

# Error Analysis of Latin-to-Balinese Script Transliteration Method Based on Noto Sans Balinese Font

L. H. Loekito<sup>1</sup>, G. Indrawan<sup>2</sup>, Sariyasa<sup>3</sup>

Dept. of Computer Science  
Ganesha University of Education  
Singaraja, Indonesia

<sup>1</sup>lucianahendrika7@gmail.com, <sup>2</sup>gindrawan@undiksha.ac.id,  
<sup>3</sup>sariyasa@undiksha.ac.id

I K. Paramarta

Dept. of Balinese Language  
Ganesha University of Education  
Singaraja, Indonesia  
ketut.paramarta@undiksha.ac.id

**Abstract**— Balinese script is a part of Balinese culture which is feared to be extinct due to the decrease in its use. This research is one of the efforts to preserve Balinese Script using technology, which focuses on the error analysis of Latin-to-Balinese script transliteration method based on Noto Sans Balinese font. This analysis can be used as a reference for the future improvement to make this method more robust. This research used accuracy testing method to test compatibility result of transliteration produced by the application with 27,506 words contained in the Balinese Language Dictionary. This experiment result shows the accuracy of the application is 54.5%. After the addition of special word and correction on the application algorithm, the result changed to 94%. There is still 6% of other incorrect transliteration because of the various case, related to a word has several different meanings (so different Balinese scripts), and related to the affixed word. In the future work, these 6% incorrect transliteration need to be accommodated to make the recent method more robust.

**Keywords**—Balinese; latin; noto sans Balinese; transliteration.

## I. INTRODUCTION

The Balinese script was introduced from an ancient Javanese script which is known as *Sanskrit*. The development of scripts starting from northern and southern India is known as the *Brahmi* script. However, the use of a Balinese script known as *Hanacaraka* has begun to be forsaken, especially by the younger generation [1]. The decline of Balinese culture, especially in the use of Balinese script among the people can cause the extinction of Balinese culture which is one of the important identities of the island of Bali. Referring to the Governor of Bali Regulation Number 80 of 2018 about the protection and use of Balinese language, script and literature, the government began circulating circular letter Number. 3172 of 2019 [2]–[4], intended for government and non-government institutions, consulates, private companies and event organizers, to use Balinese in the effort of preserving the culture of Bali as the identity of the island of Bali itself.

In order to help the government in preserving the Balinese language, transliteration [5], a research by Indrawan,

Paramarta, & Agustini in 2019 developed Balinese scripts transliteration web-based and mobile application using the Noto Sans Balinese font, which can translate Latin letter into Balinese script. This study aimed at testing the accuracy of the transliteration results of the application when translating Latin letter into Balinese script using the Noto Sans Balinese font with the words contained in the Balinese - Indonesian Dictionary. The dictionary used in this study is the Balinese - Indonesian Dictionary provided with its writing in Latin and Balinese Script collaboration between the Denpasar City Cultural Office and the Literature Development Board, and the Bali Literature of the Bali Province 2008 [6]. This study also explored errors found in the results of Latin transliteration into the Balinese script in order to provide recommendations for improving applications for application development to be more reliable.

## II. BALINESE SCRIPT

### A. Balinese Script

The Balinese script has a manuscript derived from the ancient *Brahmic* manuscripts in India. It makes the Balinese script have similarities with modern manuscripts from Southeast Asia and South Asia which are also derived from the *Brahmic* [7]. The Balinese script is often referred to as the alpha syllable (*Abugida*). This script is based on consonants with mandatory secondary vocal notation. The letters in the Balinese script are divided into two parts according to the Great Islamic School in 1963, namely *Wianjana* Script (Consonant Letter) and Voice Script (Vocal Letter). The pure Balinese language is written with 18 consonants and 9 vowels. In each consonant has a hanger that serves to make the vowels that existed in the previous syllable cancelled and punctuation such as commas, periods, colons [8].

Because the Balinese script is derived from ancient *Brahmic* manuscripts and has several complex scripts, the rendering into the Balinese script is more complex, as follows [9]: 1) Repairs and character sorting is needed, because there may be one or more separate glyphs; 2) Laying in accordance



running well or not. After the testing process was complete, the results obtained were in the form of the number of special words that must be added to the repository as additional knowledge and classification of errors produced by the application. And will be given a percentage for the accuracy of the test results of the transliteration application. Below is a percentage graph that displays the number of words by the alphabet from the alphabet A to W which is used as a test case in this study which is by the Bali-Indonesian Dictionary. Also, some letters are not found in the Bali-Indonesian Dictionary because the alphabet is an adaptation of a foreign language namely the alphabet F, Q, V, X, and Z. The testing cases of this study can be seen in Figure 3.

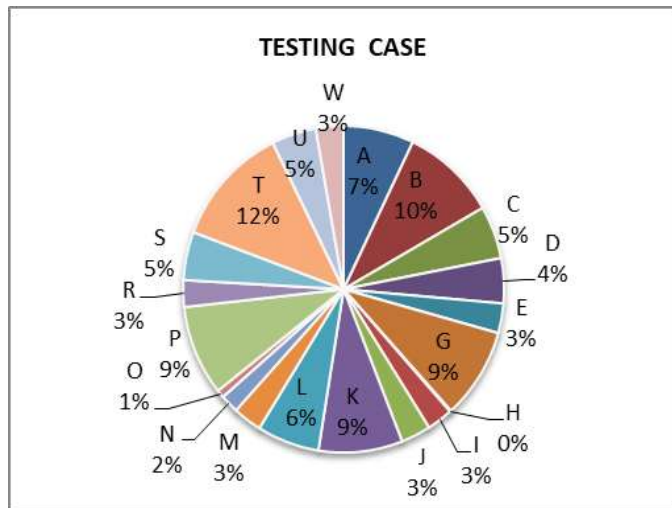


Fig 3. Testing Case

IV. RESULT AND ANALYSIS

The is a graph showing the recapitulation percentage of transliteration test results based on the Bali-Indonesian Dictionary. The results obtained in this study are the total number of words contained in the Bali-Indonesian Dictionary of 27,506 words. From the results of the testing case above there were 14,996 words in a percentage of 54.5% for all alphabets from A to W which are the correct number of Latin transliterations to Balinese script according to the dictionary as well as the Balinese language rules without any improvements or changes in the algorithm and application repository. It can be seen in Figure 4 below.

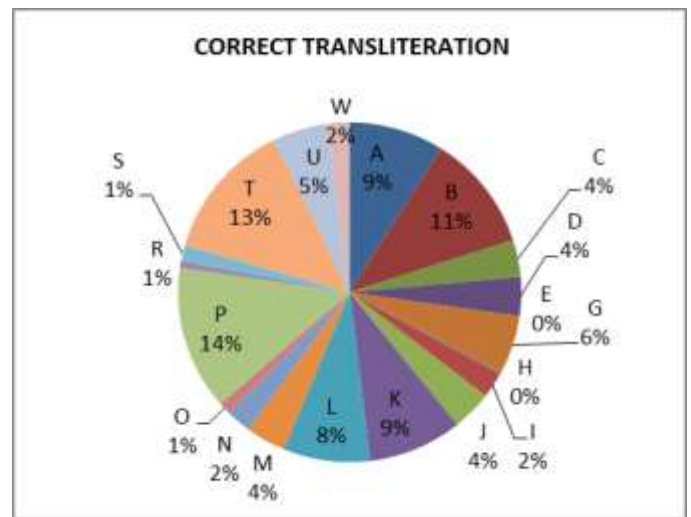


Fig 4. Correct Transliteration

Additionally, other results were obtained in the form of errors generated by the transliteration application based on a test case of 11,104 words in a percentage of 38.8% for alphabet A to W. This error was caused by application algorithms that could not recognize words entered in the application. The error graph can be seen in Figure 5 below.

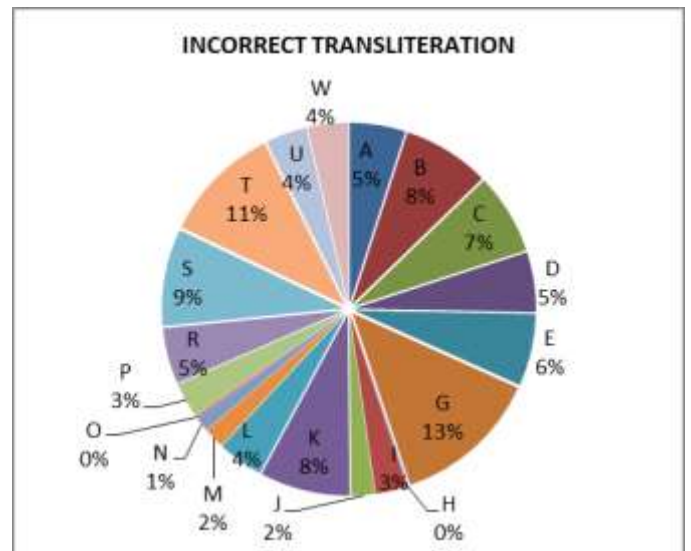


Fig 5. Incorrect Transliteration

Based on testing, the wrong word that cannot be transliterated by the application and requires improvement in transliteration applications was found. This table of error can be seen in the table below:



As the example above when the words entered are the same, there will only be one word that can be transliterated by the application. Because the Latin word has the same writing but in Balinese script, it has different meanings, one means seeing and one meaning style or action with different characters. The improvement solution in this application is the addition of special words to the application repository so that the application algorithm can distinguish each word inputted according to the Balinese script written in the Balinese dictionary. The changes made to the application algorithm can be seen in the pictures below:



Fig 9. Examples of Special Words in The Application Repository

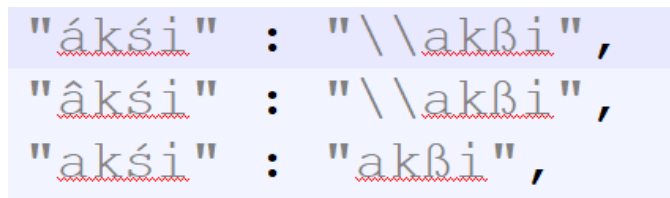


Fig 10. Example of An Algorithm in The Application Repository

After changing the algorithm in the application as discussed above and re-testing the application we get an increase in accuracy from 14,996 words in a percentage of 54.5% for the correct transliteration category to 26,100 words in this 94% percentage obtained from the results of transliteration algorithm changes that have Errors in the above categories have been fixed such as the removal of punctuation and the addition of additional knowledge from alphabet A to W according to the dictionary. It can be seen in the image below:

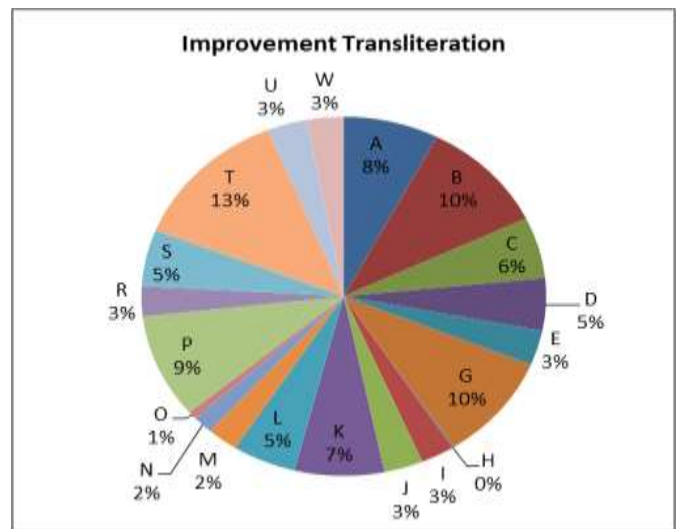


Fig 31. Improvement Transliteration

After the improvements made to the application, there are still errors of the transliteration application of 1,406 words in a percentage of 6% due to an application algorithm error that cannot be fixed in this study. These errors including:

C. Stacking of Balinese Fonts

In the transliteration results of the Balinese script, the results shown are correct but in the Balinese script display, there is a buildup or overlap between the characters with one another. This causes the results of transliteration to be confusing and not visible. For this problem, the error occurs not in the Latin word input entered into the application but in the error caused by a bug from the Noto Sans Balinese font. The following table is an example of the results of the transliteration of characters that have font buildup:

TABLE IV. DISPLAY OF TRANSLITERATION RESULTS

Words in Balinese - Indonesian Dictionary	Transliteration Results with the Noto Bali Application
<p>Luir</p>	<p>Latin Text</p> <p>luir</p> <p>Balinese Scri</p>

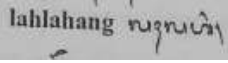
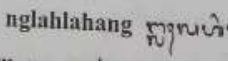

A temporary solution that can be provided for this font buildup problem is to wait for the latest version update from the application developer Font Noto Sans Balinese to be able to correct the error.

D. Transliteration Errors in Letter Splitting

At the time of transliteration from Latin words to Balinese script, the transliterated words are not only basic words but also words that have been given affixes in the form of prefix

and suffixes to indicate these words are derived from existing basic words. This application is a problem because not all basic words that have affixes can be transliterated correctly. Moreover, there are thousands of basic words in the dictionary that must be transliterated so it is not possible to add all the basic words to the dictionary in the application. Here are examples of basic words and affixed words that cannot be transliterated correctly:

TABLE V. DISPLAY OF TRANSLITERATION RESULTS

Words in Balinese - Indonesian Dictionary	Transliteration Results with the Noto Bali Application
 <p>Lahlahang</p>  <p>Nglahlahang</p>	<p>Latin Text</p> <p>lahlahang</p> <p>Balinese Script</p> 

A temporary solution that can be given to solve this problem is an improvement on the application algorithm, in the form of improvements to text processing, namely the stemming process. The stemming process is used to break words in the application into basic words so that when the affixed words are entered it will produce the correct transliteration because the basic words are already available in the application. However, attention must be paid to each word for the stemming algorithm.

**V. CONCLUSION AND FUTURE WORK**

Another error transliteration analysis based on the Noto Sans Balinese font and using the Bali-Indonesian Dictionary as a testing case achieved a fairly good accuracy of 94% (26,100 out of 27,506 words) achieved in this transliteration method. It consists of 14,996 (54.5%) words of the correct number of Latin transliterations into Balinese script according to the dictionary also the rules of the Balinese language without any improvements or changes in the algorithm and repository of the application. Another number of Latin transliterations into the Balinese script are 11.104 (38.8%) words that experience errors due to application algorithms that cannot recognize the words entered in the application. After the improvement of the application algorithm in the form of the addition of repositories in the application and improvement of punctuation in the application algorithm, the results were achieved by 94%. However, there are still 1,406 (6%) words that cannot be transliterated due to errors that cannot be corrected. In the future work, this method could be enhanced by: 1) Correction of errors that have been found in this research for future application development; 2) Conducting further exploration of testing outside of the Balinese Dictionary and seventeen Balinese script rules from Sudewa.

**ACKNOWLEDGMENT**

This research was supported by the Indonesian Ministry of Research, Technology, and High Education through Grant No. 147/UN48.16/LT/2019.

**REFERENCES**

- [1] A. A. K. Oka Sudana, K. S. Wibawa, and I. M. A. D. Tirtha, "Learning media of Balinese script writing based on augmented reality," *J. Theor. Appl. Inf. Technol.*, vol. 90, Aug. 2016.
- [2] Antara News, "Koster anjurkan acara internasional gunakan aksara dan busana Bali," [www.antaranews.com](http://www.antaranews.com), 2019. [Online]. Available: <https://www.antaranews.com/berita/830352/koster-anjurkan-acara-internasional-gunakan-aksara-dan-busana-bali>. [Accessed: 17-Jul-2019].
- [3] Gubernur Bali, *Perlindungan dan Penggunaan Bahasa, Aksara, dan Sastra Bali*. Denpasar, 2018.
- [4] Gubernur Bali, *Surat Edaran Nomor 3172 Tahun 2019 Anjuran Penggunaan Aksara dan Busana Adat Bali*. Denpasar, 2019.
- [5] I. N. Jampel, G. Indrawan, and I. W. Widiyana, "Accuracy analysis of Latin-to-Balinese script transliteration method," *Int. J. Electr. Comput. Eng.*, vol. 8, 2018.
- [6] I. G. K. Anom et al, "Kamus Bali - Indonesia beraksara Latin dan Bali. Denpasar: Badan Pembina Bahasa, Aksara, dan Sastra Bali," 2009.
- [7] I. Habibi and R. Munir, "The Balinese unicode text processing," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 1, no. 1, 2009.
- [8] C. Science, J. Campus, M. Sudarma, S. Ariyani, and M. Artana, "Balinese script ' s character reconstruction using linear discriminant analysis," vol. 4, 2016.
- [9] G. Indrawan, I. K. Paramarta, K. Agustini, and Sariyasa, "Latin-to-Balinese script transliteration method on mobile application: A comparison," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 10, 2018.
- [10] Google, "Google Noto Fonts," 2014. [Online]. Available: <https://www.google.com/get/noto/#sans-bali>. [Accessed: 17-Jul-2019].
- [11] G. Indrawan, N. N. H. Puspita, I. K. Paramarta, and Sariyasa, "LBtransbot: A Latin-to-Balinese script transliteration robotic system based on noto sans Balinese font," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 12, 2018.
- [12] B. B. Narendra, "Pembuatan smart font aksara Bali dengan graphite description language, undergraduate thesis," Institut teknologi Bandung, 2008.
- [13] I. B. A. Sudewa, "The Balinese Alphabet," 2003. [Online]. Available: [www.babadbali.com](http://www.babadbali.com). [Accessed: 17-Jul-2019].
- [14] G. Indrawan, I. K. Paramarta, and K. Agustini, "A new method of Latin-to-Balinese script transliteration based on noto sans balinese font and dictionary data structure," *2nd Int. Conf. Softw. Eng. Inf. Manag.*, 2019.