



Employing the Linguistic Corpus in Teaching Arabic Writing Skill: A Proposed Instructional Model for the Arabic Program

توظيف المدونة اللغوية في تدريس مهارة الكتابة بالعربية:

نموذج تعليمي مقترح لبرنامج اللغة العربية

Ilham Fatkhu Romadhon^{1*}, Muassomah², Zakiyah Arifa³

^{1,2,3}Universitas Islam Negeri Maulana Malik Ibrahim Malang, Indonesia

*Corresponding E-mail: ilhamfatkhu689@gmail.com

ABSTRACT

This study proposes an instructional model for teaching Arabic writing skills to students of the Program for Arabic Language (PKPBA) at Maulana Malik Ibrahim State Islamic University Malang, Indonesia, based on the use of a language corpus. The novelty of this model lies in the integration of three instructional resources (PowerPoint presentations, Sketch Engine corpus analysis software, and worksheets) grounded in a single, locally relevant corpus ('Indonesia Al-Youm'), which has not been systematically employed in previous corpus-based Arabic writing instruction. Using a descriptive-analytical method based on document analysis and expert validation, the study found that the teaching materials received an average score of 4.32 (very appropriate) and the learning design received 2.70 (appropriate), confirming the model's suitability for field implementation from an expert perspective. The model consists of five sequential steps: concept presentation, guided corpus exploration, analytical exercises, guided writing production, and reflection with feedback. As an expert-validated proposal, this model offers a conceptual framework that potentially bridges the gap between declarative knowledge and procedural skills in teaching Modern Standard Arabic writing, though empirical testing with learners remains necessary.

Keywords: Data-driven learning; language corpus; Arabic writing skills; sketch engine; Arabic language program.

مستخلص البحث

يهدف هذا البحث إلى اقتراح نموذج تعليمي لتدريس مهارة الكتابة باللغة العربية لطلبة برنامج تعليم اللغة العربية المكثف (PKPBA) في جامعة مولانا مالك إبراهيم الإسلامية الحكومية بمالانج، إندونيسيا، بالاعتماد على توظيف المدونات اللغوية. وتتمثل جدة هذا النموذج في دمج ثلاثة موارد تعليمية، هي: عروض البوربوينت (PowerPoint)، وبرنامج تحليل المدونات اللغوية (Sketch Engine)، وأوراق العمل، اعتمادًا على مدونة لغوية واحدة ذات صلة بالسياق المحلي، وهي مدونة «إندونيسيا اليوم»، التي لم تُوظف من قبل بصورة منهجية في تعليم مهارة الكتابة العربية القائم على المدونات اللغوية. واعتمدت الدراسة المنهج الوصفي التحليلي القائم على تحليل الوثائق والتحكيم العلمي من قبل الخبراء، وأظهرت النتائج أن المواد التعليمية حصلت على متوسط قدره (4,32) بدرجة «مناسبة جدًا»، في حين حصل تصميم التعلم على متوسط قدره (2,70) بدرجة «مناسب»، مما يؤكد صلاحية النموذج للتطبيق الميداني من منظور الخبراء. ويتكون النموذج المقترح من خمس خطوات متتابعة، هي: عرض المفاهيم، والاستكشاف الموجه للمدونة اللغوية، والتدريبات التحليلية، وإنتاج الكتابة بصورة موجهة، ثم التأمل والتغذية الراجعة. وبصفته نموذجًا مقترحًا خضع للتحكيم العلمي، فإنه يقدم إطارًا مفاهيميًا من شأنه أن يسهم في ردم الفجوة بين المعرفة التصريحية والمهارات الإجرائية في تدريس الكتابة باللغة العربية الفصحى المعاصرة، مع بقاء الحاجة إلى إجراء اختبارات تجريبية على المتعلمين للتحقق من فاعليته.

الكلمات المفتاحية: التعلم القائم على البيانات؛ المدونات اللغوية؛ مهارات الكتابة باللغة العربية؛ سكييتش إنجن (Sketch Engine)؛ برنامج اللغة العربية.

Citation:

Romadhon, I. F., Muassomah, & Arifa, Z. (2026). Employing the linguistic corpus in teaching Arabic writing skill: A proposed instructional model for the Arabic program. *Al-Muyassar: Journal of Arabic Education*, 5(2), 727-748.

INTRODUCTION

Writing skill is considered one of the most challenging language skills in teaching Arabic as a foreign language, especially in intensive programs where mastery is expected within a limited period. This difficulty arises from the nature of the Arabic morphological and grammatical system, which requires learners to master interconnected rules in order to produce correct texts – thereby justifying the need to develop modern instructional strategies that go beyond conventional methods (Alhawary, 2009). Traditional approaches to writing instruction have relied on artificial examples and fragmented grammatical exercises, resulting in a wide gap between what is acquired in the classroom and what is actually used in authentic communication (J. Zhang, Yoon, Williams, & Pinkas, 2024).

The urgency of this problem is particularly acute in the Indonesian context. Recent empirical studies have documented significant writing challenges among Arabic language learners in Indonesia. Burhani & Al-Khawaldeh (2026), in a study involving 1,101 Arabic learners across Islamic universities in Indonesia, found that students experience moderate levels of writing anxiety, with female learners reporting higher anxiety than their male counterparts and students in their third year exhibiting greater anxiety than those in earlier semesters. Similarly, Yasmar (2025) reported that approximately 85% of students experience difficulties in writing independently, citing insufficient practice and a lack of constructive feedback from instructors as primary causes. A descriptive analysis of written linguistic errors among non-native Arabic learners further confirmed that writing difficulties persist across grammatical, morphological, and orthographic domains, largely attributed to the complexity of Arabic grammatical rules and the over-reliance of many educational programs on theoretical instruction without sufficient practical application (Emara & Rouyan, 2025). These findings underscore the pressing need for instructional interventions that move beyond rule-based, decontextualized grammar instruction toward authentic, practice-oriented approaches.

Preliminary observation conducted by the researcher in the Arabic Language Program (PKPBA) at Maulana Malik Ibrahim State Islamic University Malang reveals a unique and seemingly paradoxical phenomenon. One of the program's distinct classes is *Fashl al-Jami'ah* (University Class, abbreviated as FJ), a special track designed for new students who demonstrate above-average proficiency in Arabic based on placement test results. Unlike regular classes organized by academic major or faculty, the FJ class brings

together students from various disciplinary backgrounds, creating a diverse and dynamic learning environment. The majority of students in this class are alumni of Islamic boarding schools (*pesantren*), where they have spent years mastering Arabic grammar (*nahwu* and *sharaf*), vocabulary, and reading classical texts. Consequently, these students possess strong declarative knowledge of Arabic rules and can read complex texts with relative ease.

Paradoxically, however, preliminary observations and informal interviews with FJ class instructors indicate that despite their strong grammatical foundation, many of these students struggle significantly with productive writing skills, particularly when required to produce coherent, stylistically appropriate Modern Standard Arabic texts. They often produce sentences that are grammatically correct but lexically unnatural or contextually inappropriate, relying on memorized patterns rather than authentic language use. This gap between receptive competence (reading) and productive performance (writing) mirrors the broader challenges documented in empirical studies (Emara & Rouyan, 2025; Yasmar et al., 2025). The FJ case thus presents a compelling paradox: learners with strong rule-based knowledge still struggle to write naturally because their instruction has been largely text-based and rule-driven, without systematic exposure to authentic corpus data or training in data-driven exploration. This observation reinforces the urgent need for an instructional model that bridges declarative and procedural knowledge—not for novice learners, but for advanced learners who already master the rules yet need to develop native-like writing proficiency through authentic language exposure.

In this context, recent studies emphasize the need to shift toward using authentic texts and language corpora to enhance learners' writing competence and bridge the gap between theory and practice through direct exposure to real language (Pérez-Paredes, 2020). Meta-analyses have further demonstrated that direct exposure to authentic language through corpora significantly improves language production (Boulton & Cobb, 2017).

A modern trend in applied linguistics has emerged known as Data-Driven Learning (DDL), which positions the learner as a language researcher who independently explores patterns from authentic language corpora rather than passively receiving ready-made rules. The importance of this approach lies in its capacity to improve writing accuracy and develop learner autonomy; experimental studies have proven its effectiveness in developing writing skills compared to traditional methods,

particularly in academic contexts (Crosthwaite, 2019). The roots of this trend trace back to the work of Johns (1991), who argued that exposing learners to hundreds of authentic examples is more beneficial than explaining a rule once. In the specific context of Arabic language teaching, systematic reviews have demonstrated the effectiveness of the corpus-based approach in overcoming instructional challenges related to writing skills (Zaki, 2017).

Over the past five years, research on corpus-based writing instruction has expanded considerably. In English as a Foreign Language (EFL) contexts, DDL has been consistently shown to improve lexical accuracy, collocational competence, and learner autonomy (Boulton & Cobb, 2017; Pascual et al., 2019; Pérez-Paredes, 2020). Recent meta-analyses confirm that DDL yields moderate to large effects on language learning outcomes, particularly when integrated into writing instruction (Lusta, Demirel, & Mohammadzadeh, 2023). For Arabic as a foreign language, however, corpus-based pedagogical research remains nascent. Zaki (2017) provided a foundational systematic review demonstrating the potential of corpus-based approaches for Arabic instruction, yet the empirical implementation has lagged behind. A handful of recent studies have begun to address this gap: Fuadi & Sani (2025) developed an Arabic learner text corpus for STAI Ma'arif Kendal Ngawi students to support DDL material development, demonstrating how concordance and word frequency analyses can identify language patterns and common errors. Maachi & Khamar (2025) demonstrated the use of a pedagogic corpus combined with Sketch Engine tools to analyze literary texts in an Arabic classroom, highlighting the potential of corpora to enhance Arabic language teaching. Ashfia (2023) implemented Sketch Engine specifically for Arabic vocabulary learning, showing that analyzing word collocations and meanings improves students' understanding of language data.

Despite these emerging efforts, several critical gaps remain. First, existing studies have predominantly focused on vocabulary acquisition and reading comprehension, with very limited attention to productive writing skills—the most challenging domain for learners. Second, most corpus-based Arabic studies have utilized general or literary corpora rather than corpora tailored to learners' cultural and linguistic backgrounds. Third, none of the existing studies have systematically developed an integrated instructional model that simultaneously combines concept presentation (e.g., PowerPoint), corpus analysis tools (e.g., Sketch Engine), and guided practice worksheets within a coherent pedagogical sequence. Fourth, the specific needs of Indonesian

learners of Arabic in intensive program settings—where mastery is expected within a constrained timeframe—have not been addressed through corpus-based writing instruction.

Thus, a clear research gap exists: there is currently no empirically validated instructional model that employs the ‘Indonesia Al-Youm’ corpus—a locally and culturally relevant media corpus covering contemporary Indonesian affairs in Modern Standard Arabic—to systematically teach Arabic writing skills to students of the Arabic Language Program (PKPBA) at Maulana Malik Ibrahim State Islamic University Malang.

The 'Indonesia Al-Youm' corpus is a contemporary Modern Standard Arabic media source covering Indonesian affairs, making it ideal for Indonesian learners who find its content culturally familiar—thereby reducing the foreignness of the target language. Using specialized, culturally relevant corpora is considered a decisive factor in enhancing learner motivation and reducing language anxiety, which in turn facilitates the assimilation of complex structures (Pérez-Paredes, 2020). This corpus provides advanced analytical tools through Sketch Engine software, including word lists, concordance lists, collocations, and a thesaurus. These tools enable precise exploration of the grammatical and semantic behavior of words, providing a rich environment for fostering self-directed exploration among students (Kilgarriff, Baisa, Bušta, Jakubíček, et al., 2014).

The novelty of the present study lies in three specific aspects. First, it proposes the first integrated instructional model based on the ‘Indonesia Al-Youm’ corpus specifically designed for Arabic writing instruction in an Indonesian intensive program context, addressing a population of learners that has been largely overlooked in existing corpus-based research. Second, unlike previous corpus-based studies that used corpora as a supplementary resource or focused on receptive skills, this model systematically integrates three teaching sources—PowerPoint (concept delivery and model presentation), Sketch Engine (linguistic exploration and data-driven discovery), and worksheets (guided practice and structured application)—into a coherent, sequential pedagogical framework that operationalizes all five stages of DDL-based writing instruction. Third, the model explicitly operationalizes the principle of cultural relevance in corpus selection by utilizing a locally sourced corpus that aligns with Indonesian learners’ background knowledge and communicative needs, an aspect largely absent in existing Arabic corpus-based pedagogical research. From this standpoint, this study seeks to present a proposed instructional model for employing the ‘Indonesia Al-Youm’

corpus in developing writing skills among students of the Arabic Language Program (PKPBA) at Maulana Malik Ibrahim State Islamic University Malang, to analyze its basic components and application mechanisms, and to verify its suitability through expert review.

RESEARCH METHOD

This study adopted the descriptive-analytical method, which involves collecting, analyzing, presenting, and interpreting data without intervening in or manipulating variables. This approach is among the leading methods in applied linguistics research that aims to diagnose the educational reality and propose application models built on solid scientific foundations (Dörnyei, 2007). This method is well-suited to the nature of the study, which seeks to analyze and present a proposed instructional model and verify its suitability through document analysis and expert review.

The study drew on two primary data sources. The first was the textbook '*Al-Arabiyya lil-Hayat 2*,' which represents the programmatic framework of the PKPBA, consisting of eight thematic units. The second was the 'Indonesia Al-Youm' corpus, accessed via the Sketch Engine platform, containing a collection of journalistic texts in contemporary Modern Standard Arabic that cover Indonesian news and affairs. Document analysis, a systematic qualitative method for reviewing and evaluating printed and electronic materials (Dalglish, Khalid, & McMahon, 2020), was employed to extract relevant vocabulary, grammatical patterns, and textual models from both the textbook and the corpus for model development (Rapley & Jenkins, 2010).

Two instruments were developed for expert validation, following established procedures in research and development (Olson, 2010). The first was an expert review form for the teaching materials, comprising twenty items distributed across four axes: (a) suitability of materials for writing skills, (b) presentation of materials, (c) language used, and (d) teaching approach. Items were rated on a five-point Likert scale (1–5). The second instrument was an expert review form for the learning design, consisting of twenty items across four axes: (a) instructional design, (b) media and appearance, (c) interactivity and ease of use, and (d) assessment and feedback. Items for this instrument used a four-point rating scale (1–4). Both instruments underwent content validation by experts in educational research and applied linguistics prior to use (Kania & Kusumah, 2024).

The teaching materials were reviewed by Professor Dr. Danial Hilmi, M.A., a specialist in Arabic language teaching technology, while the learning design was

reviewed by Dr. Muhammad Ahsanuddin, M.Pd., a specialist in language corpus learning design. The selection of multiple experts with distinct areas of specialization is a common practice to ensure comprehensive content validity (Nurhayati, Wahyudi, & Alsagaf, 2021). Their assessments resulted in both an evaluation of the model's suitability for field implementation and a list of recommendations for its improvement (Plomp & Nieveen, 2013).

The data obtained from the expert review were analyzed by calculating the arithmetic mean for each axis and for the overall score. The interpretations were based on the following category criteria: For the five-point scale - (4.0–5.0) = Very Appropriate, (3.0–3.9) = Appropriate, (2.0–2.9) = Sufficiently Appropriate, (1.0–1.9) = Inappropriate. For the four-point scale - (3.0–4.0) = Appropriate, (2.0–2.9) = Sufficiently Appropriate, (1.0–1.9) = Inappropriate. This quantitative interpretation of expert judgments is a standard validation procedure in educational development research (Mardapi, 2018).

RESULT AND DISCUSSION

Teaching Concept and Outcomes

Analysis of the PKPBA curriculum and the characteristics of the 'Indonesia Al-Youm' corpus suggests that the Data-Driven Learning (DDL) model represents a potentially suitable conceptual framework for teaching writing in this context. In principle, this model is designed to transform the learner's role from a passive recipient of rules to an active language researcher who explores patterns from authentic texts. Its proposed value lies in the fact that the journalistic corpus provides authentic rather than artificial examples, and its topics are closely connected to the reality of Indonesian learners. However, it is important to emphasize that these claims are derived from theoretical analysis and expert validation; they have not yet been empirically tested with actual learners. Therefore, the following discussion should be interpreted as a conceptual argument regarding the model's potential, not as evidence of its effectiveness.

The study's analysis yielded the concept of 'Integrated Triple Teaching Resources,' which combines: (1) PowerPoint presentations as a vehicle for presenting concepts and text models; (2) Sketch Engine software as a tool for linguistic exploration and corpus analysis; and (3) worksheets as a space for systematic practical training. This conceptualization integrates theory, data-driven exploration, and writing practice into one unified pathway. This design aligns with the principles of modern teaching materials development, which encourage integrating learners in the process of knowledge

discovery (Tomlinson, 2023). From a design perspective, this integration has the potential to support writing skill development, though its actual effectiveness remains hypothetical until the model is implemented in classroom settings.

The targeted learning outcomes of the model revolve around five axes: (1) identifying key vocabulary, collocations, and variations in their usage contexts within the corpus using Sketch Engine features; (2) employing functional grammatical structures extracted from the corpus in precise written sentences; (3) producing written texts in descriptive, report, comparison, and opinion paragraph forms; (4) independently exploring the corpus to verify the accuracy of vocabulary and structures used; and (5) self-evaluating writing based on clear quantitative criteria. These outcomes align with modern language proficiency standards that emphasize the importance of producing coherent and accurate texts using naturally collocated vocabulary (Laufer & Waldman, 2011). Nevertheless, whether learners can actually achieve these outcomes through the proposed model is a question that requires empirical investigation, as this study has not yet implemented the model with students. Consequently, the discussion presented here remains largely hypothetical and serves as a foundation for future experimental research.

Expert Review Results for Teaching Materials

Professor Dr. Danial Hilmi, M.A. reviewed the suitability of the teaching materials based on the 'Indonesia Al-Youm' corpus for developing writing skills across four main axes. The following table presents the evaluation results by axis:

Table 1. Expert Review Scores for Teaching Materials by Axis

Axis	Category	Mean Score	No. of Items	No. Score 5	No. Score 4
A. Suitability of Materials for Writing Skills	Very Appropriate	4.4	5	2	3
B. Presentation of Materials	Appropriate	4.2	5	1	4
C. Language Used	Very Appropriate	4.4	5	2	3
D. Teaching Approach	Appropriate	4.25	4	1	3
Total / Mean	Very Appropriate	4.32	20	6	13

As shown in Table 1, the overall mean score for all axes is 4.32, which falls within the 'Very Appropriate' category (range 4.0–5.0). No score of 3 (Sufficiently Appropriate) or below was recorded, indicating that the materials meet suitability criteria quantitatively. The axes of 'Suitability of Materials for Writing Skills' and 'Language Used' received the highest mean (4.4 – Very Appropriate).

The expert offered three key recommendations for improvement: first, increasing

the conceptual introduction prior to writing exercises in each unit; second, preparing a vocabulary map classified according to semantic categories; and third, providing concrete examples of vocabulary usage in correct syntactic sentences. He concluded that the materials are 'suitable for field implementation after considering the noted improvement observations.'

Components of the Teaching Materials

Designing a corpus-based curriculum that integrates multiple teaching resources is one of the most significant scientific contributions of this study. The curriculum relies on three integrated teaching sources: the PowerPoint presentation (PPT), Sketch Engine software as the corpus analysis engine, and the worksheet. These sources do not operate in isolation but form a coherent teaching structure that serves the objectives of the intensive Arabic language program (Al-Khuli, 2000).

The significance of this curriculum design lies in its connection between the living language extracted from contemporary Indonesian media—represented by the 'Indonesia Al-Youm' corpus—and the content of the prescribed textbook 'Al-Arabiyya lil-Hayat 2.' This connection provided students with the opportunity to engage with authentic texts that reflect the actual use of Arabic in the media, enhancing their motivation and raising their level of learning (Kilgarriff, Baisa, Bušta, Jakub\ \iček, et al., 2014).

A. PowerPoint Presentation (PPT)

The PowerPoint presentation consists of three main sequential sections reflecting the logical progression in teaching writing skills: the vocabulary section, the grammatical structures section, and the applied writing section. This progression is consistent with modern educational principles based on the transition from simple to complex, and from passive knowledge to productive knowledge (Hamid et al., 2020).

The curriculum comprised eight teaching units covering topics relevant to daily life and the Indonesian environment. The three main integrated sections of the presentation are: (1) key vocabulary extracted from 'Al-Arabiyya lil-Hayat 2' with their frequency in the corpus; (2) functional grammatical structures formulated in light of corpus patterns; and (3) applied writing models that consolidate the transition from comprehension to production. The eight thematic units are summarized in Table 2 below:

Table 2. Summary of Units and Competencies

Unit	Topic	Main Competency	Vocabulary	Structures	Writing Activities
1	Seasons (الفصول)	Applying sentence writing about seasons	Seasonal vocabulary - singular, dual, and plural	Temporal adverb - tens numbers	Word classification - vocabulary writing
2	The Market (السوق)	Applying word writing	Market vocabulary - colors	Mādhā (what) and bikam (how much) - adjectival expression	Word classification - word ordering
3	The Mosque (المسجد)	Applying word writing	Mosque vocabulary - synonyms	Limādhā (why) - exclamation form - past and present tense verbs	Vocabulary writing from images - image description
4	The Library المكتبة	Applying word writing	Library vocabulary - antonyms	Iḍāfa (genitive) construction - hal (yes/no) questions	Word ordering - sentence formation
5	Education & Learning التربية والتعليم	Applying sentence writing	Text reading - comprehension of text meanings	Verb + an + imperfect verb	Using structures - guided description
6	Islamic Boarding School المعهد	Applying sentence writing	Reading text about pesantren - text meanings	Lam + imperfect verb - prepositions	Sentence formation - writing daily activities
7	Occasions (المناسبات)	Applying sentence writing	Reading text about occasions - text meanings	Imperfect verb with plural wāw	Word ordering - writing activity schedules
8	Islamic Scholars in Indonesia (العلماء في إندونيسيا)	Applying sentence writing about scholars	Education vocabulary - vocabulary classification	Fī ayyi (in which) - kāna + its subject + its predicate	Word ordering - biography writing

This thematic selection is not arbitrary; it is based on the principle of relevance and meaning, which makes learners more motivated when they find a topic closely related to their personal experiences and cultural environment (Biber, Conrad, & Reppen, 1998)..

Regarding vocabulary, the researcher adopted a methodology based on selection and gradation, in which key vocabulary was extracted from 'Al-Arabiyya lil-Hayat 2' and Sketch Engine tools were then employed to uncover their authentic contexts in the 'Indonesia Al-Youm' corpus. This approach aligns with what applied linguistics researchers have confirmed: teaching vocabulary in its natural context is far more effective than learning it in isolation in abstract lists (Hamid et al., 2020).

Grammatical structures are distributed progressively across eight units, starting from simple structures such as temporal adverbs and tens numbers in Unit 1, and escalating toward more complex structures such as the 'kāna' construction with its

subject and predicate in Unit 8. This grammatical progression respects the principle of Gradual Complexity, whose effectiveness in second language acquisition has been established by educational studies (Racine, 2018). What distinguishes this curriculum in handling grammatical structures is that it does not merely present them as abstract theoretical rules, but links them to examples extracted from the 'Indonesia Al-Youm' corpus. Thus, a student studying the structure 'verb + an + imperfect verb' does not simply read it in a grammar book but encounters it employed in real journalistic sentences that reflect the contemporary use of Arabic in the Indonesian context (Hamid et al., 2020).

Writing exercises embedded in this curriculum represent the cornerstone of building students' writing competence. These exercises were designed according to a progressive sequence from lower to higher levels: beginning with word ordering to form correct sentences, then moving to guided description, short report writing, comparison and contrast, and ending with expressing opinions. This progression reflects the 'Guided Composition' model, one of the most effective approaches in writing instruction (Sinclair, 1991).

The 'report writing' exercise represents a qualitative leap toward guided creative writing, where students are required to write a short journalistic report (150–200 words) on a topic from the eight units, employing the acquired vocabulary and structures. This type of writing mimics the style of original texts in the 'Indonesia Al-Youm' corpus, making students feel that they are producing real language rather than artificial classroom texts (Rowley-Jolivet, 2002).. The 'expressing opinions' exercise represents the pinnacle of the writing exercise hierarchy, where students are required to use linguistic evaluation structures—such as 'yanbaghi' (should), 'lā budda min' (it is necessary), and 'yu'addu' (is considered)—originally extracted from the 'Indonesia Al-Youm' corpus. This embodies the fundamental goal of the curriculum: enabling students to express themselves in writing with eloquent and cohesive Arabic inspired by authentic language data (Rowley-Jolivet, 2002).

All units share the development of writing skills through word ordering, sentence formation, data completion, image description, and guided expression. The researcher used three Sketch Engine features to extract vocabulary and structures: Word Sketch for analyzing collocational relationships, Thesaurus for exploring contextual synonyms, and Concordance for presenting examples in their authentic context.

B. Sketch Engine as an Analytical Tool

Sketch Engine is a specialized computational linguistics software developed by Lexical Computing Ltd., distinguished by its ability to process millions of words in seconds. The tools utilized in this model include: (1) Word Lists, to identify the most frequently occurring vocabulary in the corpus; (2) Concordance, to display every occurrence of a word in its right and left context; (3) Collocation, to reveal statistically associated words; (4) Thesaurus, to explore words with similar contexts; and (5) Word Sketch, to display a summary of a word's grammatical and semantic behavior (Kilgarriff, Baisa, Bušta, Jakubíček, et al., 2014).

Employing Sketch Engine in corpus analysis represents a qualitative addition to the design of this curriculum. This specialized software enables both researchers and teachers to explore the language corpus with unprecedented depth, uncovering patterns that are difficult to identify through manual inspection alone.

C. Worksheet as a Space for Application

The worksheet represents the applied teaching source that translates theoretical knowledge into practical action. It comprises six progressive sections: (1) 'Vocabulary in Context,' presenting key vocabulary with their multiple translations; (2) 'Vocabulary in Collocations,' revealing collocational patterns extracted from the corpus; (3) 'Vocabulary in Contextual Agreement,' presenting words in their contexts accompanied by grammatical analysis; (4) 'Vocabulary in Synonyms,' exploring semantic relationships between similar words; (5) 'Grammatical Structures,' providing practice on the target structure; and (6) 'Guided Writing,' which represents the pinnacle of the instructional sequence.

The worksheet concludes with a 'Vocabulary List' table as a quick reference, facilitating students' recall of studied words during writing without requiring a full review of previous sections. Students are strongly encouraged to return to the language corpus to verify the correctness of collocations and structures they use, reinforcing the habit of relying on reliable linguistic evidence rather than intuition alone. Educational research has demonstrated that having easily accessible linguistic references during writing reduces errors and raises the quality of the produced text (Khalilullah, 2012).

The worksheet with its six sections embodies an integrated instructional model that achieves natural progression from reception to production, from guided learning to research independence, and from theoretical knowledge to creative application. Its greatest added value lies in its reliance on authentic examples extracted from the 'Indonesia Al-Youm' corpus, making it a living teaching tool rooted in actual language

use rather than artificial examples disconnected from linguistic reality. In this way, the worksheet achieves all five of its essential goals: integration, progression, linguistic authenticity, independence, and self-assessment.

Expert Review Results for Learning Design

Dr. Muhammad Ahsan al-Din, M.A. conducted a comprehensive review of the learning design using an evaluation instrument comprising twenty items distributed across four axes, using a modified four-point rating scale. The following table presents the evaluation results:

Table 3. Expert Review Scores for Learning Design by Axis

Axis	Category	Mean	No. of Items	Score 4	Score 3	Score 2
A. Instructional Design	Suff. Appropriate	2.60	5	0	3	2
B. Media and Appearance	Appropriate	2.83	6	1	3	2
C. Interactivity & Ease of Use	Appropriate	3.20	5	1	4	0
D. Assessment and Feedback	Suff. Appropriate	2.00	4	0	0	4
Total / Mean	Appropriate	2.70	20	2	10	8

As shown in Table 3, the overall mean score reached 2.70, falling within the 'Appropriate' category. No item received a score of 1 (Inappropriate). The 'Interactivity and Ease of Use' axis recorded the highest mean (3.20 - Appropriate), while the 'Assessment and Feedback' axis recorded the lowest mean (2.00 - Sufficiently Appropriate).

The expert provided five developmental recommendations: (1) revising the writing style and typographical consistency; (2) including references used in the model text; (3) completing the cover page with the university logo, developer's name, and supervisor's name; (4) adding a design usage guide; and (5) providing a material evaluation instrument. He concluded that 'this learning design is appropriate for field use with the application of the noted revisions.'

Learning Steps

This model implements the Data-Driven Learning approach, where corpora serve as the primary resource to trigger inductive learning through the exploration of authentic linguistic data (Pérez-Paredes et al., 2018). The following is the discussion of the instructional steps within the framework, translated into English and supported by

academic references to strengthen the theoretical foundation of the model.

Step 1 – Concept Presentation

In the initial stage, the teacher acts as a facilitator bridging the students' empirical experiences with new material. The use of live screenshots from Sketch Engine provides concrete visualizations of word frequency and real-world usage contexts. This aligns with DDL principles that emphasize the importance of exposure to rich, up-to-date linguistic data to build language awareness from the outset (Bao, 2021). Visualizations such as "Word Sketches" help students instantly see grammatical patterns and collocations before they begin their own experimentation (Foll, 2021b).

Step 2 – Guided Corpus Exploration

This step marks a shift in the student's role from a passive recipient to a language researcher (Bao, 2021). By entering keywords into Sketch Engine, students engage in a self-discovery process to observe grammatical functions and collocations. The "Word Sketch" feature is particularly effective here as it automatically categorizes collocations based on their grammatical relations, making it easier for students to understand how a word interacts with others within correct sentence structures (Foll, 2021b).

Step 3 – Analytical Exercises

The worksheet exercises (sections 1–5) are designed to reinforce "noticing" skills regarding the language patterns discovered. Focusing on synonyms, collocational patterns, and contextual agreement helps students build deep lexicogrammatical knowledge. DDL specifically encourages learning efficiency by allowing students to discover patterns for themselves that are typically presented as "ready-made" rules in traditional textbooks (Bao, 2021).

Step 4 – Guided Written Production

During the production phase, students apply their findings into a complete text following strict quantitative instructions. The requirement to use at least four vocabulary items and two collocations from the corpus ensures that students do not rely solely on intuition, but actively utilize the natural expressions found in authentic data (Foll, 2021b). This approach helps minimize unnatural language use and improves stylistic accuracy according to the target genre (D. Wong, 2018).

Step 5 – Reflection and Feedback

The learning cycle concludes with self-assessment and peer evaluation. This process is vital for developing learner autonomy. By evaluating texts based on corpus-derived criteria, students develop a better understanding of linguistic structure and style

(K. M. Wong, 2021). Teacher feedback that includes direct examples from the corpus provides data-based evidence that can strengthen the students' pedagogical confidence in the material they have learned (Sert & Aşık, 2018).

Learning Methods

The discussion of the learning methods employed in this model demonstrates a harmonious integration of student autonomy and structured instructional support. Data-Driven Learning positions students at the center as active language researchers who explore authentic data (Crosthwaite et al., 2019). Through concordance lists, students do not merely memorize rules; instead, they inductively identify lexicogrammatical patterns emerging from real-world language use (Vyatkina & Boulton, 2017). This approach allows students to discover linguistic regularities independently, which in turn enhances their language awareness and critical thinking skills when processing linguistic information (Boulton, 2010).

Inquiry-based learning is integrated through the use of intentionally blank tables in the worksheet designed to train independent search skills. This method compels students to conduct authentic investigations and in-depth data exploration, aligning closely with the core elements of language discovery strategies (Crosthwaite et al., 2019). By completing these tables independently, students develop corpus consultation literacy, helping them transition from passive recipients of information to strategic language users (O'Sullivan, 2007).

Instructional support, or scaffolding, is applied systematically to guide students from simple tasks toward complex text production. This process begins with word-ordering exercises and progresses gradually to the writing of a complete journalistic report, ensuring that the students' cognitive load remains manageable (J. Liu & Ma, 2025). The provision of scaffolding prompts during corpus searches has been shown to significantly improve student performance and confidence in self-correction and accurate text composition (Chang & Sun, 2009).

Contextual learning ensures that all materials and examples have real-world relevance through the use of the 'Indonesia AI-Youm' corpus. Utilizing a domain-specific corpus allows students to access knowledge that matches their future communicative needs (Crosthwaite et al., 2019). By drawing all data from authentic texts, the model guarantees that the language learned is functional and meaningful within real-life contexts (Foll, 2021a).

Assessment System

The assessment system in the proposed model is designed integratively to support the development of student writing skills through three primary stages. First, formative assessment plays a vital role in monitoring student performance in real-time. The use of worksheets equipped with analytical tables allows instructors to monitor students' thought processes and progress directly (Agovic, 2025). This approach aligns with research indicating that continuous monitoring and the provision of feedback during the writing process (pre-writing and while-writing) significantly improve writing quality compared to evaluating only the final product (Graham, Hébert, & Harris, 2015)

Second, self-assessment is integrated to enable students to measure their own texts against transparent quantitative criteria (Balloo, Evans, Hughes, Zhu, & Winstone, 2018). Using clear criteria helps students develop metacognitive awareness and self-regulation, which in turn enhances their accuracy in detecting errors in their own drafts (Panadero et al., 2022). Studies show that when students perform self-assessment using rubrics or data-based criteria, they tend to achieve a higher level of accuracy in meeting established academic standards (Carroll, 2020).

Finally, Product assessment is conducted using a comprehensive rubric covering four key dimensions: accuracy of targeted vocabulary, naturalness of collocations and contexts, correctness of grammatical structures, and text coherence and compliance with instructions. The dimensions of vocabulary accuracy and natural collocations are crucial, as precise word choice and the use of conventional expressions (collocations) determine the richness of expression in academic text (Jalambo, Wahab, Ismail, & Belton, 2025). Meanwhile, grammatical correctness and coherence ensure that ideas are conveyed logically and flow smoothly, ensuring the communicative purpose of the task is achieved according to the parameters specified in the instructions (Naumoska-Sarakinska, 2023).

Distinctive Features of the Proposed Model

The proposed model is distinguished by five fundamental features that set it apart from traditional instructional models: (1) Complete linguistic authenticity – all examples, sentences, and texts are based on the 'Indonesia Al-Youm' corpus rather than artificial constructions; (2) A practically applied DDL approach – students open the concordance list feature and discover patterns independently; (3) Source integration – examples in the presentation, exploration, and worksheet are drawn from a single corpus; (4) A research-oriented worksheet – with intentionally open tables that compel active inquiry; and (5) Corpus-based quantitative self-assessment criteria.

The proposed model is distinguished by its commitment to complete linguistic authenticity, utilizing the 'Indonesia Al-Youm' corpus instead of artificial constructions. This exposure to authentic linguistic corpora is essential for improving writing accuracy and learner autonomy, as it provides students with real-world language patterns rather than prescriptive, textbook-style examples (Luo, 2025). By integrating a practically applied Data-Driven Learning approach, students are encouraged to use concordance features to independently discover these patterns (X. Zhang & Zeng, 2025). Research indicates that such inductive learning strategies—where students perceive similarities and differences through data—are more engaging and effective for language acquisition than traditional teacher-led explanations (Cheng & Warren, 2007).

The model further ensures consistency through source integration, where all examples across the presentation, exploration, and worksheets are drawn from a single, unified corpus (D. Liu, Deng, & Wu, 2024). This cohesive structure is paired with a research-oriented worksheet featuring intentionally open tables that compel active inquiry. Rather than simply completing drills, students participate in discovery learning, building new knowledge as they navigate the corpus at their own pace. These classroom activities provide a meaningful context for active language discovery, transforming the writing process into a form of linguistic research where students formulate and test their own hypotheses (Cheng & Warren, 2007).

The inclusion of corpus-based quantitative self-assessment criteria empowers students to measure their progress with precision. Using corpus linguistic tools, learners can assess their own writing at the lexical, syntactic, and discourse levels, which promotes a deeper understanding of the writing process (Adorján, 2022). This data-driven approach significantly improves the accuracy of self-assessment and facilitates more strategic revisions, as students can compare their drafts against established academic standards (Qiu, 2024). By shifting from subjective evaluation to evidence-based self-regulation, the model cultivates higher levels of metacognitive awareness and writing resilience (Huang, 2025).

CONCLUSION

This study has demonstrated that employing the 'Indonesia Al-Youm' corpus through the Sketch Engine platform enables the construction of a proposed instructional model for teaching Arabic writing skills, based on three integrated sources: PowerPoint presentations, corpus analysis tools, and worksheets. Conceptually, the model offers a

structured operationalization of Data-Driven Learning principles within a locally relevant corpus, thereby contributing to the theoretical development of corpus-based pedagogy in Arabic language teaching. For classroom practices, the model provides a sequential five-step framework that potentially bridges the gap between declarative knowledge and procedural skills in writing, while fostering learner autonomy and critical linguistic inquiry. However, these contributions remain conceptual, as the study did not implement the model with actual learners. Several limitations must be acknowledged: the model has not undergone field testing, the validation involved only two experts, and no empirical data on student writing outcomes were collected. Therefore, the suitability confirmed by expert reviews (mean scores of 4.32 for teaching materials and 2.70 for learning design) should be interpreted as a formative endorsement of the model's design quality, not as evidence of its instructional effectiveness. Future research should prioritize quasi-experimental designs with pre-tests and post-tests to measure the model's impact on writing accuracy, fluency, and complexity. Further developmental research is also needed to refine the model based on classroom implementation feedback, expand the corpus database to include diverse media genres, and train teachers in using Sketch Engine for designing authentic materials.

REFERENCES

- Adorján, M. (2022). Self-assessment of academic writing in an online learning environment. *Porta Lingua*, (1), 115–121. <https://doi.org/10.48040/pl.2022.1.11>
- Agovic, M. (2025). Active Learning as an Element of Formative Assessment in the Undergraduate Biology Classroom. *The American Biology Teacher*, 87(9), 539–542. <https://doi.org/10.1525/abt.2025.87.9.539>
- Alhawary, M. T. (2009). *Arabic Second Language Acquisition of Morphosyntax*. New Haven: Yale University Press.
- Ashfia, A., Mu'awanah, E., Nur, A., & Susiawati, W. (2023). Sketch Engine Implementation in Arabic Vocabulary Learning. *Ukazh: Journal of Arabic Studies*, 4(2), 123–135.
- Baloo, K., Evans, C., Hughes, A., Zhu, X., & Winstone, N. (2018). Transparency Isn't Spoon-Feeding: How a Transformative Approach to the Use of Explicit Assessment Criteria Can Support Student Self-Regulation. *Frontiers in Education*, 3. <https://doi.org/10.3389/feduc.2018.00069>
- Bao, K. (2021). Should We Use It in Our Classrooms: An Analysis of Data-Driven Learning Research. *English Linguistics Research*, 10(3), 66. <https://doi.org/10.5430/elr.v10n3p66>
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge University Press.
- Boulton, A. (2010). Data-Driven Learning: Taking the Computer Out of the Equation. *Language Learning*, 60(3), 534–572. <https://doi.org/https://doi.org/10.1111/j.1467-9922.2010.00566.x>
- Boulton, A., & Cobb, T. (2017). Corpus Use in Language Learning: A Meta-Analysis. *Language Learning*, 67(2), 348–393.
- Burhani, N., & Al-Khawaldeh, M. A. (2026). Writing Anxiety Among Arabic Language Students at Islamic Universities in Indonesia. *Dirasat: Human and Social Sciences*, 53(6), 9647. <https://doi.org/10.35516/Hum.2026.9647>
- Carroll, D. (2020). Observations of student accuracy in criteria-based self-assessment. *Assessment & Evaluation in Higher Education*, 45(8), 1088–1105. <https://doi.org/10.1080/02602938.2020.1727411>
- Chang, W.-L., & Sun, Y. (2009). Scaffolding and Web Concordancers as Support for Language Learning. *Computer Assisted Language Learning*, 22(4), 283–302. <https://doi.org/10.1080/09588220903184518>
- Cheng, W., & Warren, M. (2007). Checking Understandings: Comparing Textbooks and a Corpus of Spoken English in Hong Kong. *Language Awareness*, 16(3), 190–207. <https://doi.org/10.2167/la455.0>
- Crosthwaite, P. (2019). Characterising postgraduate students' corpus query and usage patterns for disciplinary data-driven learning. *ReCALL*, 31(3), 255–275. <https://doi.org/10.1017/S0958344019000077>
- Crosthwaite, P., Pérez-Paredes, P., Boulton, A., Foll, E. Le, Pascual, P. P., María, S. T., ... Sun, Y. (2019). Thirty years of data-driven learning: Taking stock and charting new directions over time. *ReCALL*, 31(3), 269–286. <https://doi.org/10.53308/ide.v6i2.61>
- Dalglish, S. L., Khalid, H., & McMahan, S. A. (2020). Document analysis in health policy research: the READ approach. *Health Policy and Planning*, 35(10), 1424–1431. <https://doi.org/10.1093/heapol/czaa064>
- Dörnyei, Z. (2007). *Research Methods in Applied Linguistics*. Oxford University Press.
- Emara, F. A., & Rouyan, N. M. (2025). A Descriptive Analysis of Written Linguistic Errors among Non-Native Learners of the Arabic Language. *Afaq Lughawiyyah*, 3(2), 32–51.

- Foll, E. Le. (2021a). *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom: A step-by-step guide for (trainee) teachers using online resources*. Zenodo. <https://doi.org/https://doi.org/10.5281/zenodo.4992504>
- Foll, E. Le. (2021b). *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom*. <https://doi.org/10.5281/zenodo.4992504>
- Fuadi, A. A., & Sani, S. (2025). Design of an Arabic Learner Text Corpus for STAI Ma'arif Kendal Ngawi Students to Support Data Driven Learning Material Development. *Al-Fatih: Jurnal Studi Islam*, 13(1), 68–79.
- Graham, S., Hébert, M., & Harris, K. R. (2015). Formative Assessment and Writing. *The Elementary School Journal*, 115(4), 523–547. <https://doi.org/10.1086/681947>
- Hamid, M. A., Sutaman, Syaiful Mustofa, Danial Hilmi, Zakia Arifa, Nur Qomari, ... Nur Ila Ifawati. (2020). *كتاب العربية للحياة ٢*. Malang: UIN Maliki Press.
- Huang, S. (2025). Research on English Writing Instruction Based on Data-Driven Learning (DDL). *Journal of International Education and Development*, 9(6), 120–123. <https://doi.org/10.47297/wspiedwsp2516-250026.20250906>
- Jalambo, M. O., Wahab, M. O. A., Ismail, S. M., & Belton, B. (2025). “Assess, adapt, achieve”: investigating the role of mobile-mediated dynamic assessment in cultivating L2 learners’ writing resilience, strategic revision, and linguistic accuracy. *Language Testing in Asia*, 15(1). <https://doi.org/10.1186/s40468-025-00407-3>
- Johns, T. (1991). *Should You be Persuaded: Two Samples of Data-Driven Learning Materials* (Vol. 4). na.
- Kania, N., & Kusumah, Y. S. (2024). Constructing and providing content validity evidence through the Aiken’s V index. *REiD (Research and Evaluation in Education)*, 10(1). <https://doi.org/10.21831/reid.v10i1.12345>
- Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., ... Suchomel, V. (2014). The Sketch Engine: Ten Years on. *Lexicography. The Judicial English Eurolect*, 25(1), 7--36.
- Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., ... Suchomel, V. (2014). The Sketch Engine. *Lexicography*, 1(1), 7–36. <https://doi.org/10.1007/s40607-014-0009-9>
- Laufer, B., & Waldman, T. (2011). Verb-Noun Collocations in Second Language Writing: A Corpus Analysis of Learners’ English. *Language Learning*, 61(2), 647–672.
- Liu, D., Deng, Y., & Wu, S. (2024). Using Corpora for Teaching Grammar (pp. 268–296). <https://doi.org/10.4324/9781032650883-14>
- Liu, J., & Ma, Q. (2025). Using a custom GPT to scaffold corpus-based vocabulary learning for low-proficiency learners. *Proceedings of the International CALL Research Conference, 2025*, 77–83. <https://doi.org/10.29140/97817637116240-11>
- Luo, H. (2025). Data-Driven Learning in Second Language Writing: A Systematic Review of Efficacy, Challenges, and Future Directions. *International Journal of Education and Humanities*, 19(2), 159–165. <https://doi.org/10.54097/t1wp5032>
- Lusta, A., Demirel, Ö., & Mohammadzadeh, B. (2023). Language corpus and data driven learning (DDL) in language classrooms: A systematic review. *Heliyon*, 9(12), e22731. <https://doi.org/10.1016/j.heliyon.2023.e22731>
- Maachi, H., Khamar, H., & Abubkr, L. O. (2025). Using a Pedagogic Corpus and Sketch Engine Tools to Analyze a Literary Text in an Arabic Language Classroom. *International Journal of Arabic-English Studies*, 25(2). <https://doi.org/10.33806/ijaes.v25i2.941>
- Mardapi, D. (2018). *Pengukuran, Penilaian, dan Evaluasi Pendidikan*. Parama Publishing.
- Naumoska-Sarakinska, B. (2023). ESTABLISHING AND MAINTAINING OBJECTIVE CRITERIA IN WRITING AND SPEAKING FEEDBACK STRATEGIES IN EFL.

- Journal of Teaching English for Specific and Academic Purposes*, 233. <https://doi.org/10.22190/jtesap230126016n>
- Nurhayati, N., Wahyudi, W., & Alsagaf, S. L. H. (2021). Development of HOTS Assessment Instruments. *Jurnal Pembelajaran Fisika*, 9(2), 45–56.
- O’Sullivan, Í. (2007). Enhancing a process-oriented approach to literacy and language learning: The role of corpus consultation literacy. *ReCALL*, 19(3), 269–286. <https://doi.org/10.1017/s095834400700033x>
- Olson, K. (2010). An Examination of Questionnaire Evaluation by Expert Reviewers. *Field Methods*, 22(4), 295–318. <https://doi.org/10.1177/1525822X10379795>
- Panadero, E., García-Pérez, D., Ruiz, J. F., Fraile, J., Sánchez-Iglesias, I., & Brown, G. (2022). University students’ strategies and criteria during self-assessment: instructor’s feedback, rubrics, and year level effects. *European Journal of Psychology of Education*, 38(3), 1031–1051. <https://doi.org/10.1007/s10212-022-00639-4>
- Pascual, P. P., María, S. T., M., A. C. J., Boulton, A., Vyatkina, N., Pérez-Paredes, P., ... Foll, E. Le. (2019). Thirty years of data-driven learning: Taking stock and charting new directions over time. *ReCALL*, 31(3), 269–286. <https://doi.org/10.53308/ide.v6i2.61>
- Pérez-Paredes, P. (2020). *Corpus linguistics for education: A guide for research*. Routledge.
- Pérez-Paredes, P., Boulton, A., Bao, K., Foll, E. Le, Frankenberg-Garcia, A., Lew, R., ... Song, Y. (2018). Data-driven Learning in and out of the Language Classroom. *ReCALL*, 31(1), 55–78. <https://doi.org/10.1017/9781009511384>
- Plomp, T., & Nieveen, N. (2013). *Educational Design Research*. SLO.
- Qiu, X. (2024). Exploring the Effect of Corpus-Based Writing Instruction on Learner-Corpus Interaction in L2 Revision: A Study of Chinese EFL Disciplinary Writers. *TESOL Quarterly*, 58(3), 1108–1137. <https://doi.org/10.1002/tesq.3308>
- Racine, J. P. (2018). Lexical Approach. *The TESOL Encyclopedia of English Language Teaching*, 2, 1–7.
- Rapley, T., & Jenkins, K. N. (2010). Document Analysis. *International Encyclopedia of Education*. Elsevier.
- Rowley-Jolivet, E. (2002). Longman Grammar of Spoken and Written English. Harlow: Pearson Education Limited, 1999. *Les Cahiers de l’APLIUT. Pédagogie et Recherche*, 21(3), 91–93.
- Sert, O., & Aşık, A. (2018). A corpus linguistic investigation into online peer feedback practices in CALL teacher education. *Applied Linguistics Review*, 11(1), 55–78. <https://doi.org/10.1515/applirev-2017-0054>
- Tomlinson, B. (2023). *Developing Materials for Language Teaching*. Bloomsbury Publishing.
- Vyatkina, N., & Boulton, A. (2017). Corpora in Language Teaching and Learning. *Language Learning and Technology*, 21(3), 66–89.
- Wong, D. (2018). A Corpus-Based Study of Peer Comments and Self-Reflections. *International Journal of Online Pedagogy and Course Design*, 8(4), 65–90. <https://doi.org/10.4018/ijopcd.2018100105>
- Wong, K. M. (2021). “ A design framework for enhancing engagement in student - centered learning : own it , learn it , and share it ” by Lee and Hannafin (2016): an international perspective. *Educational Technology Research and Development*, 69(1), 93–96. <https://doi.org/10.1007/s11423-020-09842-w>
- Yasmar, R. (2025). Analisis Faktor-Faktor Kesulitan Menulis Bahasa Arab pada Mahasiswa. *Jurnal Al-Fikrah*, 13(1), 45–60.
- Zaki, M. (2017). Corpus-Based Teaching in the Arabic Classroom: Theoretical and Practical Perspectives. *International Journal of Applied Linguistics*, 27(2), 514–541.
- Zhang, J., Yoon, C., Williams, D., & Pinkas, A. (2024). Exploring the Usage of ChatGPT

Among Medical Students in the United States. *Journal of Medical Education and Curricular Development*, 11. <https://doi.org/10.1177/23821205241264695>

Zhang, X., & Zeng, J. (2025). Using large language models to enhance data-driven learning in academic writing. *Proceedings of the International CALL Research Conference, 2025*, 197–202. <https://doi.org/10.29140/97817637116240-26>