



## **QVR: Quranic Verses Recitation Recognition System using PocketSphinx**

**Hasan Ali Gamal Al-Kaf<sup>1</sup>, Muhammad Suhaimi Sulong<sup>2,3\*</sup>, Ariffuddin Joret<sup>4</sup>, Nuramin Fitri Aminuddin<sup>4</sup>, Che Adenan Mohammad<sup>2,5</sup>**

<sup>1</sup>Department of Computer and Information Sciences,  
Universiti Teknologi PETRONAS, Seri Iskandar, 32610, MALAYSIA

<sup>2</sup>Institute of Ahli Sunnah Wal Jamaah,  
Universiti Tun Hussein Onn Malaysia, Parit Raja, Batu Pahat, 86400, MALAYSIA

<sup>3</sup>Faculty of Technical and Vocational Education,  
Universiti Tun Hussein Onn Malaysia, Parit Raja, Batu Pahat, 86400, MALAYSIA

<sup>4</sup>Faculty of Electrical and Electronic Engineering,  
Universiti Tun Hussein Onn Malaysia, Parit Raja, Batu Pahat, 86400, MALAYSIA

<sup>5</sup>Centre for General Studies and Co-Curricular,  
Universiti Tun Hussein Onn Malaysia, Parit Raja, Batu Pahat, 86400, MALAYSIA

\*Corresponding Author

DOI: <https://doi.org/10.30880/jqsr.2021.02.02.004>

Received 20 September 2021; Accepted 28 October 2021; Available online 19 December 2021

**Abstract:** The recitation of Quran verses according to the actual *tajweed* is obligatory and it must be accurate and precise in pronunciation. Hence, it should always be reviewed by an expert on the recitation of the Quran. Through the latest technology, this recitation review can be implemented through an application system and it is most appropriate in this current Covid-19 pandemic situation where system application online is deemed to be developed. In this empirical study, a recognition system so-called the Quranic Verse Recitation Recognition (QVR) system using PocketSphinx to convert the Quranic verse from Arabic sound to Roman text and determine the accuracy of reciters, has been developed. The Graphical User Interface (GUI) of the system with a user-friendly environment was designed using Microsoft Visual Basic 6 in an Ubuntu platform. A verse of surah al-Ikhlâs has been chosen in this study and the data were collected by recording 855 audios as training data recorded by professional reciters. Another 105 audios were collected as testing data, to test the accuracy of the system. The results indicate that the system obtained a 100% accuracy with a 0.00% of word error rate (WER) for both training and testing data of the said audios via Quran Roman text. The system with automatic speech recognition (ASR) engine system demonstrates that it has been successfully designed and developed, and is significant to be extended further. Added, it will be improved with the addition of other Quran surahs.

**Keywords:** Quran verse, PocketSphinx, automatic speech recognition

## 1. Introduction

Reciting the Quran with proper *tajweed* is a practice for all Muslims around the world. Understanding and knowing the contents of the Quran is essential with authentication through direct links from original sources (Ta'a *et al.*, 2021). Comprehend its meaning, and having its thoughts, it will then make the reciter closer to its Creator. As stated in the Quran Surah al-Muzzammil 73, Verse 4 which carries meaning by Sahih International (Assami, Kennedy & Bantley, 1997):

(أَوْ زِدْ عَلَيْهِ وَرَتِّلِ الْقُرْآنَ تَرْتِيلًا )

(Or a little more; And recite the Quran in slow, measured rhythmic tones) (73: 4)

Thus, it is obligatory for them to learn the basic knowledge of *tajweed* in order to avoid mistakes when reciting which can lead to changes in pronunciation and meaning while both are big mistakes in reciting the Qur'an (Habash, 1986). Consequently, the reciter of the Qur'an needs an expert to check his recitation so that it follows the correct *tajweed*. According to Muhammad *et al.* (2011), this method is known as the method of talaqqi which is face-to-face recitation where it goes through the process of reciting, listening, revising, and repeating the correct *tajweed* of the Quran.

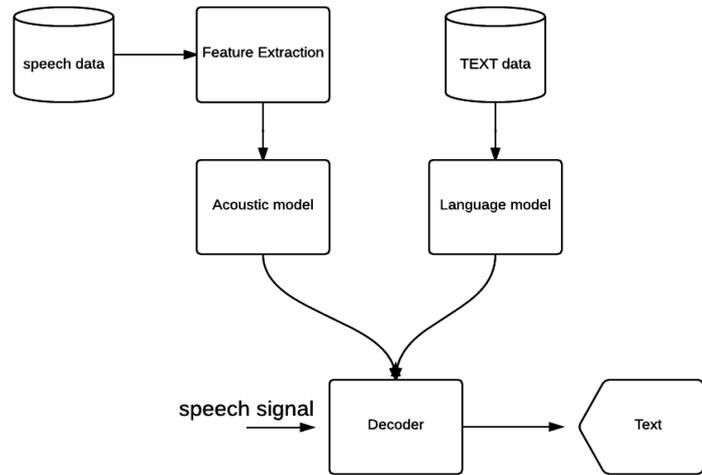
However, this current Covid-19 pandemic situation that hit the world has changed the way society lives now (Ciotti *et al.*, 2020), vastly the new norms terms is spread. Among them is the change in the method of implementation of programs and activities involving physical or face-to-face that inappropriate to carry out (Sundawa, Logayah & Hardiyanti, 2021) has shifted to virtual methods or better known as online (Ali, 2020). Therefore, the need for an online platform or system application is very necessary to meet the demand and implementation of these programs and activities, including education and learning systems (Mukhtar *et al.*, 2020). Today, technological advances are seen to be growing rapidly in the field of computer science with lots of computer-based systems developed via online (namely digital technology; Dwivedi, 2020; Ta'a *et al.*, 2021) which complement to the pandemic condition for education technology (Yousfi and Zeki, 2016; Schindler *et al.*, 2017). Therefore, the need of an accurate and efficient system that can verified Quran recitation is much required. Moreover, it is sometimes difficult for the reciters to set a suitable time to recite the Quran with the experts especially for the youth generation who are busy with their education and jobs.

The reciting and writing of the Quran have been carried out since the time of the Prophet Muhammad S.A.W. and it continues to this day (Kadir *et al.*, 2020). Recently, in the past 5 years, research on automatic speech recognition (ASR) has been at a high level of performance and standard (Ibrahim & Varol, 2020). The part of pattern recognition technology is the main research area that believes it can contribute towards the effectiveness of this study that focuses on speech recognition. Speech recognition is an important area in the processing of digital signals (Arora & Singh, 2012). The artificial intelligence of the machine or system is involved in the scope of the research area with this speech recognition, thus is able to hear and understand the spoken words of the recitation (Levis & Suvorov, 2012). Hence, in this study, a Quranic verse recitation recognition (QVR) system was developed as a practical method of recognizing Quranic words and put it into an actual software application that will be able to recognize some verse of part of the Quran and convert it to Roman text. The system was built using PocketSphinx (Daines *et al.*, 2006) to evaluate the accuracy of sentences and words error of the recitation.

## 2. QVR System Process

The Quranic verse recitation recognition system is able to convert a Quranic speech to Roman text. The developed system begins by retrieving audio speech as input data needed to convert it into text. This audio speech is then stored in a database. The databases required for the said system are text provision, speech corpus creation, transcription files, pronunciation dictionaries, language models, control files, and filler dictionaries. In order for the system to be function, the next step is by setting up the system environment such as software installation, training organization, project folders, audio files and other related items (Tabbal, El-Falou & Monla, 2006). Then, a training process is performed to train the system with the audio and data preparation. Later, decoding was performed to the trained data before obtaining system accuracy includes word errors and sentence errors. With this accuracy, API applications can be implemented to convert speech into text.

There are several different toolkits and approaches to implement speech recognition. One of the efficient toolkits is CMUSphinx. Quranic verse recitation recognition system was designed and developed using PocketSphinx – a part of the CMUSphinx toolkit for speech recognition, a well-known system recognizer (Hafeez, Mohiuddin & Ahmed, 2014). Fig. 1 shows the block diagram of a speech recognition system that has four main modules; they are feature extraction, acoustic model, language model, and decoder (Płonkowski & Urbanovich, 2018).

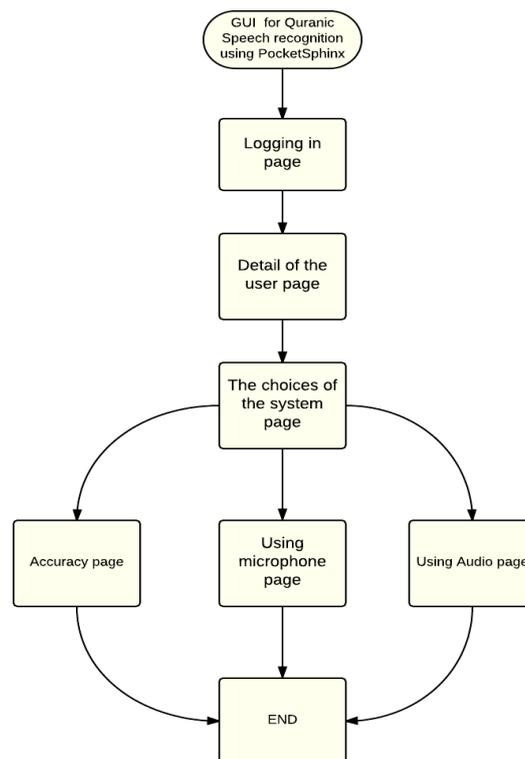


**Fig. 1 - Speech recognition block diagram**

Upon using the PocketSphinx speech recognition engine, the CMUSphinx toolkit runs on the Ubuntu platform under Linux operating system and uses Python programming language to run the scripts of the speech recognition module in order to convert speech to text. The chosen surah from the Quran verse was surah al-Ikhlās as speech data that recites by Quran reciters has been tested and applied in the QVR system. The speech data were collected by recording 855 audios as training data recorded by professional reciters. Another 105 audios were collected as testing data, to test the accuracy of the system. These speech data from both trained and untrained reciters were tested in two conditions; closed room and noisy space. Further, this study limits to some laws of *tajweed* (Habash, 1986) which are Adgham al-Ezhar, Muqalqal, Ash-Shamsi, al-Qamari, and Izhar Safari. According to Alfaries *et al.* (2013) study, the Quran text is important as the major reference of Muslim researchers where it requires to achieve 100% best accuracy following *tajweed* rules.

### 3. QVR System Development

The flow chart with several stages that illustrates the graphical user interface (GUI) of the developed system using Microsoft Visual Basic 6 is shown in Fig. 2.



**Fig. 2 - QVR GUI System**

The QVR system begins with a login page that allows the user to enter the system. It uses a simple and common security classification which are username and password. After the user is in, the system requires the user to fill in the user’s details such as name, age and etc. Then, the system displays the main page which consists of three (3) main choices either using a microphone, audio, or accuracy page for the user to select as shown in Fig. 3(a). In order to convert the speech to text, the system requires audio speech either from a microphone (real-time) or audio file (recorded). As for the real-time audio via a microphone, the system will store them in the database as previously mentioned and can be played again. It then can be displaying the results of the audio speech from the microphone. Fig. 3(b) shows the recording page using a microphone. Lastly, the system manages to present the accuracy percentage from the total percent of data correction and data errors.



Fig. 3 - GUI of the developed system (a) System main page; (b) Recording page using microphone

Fig. 4(a) illustrates the accuracy page of the system and Fig. 4(b) displays some lines of the accuracy page coding. In this window interface page, the user needs to record all the verses of surah al-Ikhlās, and then play it to make sure it has a clear speech audio. Then the system will process by pressing the “Run the Process” button to run the PocketSphinx that will give the accuracy of the recording and store it in the log file in sphinx train folder. Finally, once the “Result” button is pressed, the system will fetch the accuracy of the recording and display it on the terminal page.

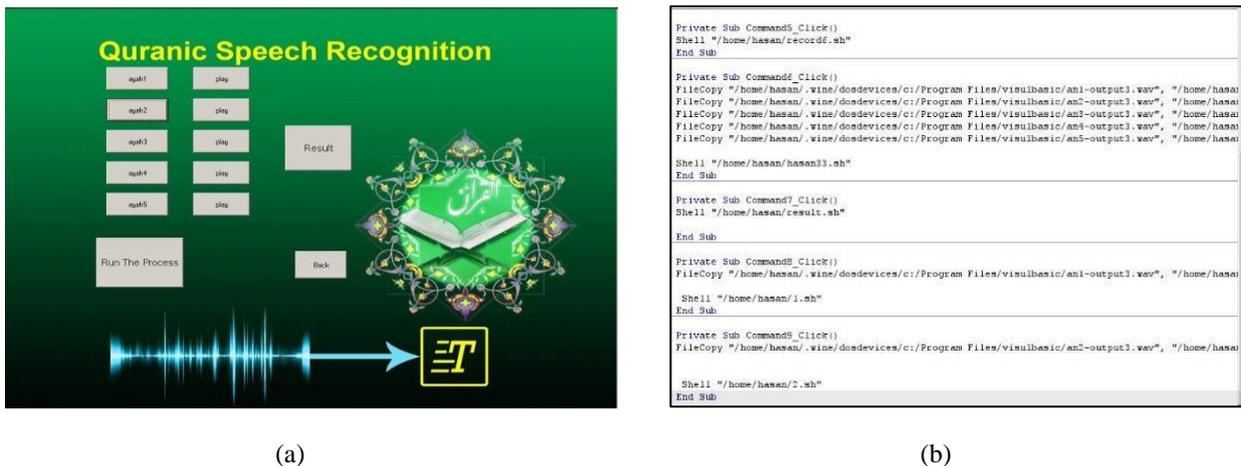


Fig. 4 - (a) GUI accuracy page; (b) Coding lines for accuracy page

#### 4. Results

The QVR system converts the Quranic verse from speech audio to text and will check the accuracy of the converted audio via text. Table 1 shows two experiments were done for 855 audios as training data and 105 audios as testing data with the results from PocketSphinx. The results indicate that both give a total of 100 percent correction and 100 percent accuracy with zero errors of word error rate (WER). This experiment results were done using the PocketSphinx by comparing the Roman text of Quran converted from both training and testing data of speech audios. Fig. 5 demonstrates the comparison results for the Roman text accuracy between the said audios ensuing from the QVR system.

**Table 1 - Trial details with results for PocketSphinx**

Experiment	Data Type	Results				
		$\Sigma\#$ Audios	$\Sigma\#$ Correct	$\Sigma\#$ Errors	$\Sigma\%$ Correct	$\Sigma\%$ Accuracy
E1	Training data	855	855	0.00	100.00	100.00
E2	Testing data	105	105	0.00	100.00	100.00

```

an4.align x
Insertions: 0 Deletions: 0 Substitutions: 0
allahou samad (tx-AN3-TX-B)
allahou samad (tx-AN3-TX-B)
Words: 2 Correct: 2 Errors: 0 Percent correct = 100.00% Error = 0.00% Accuracy = 100.00%
Insertions: 0 Deletions: 0 Substitutions: 0
lam yaled wa lam youlad (tx-AN4-TX-B)
lam yaled wa lam youlad (tx-AN4-TX-B)
Words: 5 Correct: 5 Errors: 0 Percent correct = 100.00% Error = 0.00% Accuracy = 100.00%
Insertions: 0 Deletions: 0 Substitutions: 0
wa lam yakoun lahou koufouan ahad (tx-AN5-TX-B)
wa lam yakoun lahou koufouan ahad (tx-AN5-TX-B)
Words: 6 Correct: 6 Errors: 0 Percent correct = 100.00% Error = 0.00% Accuracy = 100.00%
Insertions: 0 Deletions: 0 Substitutions: 0
bismi lahi rahmani rahim (tz-AN1-TZ-B)
bismi lahi rahmani rahim (tz-AN1-TZ-B)
Words: 4 Correct: 4 Errors: 0 Percent correct = 100.00% Error = 0.00% Accuracy = 100.00%
Insertions: 0 Deletions: 0 Substitutions: 0
koul houwa llahou ahad (tz-AN2-TZ-B)
koul houwa llahou ahad (tz-AN2-TZ-B)
Words: 4 Correct: 4 Errors: 0 Percent correct = 100.00% Error = 0.00% Accuracy = 100.00%
Insertions: 0 Deletions: 0 Substitutions: 0
allahou samad (tz-AN3-TZ-B)
allahou samad (tz-AN3-TZ-B)
Words: 2 Correct: 2 Errors: 0 Percent correct = 100.00% Error = 0.00% Accuracy = 100.00%
Insertions: 0 Deletions: 0 Substitutions: 0
lam yaled wa lam youlad (tz-AN4-TZ-B)
lam yaled wa lam youlad (tz-AN4-TZ-B)
Words: 5 Correct: 5 Errors: 0 Percent correct = 100.00% Error = 0.00% Accuracy = 100.00%
Insertions: 0 Deletions: 0 Substitutions: 0
wa lam yakoun lahou koufouan ahad (tz-AN5-TZ-B)
wa lam yakoun lahou koufouan ahad (tz-AN5-TZ-B)
Words: 6 Correct: 6 Errors: 0 Percent correct = 100.00% Error = 0.00% Accuracy = 100.00%
Insertions: 0 Deletions: 0 Substitutions: 0
TOTAL Words: 420 Correct: 420 Errors: 0

```

**Fig. 5 - Results of Roman text accuracy**

The results using a real-time microphone and recorded audio are shown in Table 2 and Table 3 respectively. The al-ikhlas surah with 22 words shows insignificant results at both conditions using a microphone despite trained and untrained speakers. On the other hand, using recorded audio in a close room condition resulted in 100 percent accuracy for both trained and untrained audios with 0.00% WER. This demonstrates that the system runs in its best performance in a close room condition with no other sound interference while recording took place.

**Table 2 - Results using microphone**

Condition	Data Type	Results					
		$\Sigma\#$ Words	$\Sigma\#$ Correct	$\Sigma\#$ Errors	$\Sigma\%$ Correct	$\Sigma\%$ Errors	$\Sigma\%$ Accuracy
Close Room	Trained speaker	22	14	8	63.63	36.36	63.63
	Untrained speaker	22	12	10	54.54	56.56	54.54
Noisy Space	Trained speaker	22	12	10	54.54	56.56	54.54
	Untrained speaker	22	6	16	27.27	72.72	27.27

**Table 3 - Results using audio**

Condition	Data Type	Results					
		$\Sigma\#$ Words	$\Sigma\#$ Correct	$\Sigma\#$ Errors	$\Sigma\%$ Correct	$\Sigma\%$ Errors	$\Sigma\%$ Accuracy
Close Room	Trained audio 1	22	22	0	100.00	0.00	100.00
	Trained audio 2	22	22	0	100.00	0.00	100.00
	Untrained audio 1	22	22	0	100.00	0.00	100.00
	Untrained audio 2	22	22	0	100.00	0.00	100.00
Noisy Space	Trained audio	22	18	4	81.81	18.18	81.81
	Untrained audio	22	18	4	81.81	18.18	81.81

## 5. Conclusion

The QVR system using PocketSphinx with ASR engine system shows that it has been successfully designed and developed in an Ubuntu platform. The GUI design using Microsoft Visual Basic 6 makes the system more readable, efficiently and user-friendly. The results successfully achieved the accuracy of a 100% for speech data using audio with 0.00% word error rate (WER) in a close room condition in spite of trained and untrained audios by comparing the

converted Quran Roman text. These remarkable results show that the developed system is significant to be implemented to the public. One of the hadiths of the Prophet S.A.W in Sahih al-Bukhari, Hadith No. 5027, Book 66, Chapter 21:

عَنِ النَّبِيِّ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ قَالَ " ( خَيْرُكُمْ مَنْ تَعَلَّمَ الْقُرْآنَ وَعَلَّمَهُ )

The Prophet (ﷺ) said, "The best among you (Muslims) are those who learn the Qur'an and teach it."

Hence, it is said to be that this developed system made it wisely practical to be performed in this pandemic era, where according to Ta'a, *et al.* (2021), the field of information technology has now greatly helped reciters read the Quran better and understand it efficiently. Further suggestions can be improved by developing an apps system in smart devices either in iOS or Android's platform. Also, additional other Quran surahs will be beneficial and useful to the users.

## Acknowledgement

The authors would like to thank the UTHM for their encouragement and acknowledge the financial support received from the government of Malaysia.

## References

- Al-Bukhari, Muḥammad ibn Ismā'īl. Shahih Al-Bukhari, 6. <https://sunnah.com/bukhari:5027> (Retrieved at 20 Sep 2021).
- Alfaries, A., Al-Bahlal, M., Almazrua, M., Almazrua, A. (2013). A rule-based annotation system to extract *tajweed* rules from Quran, Taibah University International Conference on Advances in Information Technology for the Holy Quran and Its Sciences, 281–286
- Ali, W. (2020). Online and remote learning in higher education institutes: a necessity in light of Covid-19 pandemic. *Higher Education Studies*, 10(3), 16–25
- Arora, S. J. & Singh, R. (2012). Automatic speech recognition: A review. *International Journal of Computer and Applications*, 60, 34–44
- Assami, E., Kennedy, M., & Bantley, A. (1997). *The Sahih International*. Saudi Arabia: Publishing House dar Abul Qasim
- Ciotti, M., Ciccozzi, M., Terrinoni, A., Jiang, W. C., Wang, C. B., & Bernardini, S. (2020). The COVID-19 pandemic. *Critical reviews in clinical laboratory sciences*, 57(6), 365-388
- Daines, D. H. et al. (2006). Pocketsphinx: A free, real-time continuous speech recognition system for hand-held devices, *IEEE International Conference on Acoustics, Speech and Signal Processing*, 185–188
- Dwivedi, Y.K., Hughes, D.L., Coombs, C., Constantiou, I., Duan, Y., Edwards, J.S., Gupta, B., Lal, B., Misra, S., Prashant, P.m & Raman, R. (2020). Impact of COVID-19 pandemic on information management research and practice: Transforming education, work and life. *International Journal of Information Management*, 55, p.102211.
- Habash, M. (1986). *How to memorize the Quran*. Beirut: Dar al-Khayr
- Hafeez, A. H., Mohiuddin, K., & Ahmed, S. (2014). Speaker-dependent live Quranic verses recitation recognition system using Sphinx-4 framework, *IEEE 17th International Multi Topic Conference*, 8–10
- Ibrahim, H. & Varol, A. (2020). A study on automatic speech recognition systems, *8th International Symposium on Digital Forensics and Security*, 1–5
- Kadir, M. N., Azmi, R., Mohd Saad, M. F., Zainol, N. Z. N., & Yunus, M. A. M. (2021). A Review on the Manuscript of Turkish Quran YWJ68197MSI.4 in Database Management System (DBMS). *Journal of Quranic Sciences and Research*, 1(1), 34–41
- Levis, J. M., & Suvorov, R. (2012). Automatic speech recognition. In C. Chapelle (Ed.), *Encyclopedia of applied linguistics* (pp. 1–8). New Jersey: Blackwell Publishing
- Muhammad, W. M., Muhammad, R., Muhammad, A. & Martinez-Enriquez, A. M. (2010). Voice content matching system for Quran readers, *Ninth Mexican International Conference on Artificial Intelligence*, 148–153
- Mukhtar, K., Javed, K., Arooj, M. & Sethi, A. (2020). Advantages, limitations and recommendations for online learning during covid-19 pandemic era. *Pak J Med Sci.*, 36(S4), 27–31
- Płonkowski, M., & Urbanovich, P. (2018). Using Full Covariance Matrix for CMU Sphinx-III Speech Recognition System. *Przełąd Elektrotechniczny*, 102–104
- Schindler, L. A., Burkholder, G. J., Morad, O. A., & Marsh, C. (2017). Computer-based technology and student engagement: a critical review of the literature. *International journal of educational technology in higher education*, 14(1), 1–28
- Sundawa, D., Logayah, D. S., & Hardiyanti, R. A. (2021). New normal in the era of pandemic COVID-19 in forming responsibility social life and culture of Indonesian society, *3rd International Conference on Environmental Geography and Geography Education*, 1–10
- Ta'a, A., Shaari@Hj Ashari, M. S., Che Mohamed Arif, A. S., Abu Bakar, M. S., & Mahmad, M. A. (2021). Semantic Approach for Al-Quran Knowledge Representation in Islamic Knowledge Management Framework. *Journal of Quranic Sciences and Research*, 2(1), 35–46

- Tabbal, H., El-Falou, W., & Monla, B. (2006). Analysis and implementation of a “Quranic” verses delimitation system in audio files using speech recognition techniques, IEEE Conference of 2nd Information and Communication Technologies (ICTTA), 2, 2979–2984
- Yousfi, B. & Zeki, A. M. (2016). Automatic speech recognition for the Holy Quran: A Review, International Conference on Data Mining, Multimedia, Image Processing and their Applications, 23–29.
- Yusuf Ali, A. Al-Quran Translation. <https://quranyusufali.com/73/> (Retrieved at: 20 Sep 2021)