مِمَّ (from مِنْ مَا), idgham produces pronunciation without *nun* (Alhaider, 2025). Al Hanai and Glass (2014) describe this as a classic challenge of Arabic deep orthography: students must mentally reconstruct the original word form before writing it. Asadi et al. (2017) recommend explicit modelling that presents the original word form before its *Ḥażf* form.

Factor 4 adds further complexity: in *Ḥażf Yā',* students must differentiate three possible orthographic realizations involving *yā'*: explicitly written *yā'* (ي), zero realization resulting from deletion, and representation through *Alif Maqṣūrah* (ى). The absence of an explicit visual marker for a deleted letter may further increase the difficulty of applying *Ḥażf rules* (Al-Qali et al., 2021). This is reflected in the consistency of errors across all seven *Ḥażf* sub-rules; none reached complete mastery as *Alif Yabisah* did in the *Alif* category.

Ziyādah: Silent Letters and Narrow Lexical Distribution

Table 9. Ziyādah Sub-Rule Error Mapping

No.	Orthographic Sub-Rule	Σ Tested Items	Σ Incorrect Items	Error Share (%)	Item-Level Error Rate (%)
1	*Ziyādah* Alif	5	1	12.50%	20.00%
2	*Ziyādah* Wāw	8	2	25.00%	25.00%
3	*Ziyādah* Hā' Saktah	9	5	62.50%	55.56%
	Total	**22**	**8**	**100.00%**	**36.36%**

Table 9 shows that 8 incorrect responses were recorded among the 22 tested items, corresponding to an overall item-level error rate of 36.36%. Within the error set, *Hā' Saktah* contributed the largest share (5 of 8; 62.50%) and also the highest sub-rule item-level rate (5 of 9; 55.56%), followed by *Ziyādah Wāw* (25.00%; 25.00%) and *Ziyādah Alif* (12.50%; 20.00%). Thus, both error share and item-level error rate identify *Hā' Saktah* as the most prominent source of difficulty within the *Ziyādah* category. Unlike the other four orthographic categories, however, the present analysis does not indicate whether this pattern is attributable to a broader rule-level characteristic without considering the distribution and nature of the tested items.

What makes *Ziyādah* distinctive is the interaction between Factor 2 and Factor 3. In Factor 2, *Hā' Saktah* has a relatively narrow lexical distribution: imperative forms such as (قِهْ، عِهْ، رَهْ) are infrequent in ordinary instructional texts, whereas forms such as (سُلْطَانِيَهْ) and (مَاهِيَهْ) are more typically encountered in Qur'anic contexts, particularly in *waqf*. In the present dataset, this relatively limited lexical distribution coincided with the highest concentration of errors within the *Ziyādah* category, although the observed pattern does not by itself establish that limited exposure caused the errors. More generally, Arabic spelling is influenced by the interaction of phonological, morphological, and orthographic knowledge rather than by sound-to-letter correspondence alone (Yassin et al., 2020).

Factor 3 concerns the mismatch between auditory input and written representation. In several *Ziyādah* forms, the required grapheme is not fully predictable from pronunciation alone. In *Hā' Saktah*, for example, the final hā' is associated with word-final *waqf*, while in *Ziyādah Wāw*, (أُولِي) contains an orthographic *wāw* that is not transparently represented in the auditory signal, and (عَمْرٌو) takes an additional *wāw* to distinguish it orthographically from (عُمَر). Such cases illustrate the limitations of a purely "write-as-you-hear" strategy. Research on Arabic spelling likewise shows that phonological processing constitutes a major source of spelling difficulty and that accurate spelling requires learners to integrate phonological information with language- and script-specific orthographic knowledge (Abu-Rabia & Taha, 2006; Yassin & Shalhoub-awwad, 2025).

Ziyādah Alif adds a conventional-historical dimension. The alif in (مائة) is retained in the classical orthographic tradition, as in (ثلاثمائة، أربعمائة، خمسمائة), whereas contemporary usage also permits the form (مئة). Mastery of this pattern therefore depends not only on phonological processing but also on knowledge of

conventional orthographic forms. Factor 4 further contributes to the difficulty because ه (*hāʾ saktah*) and ة (*tāʾ marbūṭah*) are visually similar, differing primarily in dotting. This visual similarity is relevant to Arabic spelling because script-specific features such as letter-form similarity, positional allography, ligaturing, and non-linearity have been shown to constitute substantial sources of orthographic difficulty (AlJassmi & Perea, 2024).

Ziyādah showed the largest improvement of the five rules, with the mean score rising from 68.55 to 80.18, an increase of 11.63 points and a relative gain of 17.0%. The item distribution provides one possible context for interpreting this pattern: 9 of the 22 *Ziyādah* items targeted *Hāʾ Saktah*, the sub-rule in which the written form is particularly difficult to recover from auditory input alone. The intervention nevertheless provided repeated auditory exposure and sub-rule-specific feedback, including practice with forms for which direct sound-to-writing mapping was insufficient. This instructional alignment may have contributed to the observed improvement, although the present design does not permit the gain to be attributed to the intervention alone.This interpretation is consistent with research showing that Arabic spelling depends on the integration of phonological, morphological, and orthographic information (Abu-Rabia & Taha, 2006; Issa, 2023; Yassin et al., 2020) Nevertheless, the rule-level estimate should be interpreted cautiously because the *Ziyādah* subscale contains only 22 items and its reliability estimate was the lowest among the five rule categories.

***Waṣl-Faṣl*: Idgham and Functional Variation of Particles**

Table 10. *Waṣl-Faṣl* Sub-Rule Error Mapping

No.	Orthographic Sub-Rule	Σ Tested Items	Σ Incorrect Items	Error Share (%)	Item-Level Error Rate (%)
1	*Murakkab Majzī*	3	1	6.25%	33.33%
2	*Raʾs*	3	1	6.25%	33.33%
3	*Idhin*	5	3	18.75%	60.00%
4	*Ḥabbā*	1	0	0.00%	0.00%
5	*Mā Istifhāmiyyah*	8	3	18.75%	37.50%
6	*Mā Maṣdariyyah*	4	1	6.25%	25.00%
7	*Mā Zāʾidah*	9	3	18.75%	33.33%
8	*Man Istifhāmiyyah*	5	2	12.50%	40.00%
9	*Man Mawṣūliyyah*	4	1	6.25%	25.00%
10	*Lā Nāfiyah*	2	0	0.00%	0.00%
11	*Kayfa*	1	1	6.25%	100.00%
	Total	**45**	**16**	**100.00%**	**35.56%**

Table 10 shows that *6* incorrect responses were recorded among the 45 tested items, corresponding to an overall item-level error rate of 35.56%. this was second highest after *Ziyādah* (36.36%). The largest error shares were observed for *Idhin* (3 errors, 18.75%), *Mā Istifhāmiyyah* (3 errors, 18.75%), and *Mā Zāʾidah* (3 errors, 18.75%), which together accounted for 56.25% of all errors. These were followed by *Man Istifhāmiyyah* (2 errors, 12.50%). The remaining errors were distributed across *Murakkab Majzī, Raʾs, Mā Maṣdariyyah, Man Mawṣūliyyah,* and *Kayfa* (1 error each, 6.25%). No errors were recorded for Ḥabbā or Lā Nāfiyah in the present dataset.

The *Waṣl-Faṣl* category differs from all previous categories in its difficulty: Factor 1 appears in its most complex mode, integrating four layers of knowledge, and Factor 3 appears as phonological-orthographic mismatch caused by idgham. In Factor 1, the determination of waṣl versus faṣl is governed by the grammatical function of particles in a sentence, not by the phonetic or morphological form of the word. Idhin is joined when it functions as a temporal adverb forming a murakkab ẓarfi, as in (حِيْنَئِذٍ, عندئذٍ), and (ساعتئذٍ); without this syntactic knowledge, students tend to write it separately as (حين إذٍ). Ma istifhamiyyah is fused with a preceding preposition (حَتَّامَ= حَتَّى +مَا, فِيْمَ= فِي +مَا, مِمَّ= مِن +مَا), while ma zaidah functions as an emphatic particle without lexical meaning (طَالَمَا, قَلَّمَا, نِعِمَّا). These items require coordination of grammatical function and orthographic form; in the present dataset, this makes *Waṣl-Faṣl* qualitatively different from categories governed primarily by a single graphemic decision.

The difficulty is intensified by functional variation: one particle may have multiple grammatical functions. The particle mā, for example, may function as interrogative, relative, negative, maṣdariyyah, additional, conditional, or exclamatory, each with different implications for *Waṣl-Faṣl*; man likewise has several functions. In the present data, this functional multiplicity plausibly contributes to error because students must identify the grammatical function before applying the orthographic rule. This interpretation is grounded in the observed distribution across *Waṣl-Faṣl* sub-rules rather than attributed to a non-matching external study.

In Factor 3, *Waṣl-Faṣl* displays a unique phenomenon: *idgham*, or sound fusion, which obscures word boundaries in pronunciation. In (مِمَّ) (مِن + مَا), the sound nun is assimilated into *mim* so that students hear a single unit rather than two joined words; the same applies to (فِيْمَ). Saiegh-Haddad et al., (2023) describe this as phonological-orthographic mismatch, where pronunciation does not reflect orthographic structure. A write-as-you-hear approach therefore fails

to reveal whether a phonetic unit derives from one word or two joined ones, as in (طَالَمَا، قَلَّمَا) where *ma zaidah* is not transparently recoverable from pronunciation alone.

Factor 4 appears in the present data as word-boundary ambiguity: (مِمَّ) differs from (مِنْ مَا) only by spacing, as does (حِيْنَئِذٍ) from (حين إذٍ) (Maciuszek, 2018). Arabic spelling research shows that script-specific features such as ligaturing, positional allography, and non-linearity can create additional visual and orthographic demands(Yassin & Shalhoub-awwad, 2025). In the present data, spacing and letter connection were especially relevant to *Waṣl-Faṣl*, with errors occurring in 9 of 11 sub-rules.

An interesting finding in this category is the coexistence of two sub-rules with no recorded errors for *Ḥabbā and Lā Nāfiyah*, alongside sub-rules with very high error rates, such as *Idhin* (60.00%) and *Kayfa* (100.00%). This contrast suggests uneven development of orthographic knowledge within the same learners, with some sub-rules already mastered while others remain highly error-prone. In the present data, differences in familiarity and instructional exposure may therefore help explain why mastery varied substantially across *Waṣl-Faṣl* sub-rules.

Formulation of the Audio-Laboratory-Based Evaluative Orthography Model

Based on the findings of the first and second research questions, this study formulates the Audio-Laboratory-Based Evaluative Orthography Model (OEA-Lab). This model is a design for both learning and evaluating Arabic orthography that positions auditory input as the starting point, Arabic writing production as the core activity, and *Qawāʿid Khamsah*-based error analysis as the basis for diagnostic feedback. The formulation is derived from empirical evidence showing that the mean score increased from 83.28 to 87.03 (Δ = 3.75 points), with a statistically significant Wilcoxon signed-rank test result ($Z = -4.382$, $p < .001$) and a large effect size ($r = .619$).

Conceptually, the model rests on the assumption that Arabic orthographic errors are not caused solely by weak memorization of rules, but by students' limited ability to connect sounds, word structure, graphic forms, and writing rules. In Arabic, sound-writing relationships are complex: not all sounds are represented transparently in writing, not all written letters are pronounced clearly, and several different graphic forms can produce the same sound (Cherifi & Guerti, 2021). Therefore, Arabic orthography learning must move beyond text-copying exercises; students need to be trained to listen, recognize sound

structures, analyze word forms, write correct orthographic forms, and evaluate errors based on rule categories (Dramé, 2021).

The OEA-Lab Model consists of five main components that also function as operational stages: (1) auditory input; (2) phonological-orthographic processing; (3) writing production or transcription; (4) *Qawā'id Khamsah*-based error analysis; and (5) remedial feedback and developmental evaluation (Muñoz, 2022).

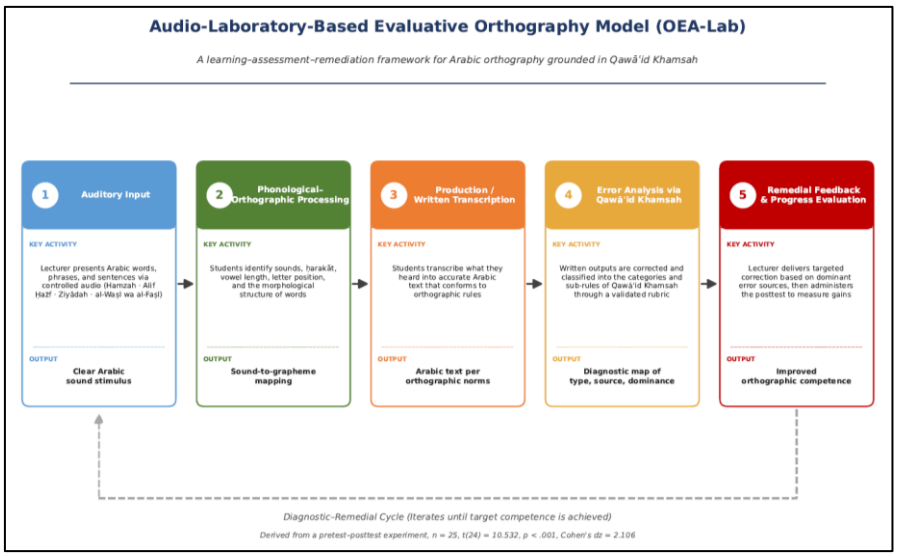

Figure 2. The Audio-Laboratory-Based Evaluative Orthography Model (OEA-Lab)

Based on the four-factor framework discussed above, the OEA-Lab Model has different remedial orientations for each rule: for *Hamzah*, identification of the *Hamzah* vowel, the preceding letter's vowel, and carrier selection; for *Alif*, etymological-morphological analysis to distinguish *Maqṣurah* and *mamdudah*; for *Ḥażf*, reconstruction of the original word form before deletion; for *Ziyādah*, recognition of added letters that are not heard in the audio, especially *Hā' Saktah*; and for *Waṣl-Faṣl*, differentiation between joined and separated elements through syntactic-semantic analysis. The syntax of the OEA-Lab Model is summarized in Table 11.

Table 11. Audio-Laboratory-Based Evaluative Orthography Model

No.	OEA-Lab Model Stage	Main Activity	Purpose
1	Auditory Input	students listen to controlled Arabic audio containing words, phrases, or short sentences through controlled audio containing *Hamzah, Alif, Ḥażf, Ziyādah,* and *Waṣl-Faṣl* elements.	To provide standardized auditory stimuli as the basis for subsequent orthographic processing.
2	Phonological-Orthographic Processing	Students identify sounds, harakat, long and short vowels, letter positions, and word morphology.	To develop the ability to connect auditory information with Arabic orthographic rules.

No.	OEA-Lab Model Stage	Main Activity	Purpose
3	Writing Production/Transcription	Students write the words or sentences they hear in correct Arabic script.	To transform auditory input into written representation according to Arabic orthographic conventions.
4	*Qawāʿid Khamsah*-Based Error Analysis	Students' written products are corrected and classified into *Hamzah, Alif, Ḥażf, Ziyādah, and Waṣl-Faṣl* categories and their sub-rules.	To identify and map the types and distribution of orthographic errors for diagnostic purposes.
5	Remedial Feedback and Developmental Evaluation	students receive rule-specific corrective feedback and revise their responses, followed by an evaluation of subsequent performance.	To address identified difficulties and monitor changes in orthographic competence.

The strength of this model lies in its diagnostic, measurable, and remedial character. It identifies specific error types, tracks development through pretest-posttest comparison, and turns error analysis directly into instructional improvement. The model also carries four practical implications for imlaʾ learning in PBA programs. First, *imlaʾ* instruction should be designed from actual error data rather than textbook sequence. Second, listening practice should be integrated with writing practice, since many orthographic errors may reflect difficulties sound-to-writing mapping. Third, language laboratories can serve not only istimāʿ and muḥadathah but also evaluation-based orthography learning. Fourth, lecturers should provide correction rubrics with *Qawāʿid Khamsah* categories and subcategories so that error evaluation becomes more objective and evidence-based.

The finding that audio-laboratory-based evaluative orthography was associated with a statistically significant improvement in students' orthographic competence ($Z = -4.382$, $p < .001$, $r = .619$) is consistent with the theoretical premise that dictation is not a single mechanical act but the integration of several cognitive processes operating simultaneously. Recent work confirms that Arabic dictation requires the coordination of phonological awareness, working memory, and orthographic knowledge, so that accuracy reflects deeper language processing rather than rote recall (Issa, 2023).

The model's emphasis on sound–grapheme mapping aligns with studies showing that the complexity of Arabic orthography lies in the inconsistent correspondence between phonemes and graphemes, including positional allographs, the high functional load of short-vowel diacritics, and root–pattern morphology that complicates grapheme–phoneme mapping (H. Almusawi, 2023). This explains why, in the present study, the *Hamzah* and *Alif* categories, both

heavily dependent on distinguishing sound realization from written form, were already comparatively well mastered yet still benefited from repeated auditory-graphemic exposure. The result corroborates error-analysis research reporting that *Hamzah waṣl* and *Hamzah qaṭ'* remain among the most frequent sources of Arabic spelling error even in prepared writing and therefore require explicit, structured instruction rather than incidental acquisition (Ghazouani et al., 2024).

The largest mean improvement occurred in the *Ziyādah* category, the aspect identified as most difficult in the pretest. This pattern resonates with morphology-focused studies demonstrating that Arabic spelling errors are strongly tied to morphological structure, since learners must reconstruct the underlying word form to decide whether a letter is added or omitted (Yassin & Shalhoub-awwad, 2025). The diagnostic-remedial character of the OEA-Lab Model, which directs learners' attention to letters that are written but not heard such as *Alif* added after the plural waw or the silent *Hā' Saktah* appears to close mastery gaps in exactly the categories where sound input alone is insufficient, supporting the claim that explicit rule-based remediation is essential for the most structurally complex orthographic rules (Kassem et al., 2025).

The remedial-feedback component of the model is grounded in the theory of diagnostic and formative assessment, which holds that feedback is most effective when it is criterion-referenced and maps specific strengths and weaknesses rather than merely marking errors (Lee & Jang, 2025). By classifying each error into *Qawā'id Khamsah* categories and sub-rules, the model transforms correction into targeted diagnostic information that becomes the basis for instructional improvement, echoing evidence that diagnostic feedback fosters deeper engagement and self-regulated learning than generic corrective marking (Gebremariam, 2024).

Finally, the use of the language laboratory as a medium for evaluation-based orthography learning provides a controlled environment for standardized auditory input and repeated listening-to-writing practice. In the present study, all participants received the same recorded stimuli through individual laboratory devices, with each audio file replayed up to a maximum of three times, allowing auditory exposure to be standardized across participants. The magnitude of improvement reported here parallels findings from comparable Indonesian pre-experimental studies; for example, an AI-driven Arabic listening e-module raised second-semester students' mean scores from 55.4 to 85.9 with an N-Gain of 0.68 (Umri et al. 2026). Recent digital *imlā'* research suggests that technology-supported instruction can facilitate more structured and interactive practice (Hakim & Wahid, 2026; Kusnawati et al., 2025). The present study extends this line of work by integrating controlled auditory input, Qawā'id Khamsah-based

Please cite this article as Hula, I.R.N., Hasaniyah, N., Subait, W.F., Hairuddin, & Mariana, A. (2026). Audio-Laboratory-Based Evaluative Orthography as a Diagnostic Approach for Errors in Writing Arabic Texts. *Izdihar : Journal of Arabic Language Teaching, Linguistics, and Literature*, 9(2), 149-178. DOI: https://doi.org/10.22219/jiz.v8i2.45039

diagnostic error mapping, and sub-rule-specific remedial feedback in a laboratory setting.

CONCLUSIONS

This study demonstrates that audio-laboratory-based evaluative orthography functions as a diagnostic solution in learning to write Arabic texts. This is evident from the increase in students' mean score from 83.28 in the pretest to 87.03 in the posttest, with a difference of 3.75 points. The Wilcoxon signed-rank test showed a statistically significant difference ($Z = -4.382$, $p < .001$), with a large effect ($r = .619$). Thus, the model not only supports improvement in writing accuracy but also helps map the types and sources of students' orthographic errors.

Students' errors were not evenly distributed across the Qawāʿid Khamsah categories, but were concentrated in specific sub-rules. The largest error shares were found in *Hamzah Qaṭʿ Waṣaṭ ʿalā Wāw* (41.94%) and *Qaṭʿ Waṣaṭ ʿalā Yāʾ* (22.58%), *Alif Layyinah fī al-Afʿāl* (33.33%), *Ḥażf Yāʾ* (28.57%), and *Ziyādah Hāʾ Saktah* (62.50%). In *Waṣl-Faṣl, Idhin, Mā Istifhāmiyyah,* and *Mā Zāʾidah* each contributed 18.75% of the category's errors.

Based on these findings, the study formulates the OEA-Lab Model, consisting of five stages: auditory input, phonological-orthographic processing, writing production or transcription, Qawāʿid Khamsah-based error analysis, and remedial feedback and developmental evaluation. The model positions imlāʾ not merely as dictation practice, but as an evaluative, diagnostic, and remedial instrument for directing instructional follow-up.

Theoretically, this study strengthens the view of Arabic orthography as a writing system that requires the integration of auditory, phonological, morphological, visual, and cognitive dimensions. Practically, *imlāʾ* instruction should be designed on the basis of students' actual error profiles and should utilize the language laboratory as a setting for data-based orthographic evaluation.

This study has several limitations, including the absence of a control group, a relatively small intact sample of 25 participants without an a priori power analysis, and the use of a single primary rater without inter-rater reliability estimation. Variation in audio replay frequency and the weaker internal consistency of the relatively small Ziyādah subscale should also be considered when interpreting the findings. Future studies should therefore employ larger and adequately powered controlled samples, multiple raters, and more balanced rule-level instruments.

ACKNOWLEDGMENT

The authors express their gratitude for the support provided by the Faculty of Tarbiyah and Teacher Training and the Academic Development Unit (UPA) for Language at IAIN Sultan Amai Gorontalo.

BIBLIOGRAPHY

Abandah, G., Suyyagh, A., & Khedher, M. Z. (2022). Correcting Arabic Soft Spelling Mistakes using BiLSTM-based Machine Learning. *International Journal of Advanced Computer Science and Applications*, *13*(5), 815–829. https://doi.org/10.14569/IJACSA.2022.0130594

Abu-Rabia, S. (2021). Syntax Functions and Reading Comprehension in Arabic Orthography Syntax Functions and Reading Comprehension in Arabic Orthography. *Reading Psychology*, *0*(0), 1–30. https://doi.org/10.1080/02702711.2021.1912866

Abu-Rabia, S., & Taha, H. (2006). Phonological Errors Predominate in Arabic Spelling Across Grades 1–9. *Journal of Psycholinguistic Research*, *35*(2), 167–188. https://doi.org/10.1007/s10936-005-9010-7

AbuZeina, D., & Elshafei, M. (2012). Modeling of Arabic Cross-Word Pronunciation Variations. In *SpringerBriefs in Speech Technology* (pp. 47–58). https://doi.org/10.1007/978-1-4614-1213-7_5

Al-Gani, A. A. A. (2012). *Al-Kafi fi Qawaid al-Imla'*. Dar al-Taufiq li al-Turats. www.moswarat.com

Al-Qali, M., Abandah, G., & Suyyagh, A. (2021). Correcting Arabic Soft Spelling Mistakes Using Transformers. *2021 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology, JEEIT 2021 - Proceedings*, 146–151. https://doi.org/10.1109/JEEIT53412.2021.9634142

Al Hanai, T., & Glass, J. (2024). Lexical modeling for Arabic ASR: A systematic approach. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 2605–2609. https://doi.org/10.21437/Interspeech.2014-558

Al Qunayeer, H. S. (2020). The perils of Arabic consonants: What do the English speakers have to say? *Asian EFL Journal*, *27*(42), 22–40. https://www.scopus.com/inward/record.uri?eid=2-s2.0-

85096048015&partnerID=40&md5=3287fa31121cf665280266e5d6988752

Alhaider, S. (2025). Optimality and correspondence theories in phonological shifts: a case study on Arabic guttural consonants in English loanwords. *Frontiers in Language Sciences*, *4*. https://doi.org/10.3389/flang.2025.1535723

AlJassmi, M. A., & Perea, M. (2024). Visual similarity effects in the identification of Arabic letters: evidence with masked priming. *Language and Cognition*, *16*(4), 1618–1638. https://doi.org/10.1017/langcog.2024.20

Almusawi, H. (2023). Factors Affecting the Writing Performance in Hearing and Deaf Children: An Insight into Regularities and Irregularities of the Arabic Orthographic System. *Language and Speech*, *66*(1), 246–264. https://doi.org/10.1177/00238309221097714

Almusawi, H. M. (2019). Determinants of spelling proficiency in hearing and deaf graduate students: The presentation of medial glottal stop. *Ampersand*, *6*. https://doi.org/10.1016/j.amper.2019.100050

Aloufi, S. (2022). Writing the (Hamza) at the beginning of a Word in the Inscriptions of Makkah and Madinah during the First Two Hijri Centuries. *Dirasat: Human and Social Sciences*, *49*(6), 413–432. https://doi.org/10.35516/hum.v49i6:.4038

Asadi, I. A., Ibrahim, R., & Khateb, A. (2017). What contributes to spelling in Arabic? A cross-sectional study from first to sixth grade. *Writing Systems Research*, *9*(1), 60–81. https://doi.org/10.1080/17586801.2016.1218748

Balshe, R., & Masterson, J. (2026). Word learning in Arabic-speaking primary school students. *European Journal of Psychology of Education*, *41*(1). https://doi.org/10.1007/s10212-025-01061-2

Brierley, C., Sawalha, M., Heselwood, B., & Atwell, E. (2016). A Verified Arabic-IPA Mapping for Arabic Transcription Technology, Informed by Quranic Recitation, Traditional Arabic Linguistics, and Modern Phonetics. *Journal of Semitic Studies*, *61*(1), 157–186. https://doi.org/10.1093/jss/fgv035

Cheragui, M. A., & Hiri, E. (2020). Arabic Text Segmentation using Contextual Exploration and Morphological Analysis. *Proceedings of the 2020 International Conference on Mathematics and Information Technology, ICMIT 2020*, 220–225. https://doi.org/10.1109/ICMIT47780.2020.9046976

Cherifi, E.-H., & Guerti, M. (2021). Conditional random fields applied to Arabic orthographic-phonetic transcription. *Archives of Acoustics*, *46*(2), 237–247. https://doi.org/10.24425/aoa.2021.136574

Cuevas, J. A. (2016). Visual and auditory learning: Differentiating instruction via sensory modality and its effects on memory. In *Student Achievement: Perspectives, Assessment and Improvement Strategies* (pp. 29–54). https://www.scopus.com/inward/record.uri?eid=2-s2.0-85029949302&partnerID=40&md5=8d5da376c7256443d4c348371469feb2

Dramé, D. (2021). Soninke in Islamic manuscripts: Correlation between graphemic and phonemic correspondences. *Mandenkan*, *66*, 187–203. https://doi.org/10.4000/mandenkan.2634

Faisal Umri, Muhbib Abdul Wahab, Maswani. (2026). Developing and Evaluating an AI - Driven Arabic Listening E - Module : Integrating Intensive and Extensive Listening Through Elevenlabs. *Al-Ta'rib*, *14*(1), 151–168. https://doi.org/10.23971/altarib.v14i1.11306

Gebremariam, H. T. (2024). Exploring the Effects of Written Corrective Feedback Types on Grammatical Accuracy in L2 Writing : Evidence From Ethiopian High School Students. *SAGE Open*, *2*(September), 1–16. https://doi.org/10.1177/21582440241274331

Ghazouani, H. A. S., Khalil, A. A., & Abu-Ayyash, E. A. S. (2024). EFL Arab Learners' Orthographic Errors: A Contrastive Analysis. In *English Language Education* (Vol. 35, pp. 3–17). https://doi.org/10.1007/978-3-031-69789-0_1

Hacimisiroğlu, M. (2024). The Phenomenon of Istidrak in Muṣḥaf Orthography Literature: Example of Ibn al-Qāḍī al-Miknāsī (d. 1082/1672) and His Book Named Bāyānu'l-Khilāf. *Hitit Theology Journal*, *23*(1), 280–299. https://doi.org/10.14395/hid.1418607

Haddad-Najjar, L., & Abu-Rabia, S. (2025). Exploring cross-linguistic transfer of phonology , morphology , syntax , and orthography from Arabic (L1) to English (FL): An intervention study. *International Journal of Bilingualism*, *29*(X), 1657 –1680. https://doi.org/10.1177/13670069241300286

Hakim, A., & Wahid, M. (2026). E-Module Based Learning as an Effort to Improve Understanding of Imla ' Among Islamic Religious Education Students. *Halaqa: Islamic Education Journal*, *10*(1), 1–11. https://doi.org/10.21070/halaqa.v10i1.1758

Halpern, J. (2008). Exploiting lexical resources for disambiguating CJK and Arabic orthographic variants. *Proceedings of the 6th International Conference on Language Resources and Evaluation, LREC 2008*, 429–435. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85037539334&partnerID=40&md5=80371cbf17de7c2d1ba3710584225de5

Haluti, A., Mile, M. P., Mukmin, M., Hardianti, H., & Arrasul, A. R. (2024). Exploring the Use of Audio-Visual Teaching Media to Improve Students' Ability in Listening Comprehension. *Pakistan Journal of Life and Social Sciences*, *22*(2), 11081–11088. https://doi.org/10.57239/PJLSS-2024-22.2.00839

Hansen, G. F. (2020). Word recognition in Arabic as a foreign language. *Modern Language Journal*, *94*(4), 567–581. https://doi.org/10.1111/j.1540-4781.2010.01094.x

Harun, A. M. (2005). *Qawaid al-Imla wa Alamah al-Tarqim* (p. 79). Maktabah al-Sa'iy.

Hassan, T. (2013). Marginal phenomena in the grammar system, with supporting evidence from the Qur'an. *Journal of Qur'anic Studies*, *15*(2), 203 – 215. https://doi.org/10.3366/jqs.2013.0108

Ibrahim, A. al-A. (1975). *Al-imlāʾ wa al-Tarqīm fī al-Kitābah al-ʿArabiyyah.* Maktabat Gharīb.

Issa, I. (2023). Morphological Complexity in Arabic Spelling and Its Implication for Cognitive Processing. *Journal of Psycholinguistic Research*, *52*(1), 331–357. https://doi.org/10.1007/s10936-022-09896-2

Kassem, A. S., Schiff, R., & Saiegh, E. (2025). Development of morphological awareness in Arabic : the role of morphological system and morphological distance. In *Reading and Writing* (Vol. 38, Issue 8). Springer Netherlands. https://doi.org/10.1007/s11145-024-10581-0

Khanabadi, T., Motaghizadeh, I., Ameri, H., & Monazam, H. N. (2026). A Classification of Arabic Grammar Based on Typology for the Writing Course of Undergraduate Students of Arabic Language and Literature. *Language Related Research*, *16*(6), 73–114. https://doi.org/10.48311/LRR.16.6.3

Khresat, M. S., & Ahmad, S. H. (2025). Some Transformations of al'ajwaf "the mid-voweled" and alnaaqis "the end-voweled" verbs: A Descriptive Analytical Applied Study in Light of Modern Linguistics. *Dirasat: Human and Social Sciences*, *52*(4). https://doi.org/10.35516/hum.v52i4.6646

Kusnawati, Y., Miswandono, A., & Luthfi, A. (2025). Reimagining Arabic Imlā' Instruction in the Digital Era:A Systematic Review of Mobile Learning Innovations. *EL-IBTIKAR Jurnal Pendidikan Bahasa Arab*, *14*(2), 196–210.

Lee, M., & Jang, E. E. (2025). Automated Diagnostic Feedback vs . Self-Assessment : Rethinking Feedback Development. *TESOL Quarterly*, *59*(December 2025), 280–317. https://doi.org/10.1002/tesq.70032

Maciuszek, J. (2018). Lexical Access in the Processing of Word Boundary Ambiguity. *Social Psychological Bulletin*, *13*(4). https://doi.org/10.32872/spb.v13i4.28690

Masrai, A. M. (2016). The influence of morphological knowledge on lexical processing and acquisition: The case of Arab EFL learners. *Ampersand*, *3*, 52–60. https://doi.org/10.1016/j.amper.2016.04.001

Milo, T., & Martínez, A. G. (2019). A new strategy for Arabic OCR archigraphemes, letter blocks, script grammar, and shape synthesis. *ACM International Conference Proceeding Series*, 93–96. https://doi.org/10.1145/3322905.3322928

Mohamed, E., & Sayyed, Z. A. (2019). Arabic-SOS : Segmentation, Stemming, and Orthography Standardization for Classical and pre-Modern Standard Arabic. *DATeCH*, *19*(2), 27–32. https://doi.org/10.1145/3322905.3322927

Mohsen, M., Alsudairy, N. A., Alhamami, M., & Al-Hoorie, A. H. (2026). Exploring cognitive processes in Arabic dictation: A study on writing challenges among dyslexic children. *Acta Psychologica*, *262*. https://doi.org/10.1016/j.actpsy.2025.106118

Muna Al-Razgan, H. M. A. (2022). Integrating Mobile Games in Arabic Orthography Classrooms. *Arab World English Journal (AWEJ)*, *10*(8), 146–165. https://doi.org/10.24093/awej/call8.10

Muñoz, C. (2022). Audiovisual input in L2 learning. *LIA Language, Interaction and Acquisition*, *13*(1), 125–143. https://doi.org/10.1075/lia.22001.mun

Retnawati, H., Rahmatullah, S., Djidu, H., & Apino, E. (2020). Has Arabic language learning been successfully implemented? *International Journal of Instruction*, *13*(4), 715–730. https://doi.org/10.29333/iji.2020.13444a

Saiegh-Haddad, E., Ghawi-Dakwar, O., Haj, L., Farraj-Bsharat, R., & Laks, L. (2023). SPELLING ARABIC: When Does Orthographic Knowledge End and Language Knowledge Start? In *Routledge International Handbook of Visualmotor Skills, Handwriting, and Spelling: Theory, Research, and Practice* (pp. 276–292). https://doi.org/10.4324/9781003284048-25

Saiegh-Haddad, E., & Schiff, R. (2025). Diglossic and Orthographic Features of Reading Comprehension in Standard Arabic: The Primacy of the Spoken Language. *Reading Research Quarterly*, *60*(1). https://doi.org/10.1002/rrq.598

Suleman Bouti, Asna Ntelu, Muziatun. Ela Budiarti, Ari, D. D. (2025). Orthographic Representation of Long Vowels in Gorontalo : A Sociolinguistic Perspective on Tradition and Technology. *The International Journal of Communication and Linguistic Studies*, *23*(3), 1–22. https://doi.org/10.18848/2327-7882/CGP/v23i03/1-22

Taha, H. Y. (2013). Reading and spelling in Arabic: Linguistic and orthographic complexity. *Theory and Practice in Language Studies*, *3*(5), 721–727. https://doi.org/10.4304/tpls.3.5.721-727

Wali, H. (2015). *Al-Syamil fi Qawa'id al-Kitabah wa al-Im;a'* (A. S. H. Al-Barjal (ed.); 1st ed.). Dar al-Fadhilah.

Yakup, M., Abliz, W., Sereno, J., & Perea, M. (2014). How is letter position coding attained in scripts with position-dependent allography? *Psychonomic Bulletin and Review*, *21*(6), 1600–1606. https://doi.org/10.3758/s13423-014-0621-6

Yassin, R., & Shalhoub-awwad, Y. (2025). Characteristics of early spelling errors in Arabic : the impact of visual-orthographic features. *Reading and Writing*. https://doi.org/10.1007/s11145-025-10731-y

Yassin, R., Share, D. L., & Shalhoub-Awwad, Y. (2020). Learning to Spell in Arabic: The Impact of Script-Specific Visual-Orthographic Features. *Frontiers in Psychology*, *11*(August). https://doi.org/10.3389/fpsyg.2020.02059

Zarytovskaya, V. N., & Al-Rahbi, A. M. (2026). Arabic Homographs: Systemic Language Difficulties In Text Perception. *Vostok (Oriens)*, *2026*(1), 184–196. https://doi.org/10.31696/S086919080034971-5

Zeghli, H., Bouanani, M., & Lété, B. (2022). Orthographic processing in Arabic. *Psychologie Francaise*. https://doi.org/10.1016/j.psfr.2022.01.002