Automated Generation of Arabic Verb Conjugations with Multilingual Urdu Translation: An NLP Approach

Haq Nawaz[1], Manal Elobaid[2], Ali al-Laith[3], Saif Ullah[4]
COMSATS University, Lahore, Pakistan[1]
Qatar University, Qatar[2]
Copenhagen University, Denmark[3]
ILI.digital, Lahore, Pakistan[4]

Abstract

This paper presents a rule-based automated system for generating both Arabic verb conjugations and their corresponding Urdu translations. The system processes triliteral, non-weak Arabic roots across key tenses Past Simple, Past Simple Negative, Present Simple, and Present Simple Negative. Addressing the challenges posed by Arabic morphology, our rule-based approach applies patterns and morphological rules to accurately produce verb conjugations, capturing essential grammatical variations in gender, number, and person. Simultaneously, the system generates Urdu translations using predefined patterns that is aligned with the grammatical nuances of Arabic, ensuring semantic consistency. As the first system of its kind, it uniquely provides a cross-lingual resource that bridges two linguistically similar but distinct languages. By focusing on rule based precision and dual-language outputs, it addresses critical gaps in NLP resources, serving as a valuable tool for linguists, educators, and NLP researchers in academic and religious contexts where Arabic and Urdu coexist.

1 Introduction

The Arabic language, deeply rooted in the Semitic language family, presents significant challenges in natural language processing (NLP) due to its intricate morphological structure, particularly in verb conjugations [1]. Arabic verbs are inflected for tense, voice, gender, number, and person, creating complex conjugation tables that reflect each verb's nuanced forms [2]. These conjugations involve the combination of triliteral roots, prefixes, and suffixes to indicate grammatical distinctions, which are further diversified across Classical Arabic, Modern Standard Arabic, and colloquial dialects. Such features, along with unique tenses and forms, make the automation of Arabic verb conjugation a challenging task for NLP applications [3].

In this paper, we address these challenges by developing a rule-based automation system for generating Arabic verb conjugations from triliteral, non-weak root words across four specific tense categories: Past Simple, Past Simple Negative, Present Simple, and Present Simple Negative and generating Urdu translations for these sentences. This approach leverages the systematic structure of Arabic morphology to handle common triliteral root patterns, or Baab patterns, ensuring accurate inflection across gender, number, and tense. Additionally, to support cross-lingual applications, each generated conjugated form is provided with an equivalent Urdu translation, allowing for broader accessibility and use in multilingual NLP settings. This work focuses on triliteral, non-weak roots in Classical Arabic, but future extensions could address weak roots, dialectal Arabic, and languages like Persian and Pashto by adapting morphological rules. This aligns with the growing need for computational resources that bridge less-resourced languages with modern NLP advancements, particularly in religious and educational contexts where both Arabic and Urdu are commonly used. The dataset is available in GitHub[1].

[1] https://github.com/haqnawaz99/Arabic-Urdu-Conjugation-Dataset

Processings of the 1st Workshop on NLP for Languages Using Arabic Script (AbjadNLP), pages 136–143

2 Contributions

Our contributions are as follows:

2.1 A Rule-Based Automation System for Arabic Verb Conjugation

Our system currently focuses on Arabic triliteral, non-weak root verbs for Past Simple, Past Simple Negative, Present Simple, and Present Simple Negative forms. Future work includes expanding coverage to weak roots and quadrilateral roots, requiring tailored morphological rules and pattern adaptations. The system applies Arabic morphological rules, addressing gender, number, and tense variations in accordance with different patterns. Our focus on non-weak roots allows us to refine conjugation accuracy and computational efficiency.

2.2 Integration of Urdu Translations for Conjugated Forms

Each Arabic verb form is paired with its corresponding Urdu translation, fostering cross-lingual applications and enabling Urdu speakers to leverage computational insights into Arabic verb morphology. This integration enhances accessibility and understanding, especially in contexts demanding a precise grasp of Arabic grammar.

2.3 Detailed Analysis of Accuracy and Impact

We conducted a detailed analysis of the system's accuracy, evaluating its effectiveness in correctly applying Arabic morphological rules. This assessment highlights the system's significant contributions to computational linguistics, particularly for less-resourced languages such as Arabic and Urdu. Our work involved 200 Arabic root words, generating 78 variations for each verb, resulting in a comprehensive dataset of 14,400 entries, accompanied by their Urdu translations. The dataset underwent meticulous review by a team of religious scholars of Jamia Ashrafia[2] Lahore Pakistan, with deep expertise in Arabic morphology and the Urdu language. This rigorous approach underscores the system's ability to handle specific verb forms, achieving high linguistic fidelity in the generated outputs and ensuring precision and relevance across both languages.

3 Background and Related Work

Research in Arabic natural language processing (NLP) has advanced significantly over recent years, driven by the need to develop robust computational tools for processing Arabic's complex morphological structure [4]. Arabic morphology, particularly verb conjugation, poses unique challenges due to the language's rich inflectional system, including variations across tense, gender, number, and voice [2].

However, while Arabic NLP has made strides in verb root extraction and morphological analysis, little attention has been given to the integration of Urdu translations. Urdu, despite its shared script and vocabulary with Arabic, lacks the level of computational resources available for Arabic NLP. Prior work in Arabic NLP has primarily focused on Arabic syntax and morphology within Arabic contexts [5]. There is very limited exploration into cross-linguistic applications or extensions for Urdu language for translation purposes. Our system bridges grammatical alignment between Arabic and Urdu, addressing the lack of bilingual resources for these languages. This gap underscores the need for focused efforts to bridge Arabic NLP methodologies with Urdu linguistic resources to enhance their mutual computational potential. Another research introduced a novel approach to help Arabic learners understand and memorize the meanings derived from morphological changes (wazn al-ṣarf), which often lead to translation errors. By identifying key morphological constructions and their associated meanings, the study represents them in a didactic poetic form (naẓm) using the Rajz prosodic structure. An Android application was developed to deliver the naẓm, comprising 9 chapters and 32 verses, enhancing accessibility and usability. Validation tests rated the application highly (87.61%), confirming its effectiveness. This tool offers an innovative solution for teachers and learners to master Arabic morphology more efficiently.

[2] https://jamiaashrafia.org/

4 Methodology

4.1 Data source:

Our system relies on a dataset of Arabic triliteral roots, we selected the most used trilateral root words used in Quran and Classical Arabic literature. These roots serve as the foundation for generating full verb conjugation tables in Arabic. To ensure linguistic diversity, the roots were carefully selected from Quranic and classical Arabic sources, representing a range of grammatical contexts, including variations in diacritics and negations. To ensure the validity and linguistic accuracy of these roots, we utilize well-established Arabic linguistic resources, including the Quranic Arabic Corpus, which provides comprehensive morphological data and annotations specifically for Classical Arabic.

4.2 Conjugation Generation Process:

Our approach involves a systematic rule-based generation of Arabic verb conjugations from triliteral, non-weak root words across selected tenses: Past Simple, Past Simple Negative, Present Simple, and Present Simple Negative. The generation of conjugated forms begins with the application of Arabic morphological rules that define the specific pattern each root follows. By employing this structure, our system generates various verb forms by appending the correct prefixes and suffixes based on tense, person, gender, and number. For instance, given the root سمع meaning “to hear”, the system produces conjugations like سَمِعَ (he heard) in Past Simple and لَا يَسْمَعُ (he does not hear) in Present Simple Negative. This automated approach ensures that all conjugated forms are syntactically accurate and align with the morphological standards of Classical Arabic.

4.3 Urdu Translation Generation:

Complementing the Arabic conjugation system, we developed an Urdu translation module to provide accurate translations for each conjugated form. For every generated Arabic conjugation, the system maps it to a corresponding Urdu phrase by using predefined translation patterns that respect Urdu grammar rules. These patterns ensure that each translated verb accurately reflects the gender, tense, and grammatical number of the Arabic original. Given the linguistic similarities between Arabic and Urdu, particularly their shared root system, this module allows us to generate semantically precise Urdu translations that mirror the Arabic verb’s grammatical structure. The Urdu translation module thereby extends the utility of our system, enabling seamless cross-lingual understanding of Arabic verbs in Urdu, a feature valuable for religious, academic, and educational contexts

4.4 Importance of Diacritics in Classical Arabic:

Diacritics in Arabic play a pivotal role in conveying meaning and grammatical context, particularly in verb conjugation [6, 7]. The root ن ص ر, combined with three فتحہ (fatha) diacritics, transforms into نَصَرَ, which signifies "he helped." In Urdu, this translates to "اس ایک مرد نے مدد کی." To create the dual form in Arabic, an additional "ا" is appended, resulting in نَصَرَا, meaning "they two helped." In Urdu, however, the corresponding translation becomes "ان دو مردوں نے مدد کی," requiring structural adjustments in sentence formation. This example highlights the linguistic transformations necessary when moving between Arabic and Urdu. While Arabic utilizes diacritics and morphological changes to convey grammatical details, Urdu requires explicit word additions and reordering to preserve the intended meaning. Incorporating these linguistic nuances, in our system we ensure accurate cross-lingual conjugation and translation, preserving the grammatical integrity of both languages.

5 Model Lay Out for Conjugation generation

The general layout of our model is structured as follows:

5.1 Lexeme Parsing:

Each Arabic root word undergoes a detailed parsing process to break it down into a lexeme structure, which serves as a framework for identifying its grammatical features, such as tense, gender, and number. This process involves analyzing the root and comparing it against predefined patterns or regular expression-based templates. These templates are meticulously designed to capture and define the allowable transformations for each specific conjugation form, ensuring that the resulting lexeme exactly follows to the grammatical rules of the language. This step is essential for accurately modeling the complex morphology of

Arabic, enabling precise identification and generation of all valid verb forms derived from a given root.

5.2 Pattern Matching and Transformation:

Morphological rules are applied based on the grammatical tags of the lexeme. At this stage, structured rules are used to add suffixes, prefixes, and make specific phonetic changes, ensuring that the word forms follow the established principles of Arabic morphology. These rules are carefully designed to align with the patterns and structures outlined in existing linguistic models, maintaining accuracy and consistency in the generation of word forms.

5.3 Validation:

To guarantee the accuracy of the generated conjugations, a rigorous validation process was carried out. This involved comparing the system-generated conjugations with verified patterns documented in authoritative reference corpora. Additionally, the dataset and outputs were thoroughly reviewed and annotated by a team of religious scholars from Jamia Ashrafia Lahore Pakistan, who possess deep expertise in Arabic morphology and linguistics. By aligning the outputs with established standards and leveraging expert validation, this process significantly reduced false positives and negatives. Furthermore, it refined the rules for root-word disambiguation, ensuring that each conjugated form was not only linguistically valid but also contextually appropriate, enhancing the overall reliability and precision of the system
Figures and tables

6 Results

Table 1 presents the conjugation results derived from the root word ع م س, showcasing the extensive variations generated using a rule-based algorithm. These conjugations demonstrate the application of Arabic morphological rules to produce forms that vary by number, gender, and person. The inclusion of detailed grammatical tags, such as singular, dual, plural (number), masculine and feminine (gender), and third-person, second-person, and first-person (person), highlights the system's capacity to systematically generate precise verb forms.

To enhance accessibility for Urdu language users and learners, grammatical attributes are provided in Urdu. This facilitates a better understanding of the conjugation patterns, bridging the gap between Arabic linguistic structures and Urdu-speaking learners. The detailed output underscores the effectiveness of the rule-based algorithm in capturing all permissible variations from a given root word, ensuring high accuracy and linguistic fidelity in the generated results.

Table 2 provides generated Urdu translations of the conjugated forms generated from the root word ع م س emphasizing the variations in number, gender, and person. These translations have been meticulously crafted to align with the corresponding Arabic conjugations, ensuring linguistic accuracy and contextual relevance. The translations include explicit annotations such as singular, dual, and plural (number), masculine and feminine (gender), and third-person, second person, and first-person (person), reflecting the comprehensive application of morphological rules.

Arabic	Number	Gender	Person
سَمِعَ	واحد	مذکر	غائب
سَمِعَا	تثنیہ	مذکر	غائب
سَمِعُوْا	جمع	مذکر	غائب
سَمِعَتْ	واحد	مونث	غائب
سَمِعَتَا	تثنیہ	مونث	غائب
سَمِعْنَ	جمع	مونث	غائب
سَمِعْتَ	واحد	مذکر	حاضر
سَمِعْتُمَا	تثنیہ	مذکر	حاضر
سَمِعْتُمْ	جمع	مذکر	حاضر
سَمِعْتِ	واحد	مونث	حاضر
سَمِعْتُمَا	تثنیہ	مونث	حاضر
سَمِعْتُنَّ	جمع	مونث	حاضر
سَمِعْتُ	واحد	مذکر	متکلم
سَمِعْنَا	تثنیہ	مذکر	متکلم
سَمِعْنَا	جمع	مذکر	متکلم
سَمِعْتُ	واحد	مونث	متکلم
سَمِعْنَا	تثنیہ	مونث	متکلم
سَمِعْنَا	جمع	مونث	متکلم

Table 1: Arabic Conjugations.

To cater specifically to Urdu-speaking users and learners, the details are presented in Urdu script with contextual examples for clarity. Each conjugated form is not only grammatically accurate but also contextually expressive, facilitating a deeper understanding of the relationship between Arabic and Urdu linguistic structures. This systematic presentation highlights the ability of the rule-based algorithm to generate accurate conjugations and demonstrates the effectiveness of the system in bridging Arabic morphology with its Urdu translations, fostering language learning and comprehension across both languages.

Urdu	Number	Gender	person
اس (ایک مرد) نے سنا	واحد	مذکر	غائب
ان (دو مردوں) نے سنا	تثنیہ	مذکر	غائب
ان (سب مردوں) نے سنا	جمع	مذکر	غائب
اس (ایک عورت) نے سنا	واحد	مونث	غائب
ان (دو عورتوں) نے سنا	تثنیہ	مونث	غائب
ان (سب عورتوں) نے سنا	جمع	مونث	غائب
آپ (ایک مرد) نے سنا	واحد	مذکر	حاضر
آپ (دو مردوں) نے سنا	تثنیہ	مذکر	حاضر
آپ(سب مردوں)نے سنا	جمع	مذکر	حاضر
آپ (ایک عورت) نے سنا	واحد	مونث	حاضر
آپ (دو عورتوں)نے سنا	تثنیہ	مونث	حاضر
آپ (سب عورتوں) نے سنا	جمع	مونث	حاضر
میں (ایک مرد) نے سنا	واحد	مذکر	متکلم
ہم (دو مردوں) نے سنا	تثنیہ	مذکر	متکلم
ہم (سب مردوں) نے سنا	جمع	مذکر	متکلم
میں (ایک عورت) نے سنا	واحد	مونث	متکلم
ہم (دو عورتوں) نے سنا	تثنیہ	مونث	متکلم
ہم (سب عورتوں) نے سنا	جمع	مونث	متکلم

Table 2: Urdu Translations.

7 Demonstration and Accessibility

To provide an interactive experience and allow users to explore the capabilities of our system, we have developed a live demonstration available at the following link: https://mhasham.pythonanywhere.com/. This demo enables users to input Arabic root words, view the generated conjugations, and access their Urdu translations.

The demo serves as a practical extension of the research, illustrating the system's functionality and accuracy in real-time. By making the system accessible online, we aim to support both researchers and learners in exploring Arabic morphological structures and their Urdu translations. The tool also allows users to compare translations with existing solutions like Google Translate, further highlighting the linguistic precision and contextual fidelity of our approach.

8 Comparison Analysis of Al-Tasreef Translations and Google Translate Outputs

To evaluate the accuracy and reliability of our system, we conducted a comparative analysis of conjugations derived from the Arabic verb forms with their corresponding Urdu translations generated by our rule-based Al-Tasreef system and those produced by Google Translate. This comparison highlights the linguistic fidelity of our approach, particularly in preserving the nuanced grammatical structures of Arabic, which include variations in gender, number, and person.

The Arabic conjugations in this study were systematically processed using our rule-based algorithm, which applies precise morphological rules to generate contextually accurate verb forms. These were translated into Urdu while maintaining grammatical integrity, ensuring that each translation aligns with the original meaning and context. The translations generated by Al-Tasreef were further validated by a team of religious scholars from Jamia Ashrafia, Lahore, known for their expertise in Arabic morphology and Urdu language.

In contrast, Google Translation output often displayed inaccuracies stemming from a lack of sensitivity to Arabic's complex morphological features, such as handling gendered plurals and negations. For example, while Google Translate failed to distinguish between masculine and

feminine plural forms, our system produced accurate and semantically aligned translations see Table 3. Issues included misinterpretation of negations, improper handling of gender-specific forms, and incorrect contextual mapping. These errors underscore the limitations of general-purpose translation systems when applied to morphologically rich languages like Arabic.

Arabic	Al Tasreef Translation	Google Translation
حَلَقَتْ	اس (ایک عورت) نے سر مونڈھا	یہ اڑ گیا۔
حَلَقْتَ	آپ ایک مرد نے سر مونڈھا	مونڈنا
تَحْلِقُ	آپ (ایک مرد) سر مونڈھتے ہو	پرواز
مَا قَطَعَتْ	اس (ایک عورت) نے نہیں کاٹا	اسے کاٹا نہیں گیا تھا۔
مَا قَطَعْنَا	ان (دو عورتوں) نے نہیں کاٹا	وہ ٹوٹے نہیں تھے۔
مَا قَطَعْتَ	آپ ایک مرد نے نہیں کاٹا	جو آپ نے کاٹ دیا۔
مَا قَطَعْتُنَّ	آپ (سب عورتوں) نے نہیں کاٹا	تم نے مجھے نہیں کاٹا
بَلَغْتُنَّ	آپ (سب عورتیں) پہنچیں	آپ اپنی عمر کو پہنچ چکے ہیں
بَلَغْتِ	آپ (ایک عورت) پہنچی	وہ پہنچ گیا
مَا خَرَجْتُمَا	آپ (دو مرد) نہیں نکلے	جب تم دونوں چلے گئے
مَا خَرَجْتُمْ	آپ (سب مرد) نہیں نکلے	جب تک تم چلے جاؤ
تَصْفَحُ	وہ (ایک عورت) درگزر کرتی ہے	براؤز کریں
مَا شَفَعْنَا	ہم (سب عورتوں) نے نہیں سفارش کی	ہم شفاعت نہیں کریں گے۔
لَا تَبْخَسُ	آپ (ایک مرد) نہیں کمی کرتے ہو	کم نہ سمجھیں

Table3: Translation Comparison.

Table 3 provides a detailed comparison, showcasing the Arabic conjugation, the corresponding Al-Tasreef translation, and the output from Google Translate. It highlights the instances where Google Translate diverges from the intended meaning, offering clear evidence of the strengths of our approach in achieving accurate and contextually appropriate translations.

9 Applications

The purposed system can be utilized in the following areas

Educational Tools: This system serves as an invaluable resource for students learning both Arabic and Urdu. By offering comprehensive conjugation tables alongside accurate translations, it provides learners with a clear understanding of the grammar, syntax, and structure of both languages simultaneously. This side-by-side approach not only simplifies language learning but also makes it more effective and engaging. Furthermore, the Arabic-Urdu conjugation generator has the potential to be utilized in the development of various educational tools and applications, supporting language learning in formal educational settings as well as self-study environments.

Machine Translation: By integrating Urdu translations for Arabic conjugations, this system lays the groundwork for developing machine translation tools. These tools can bridge the gap for low-resource languages, especially for language pairs like Arabic and Urdu that currently lack robust computational resource

Linguistic Research: This system provides a powerful tool for researchers in linguistics to explore and analyze the morphological similarities and differences between Arabic and Urdu. By systematically examining the grammatical structures, conjugation patterns, and linguistic nuances of both languages, the model enables a deeper understanding of their connections. Such analysis not only sheds light on the shared features and divergences between these languages but also offers valuable insights into the broader linguistic relationships between the Semitic language family, to which Arabic belongs, and the South Asian linguistic tradition, represented by Urdu. This cross-linguistic study significantly contributes to the field of comparative linguistics, paving the way for further research into how languages evolve, influence one another, and develop across different cultural and historical contexts.

Religious Studies: The interplay of Arabic and Urdu holds immense significance in religious contexts, especially for Islamic texts, where both languages serve as crucial mediums for understanding and interpretation. This system offers a precise and systematic tool for scholars, researchers, and readers by generating accurate conjugations and corresponding translations. By adhering closely to traditional linguistic rules and interpretations, the system ensures that the generated outputs remain faithful to the original meanings of sacred texts.

In addition, this resource is particularly valuable for institutions and scholars engaged in the study of Islamic texts, such as Religious Madaris across regions, as highlighted in the Pakistan Education Statistics (2021-22) [3] . These institutions, with significant enrollments and teaching faculties dedicated to Arabic and Urdu studies, can utilize this tool to streamline linguistic analysis and improve access to both languages' grammatical structures. The system's ability to bridge Arabic's intricate morphology with Urdu's expressive semantics facilitates a deeper understanding of classical religious literature, enhancing educational and theological research efforts across Pakistan and similar regions.

10 Conclusion

In this study, we present a rule-based system designed to generate Arabic verb conjugations and their corresponding Urdu translations. By addressing the complex morphological challenges of Arabic and aligning them with Urdu's linguistic structures, our approach makes a significant contribution to the fields of computational linguistics, language education, and translation for under-resourced languages. Validated by expert linguists and scholars, the system has proven highly accurate in producing linguistically and contextually appropriate outputs, positioning it as a valuable resource for educators, linguists, and researchers, particularly in academic and religious contexts.

A comparative analysis with Google Translate revealed the limitations of general-purpose machine translation systems in capturing the subtle grammatical nuances of Arabic and their precise translation into Urdu. In contrast, our system offers a domain-specific solution, tailored to the intricacies of Arabic morphology and its contextual relevance in Urdu. This makes it a robust tool for multilingual NLP applications, providing a more accurate and reliable translation experience.

11 Future Work

While the proposed system has shown promising results, several avenues for further development exist, which could enhance its functionality and broaden its impact.

One key area for future improvement is the expansion of the dataset and root coverage. Currently, the system focuses on triliteral, non-weak Arabic roots, but there is significant potential to include weak verbs and quadrilateral (four-letter) roots. This expansion would allow the system to handle a broader array of verb conjugations and translations, making it more versatile and capable of covering a larger portion of the Arabic verb system. By incorporating these additional roots, the system's utility would increase, benefiting both practical applications and theoretical linguistic analysis.

Another exciting direction is the integration with advanced machine learning models. While the current system is rule-based, incorporating transformer-based models such as AraBERT or multilingual BERT could significantly enhance its performance. These models are particularly powerful in their ability to generalize across unseen verb roots, improving the system's ability to generate accurate conjugations and translations. By integrating these models, the system could also refine the contextual accuracy of its translations, particularly for more complex sentence structures or less common verb forms.

Moreover, there is considerable potential in extending the system to support additional languages. While the current focus is on Arabic and Urdu, languages such as Persian or Pashto, which share similar linguistic characteristics, would benefit from the system's capabilities. Modifications would involve adapting morphological rules and translation patterns to accommodate the syntactic and semantic nuances of these languages. Adding support for these languages would significantly broaden the system's applicability and allow it to serve as a valuable tool in a wider range of linguistic contexts. This extension would also provide insights into the

[3] https://pie.gov.pk/

similarities and differences between these languages, offering a comparative perspective for linguists and researchers.

An important aspect of enhancing the system's usability is the development of an interactive user interface. Although the current demo offers basic functionality, a more interactive and user-friendly interface would greatly facilitate adoption among educators, students, and researchers. A well-designed interface would allow users to easily input verbs, view conjugations and translations, and explore the underlying rules and patterns. This would not only improve the system's accessibility but also make it more effective as a teaching tool in academic environments.

Finally, there is significant potential for the system's application in religious and cultural studies. Expanding the system to handle more complex Quranic or classical Arabic structures would be highly beneficial for scholars working with sacred texts. By accurately processing these more intricate forms of Arabic, the system could provide deeper linguistic and theological insights. This expansion would enable researchers to study sacred texts with greater precision, further enhancing the system's value in both religious and linguistic fields.

12 Acknowledgement

We extend our sincere gratitude to the scholars of Jamia Ashrafia, Lahore, for their invaluable feedback and meticulous evaluation of our work. Their expertise in Arabic morphology and Urdu linguistics significantly contributed to the precision and authenticity of this study.

References

J. Alasmari, J.C.E. Watson, and E. Atwell. *Investigating the Rate of Agree- ment and Disagreement of Tense and Aspect of Quranic Verbs in Arabic to English Translations: Experimental Results and Analysis,"* International Journal on Islamic Applications in Computer Science and Technology, vol. 6, no. 1, pp. 1-10, 2018.

M.T. Ben Othman, M.A. Al-Hagery, and Y.M. El Hashemi*, Arabic Text Processing Model: Verbs Roots and Conjugation Automation,* IEEE Access, vol. 8, pp. 103913–103923, 2020.

A.N. Alsaleh, E. Atwell, and A. Altahhan, *Quranic Verses Semantic Relat- edness Using AraBERT*, University of Leeds, 2021.

K. Shaalan, S. Siddiqui, and M. Alkhatib, *Challenges in Arabic Natural Language Processing,* in Computational Linguistics, Speech and Image Processing for Arabic Language, British University in Dubai, 2018.

Bashir, M.H., Azmi, A.M., Nawaz, H. *et al.* Arabic natural language processing for Qur'anic research: a systematic review. *Artif Intell Rev* **56**, 6801–6854 (2023). https://doi.org/10.1007/s10462-022-10313-2

Adany, Mohamed & Atwell, Eric. (2017). *Quran Question Answering System Using Arabic Number Patterns (Singular, Dual, Plural).* International Journal on Islamic Applications in Computer Science and Technology, Vol. 5, Issue 2, June 2017, 01-12. 5. 1-12.

E. L. M. Hassan*, The impact of standard Arabic verb phrase structure on Moroccan EFL learner's writing*, J. Humanities Social Sci., vol. 24, no. 1, pp. 6067, 2019.

Muh Syahri Romadhon, Moh Khasairi, and Achmad Tohe. 2024. *Android-based media development for memorizing Arabic verbal conjugation and its functional meaning*. Ijaz Arabi Journal of Arabic Learning, 7(2).