

Accuracy Analysis of Latin-to-Balinese Script Transliteration Method

I N. Jampel, G. Indrawan, I W. Widiana
Ganesha University of Education, Bali, Indonesia

Article Info

Article history:

Received Nov 3, 2017

Revised Jan 19, 2018

Accepted Jan 26, 2018

Keyword:

Accuracy

Balinese

Latin

Mobile application

Transliteration

ABSTRACT

Balinese script writing, as one of Balinese cultural richness, is going to extinct because of its decreasing use. This research is one of the ways to preserve Balinese script writing using technological approach. This research focused on accuracy analysis of Latin-to-Balinese script transliteration method on mobile application since mobile computing becomes more important way on preservation effort. This analysis used an existing method from few research in this area, i.e. method on mobile application that was called Transliterasi Aksara Bali (TAB). The analysis was based on certain document of Balinese script writing rules and examples. Through the experiment, TAB was considered has accuracy over 68% (103 of 151) cases and improvement recommendation on thirteen kind of special words. This research contributes on the accuracy analysis and that recommendation for future development of Latin-to-Balinese script transliteration method.

*Copyright © 2018 Institute of Advanced Engineering and Science.
All rights reserved.*

Corresponding Author:

I N. Jampel,
Department of Educational Technology,
Ganesha University of Education,
Jl. Udayana 11, Singaraja 81116, Bali, Indonesia.
Email: jampel@undiksha.ac.id

1. INTRODUCTION

Balinese script writing, as one of Balinese cultural richness, is going to extinct. As reported by The Indonesia News Agency [1], Bali Governor, Made Mangku Pastika, said that although Balinese language education has been included in the educational curriculum at the school, but less and less people who are fluent in Balinese language. This is because for the simplicity, daily communication in Bali tend to use national language, i.e. Indonesian language, rather than the local language. Since long time, the less of users or speakers of Balinese language has caused concern over the threat of extinction of this local language.

The decreased use of the Balinese language has big influence to the decreased use of its Balinese script writing, as this research concerned about. Stern [2] said that *"As with endangered species, the goal with languages should not be to wait until there are only a few remaining survivors and then place them under protection (i.e. make recordings of the last speakers). Rather, we should take the more sustainable path of preserving the diverse natural habitats where minority languages are spoken. This means taking a cultural, political and even economic approach to saving languages ... and starting as soon as possible."*

For preservation of Balinese language, more specifically on its Balinese script writing, another approach was taken, i.e. technological approach by this research. This research focused on accuracy analysis of Latin-to-Balinese script transliteration method on mobile application, as it eases people by providing handy transliteration learning on mobile device. Transliteration itself is the conversion of a text from one script to another [3]. This analysis using an existing method from few research in this area, i.e. method on mobile application that was called Transliterasi Aksara Bali (Balinese Script Transliteration) [4]. The analysis was based on Balinese script writing rules and examples by Sudewa [5], as a project script committee related to the proposal by Eversen and Suatjana [6] for encoding the Balinese script in ISO [7].

Not so many references in this research area. A work related to the same object, i.e. Balinese script, was conducted by Sudana et al. [8] but their work is on Augmented-Reality-based learning media application that was focused on learning on how to write Balinese script. On Latin-to-Balinese script transliteration research area, Sartini et al. [9] has developed a text-to-digital-image converter method. The output is the Balinese script represented by using pre-collected images that were captured from Bali Simbar font [10] display at word processor. Arimbawa et al. [11] has developed a Latin-to-Balinese script transliteration method where the output Balinese script pattern was configured to be written by the robotic system.

2. BALINESE SCRIPT

Through the effort of a non-profit organization in Balinese culture preservation, i.e. Yayasan Bali Galang [12], Balinese script was added to the Unicode Standard. Table 1 shows that the Unicode Consortium [13] allocate the Unicode block for Balinese at U+1B00 – U+1B7F. Non-assigned *code points* was indicated by grey areas.

Table 1. Balinese Unicode

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
U+1B0x	ꦲ	ꦱ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ
U+1B1x	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ
U+1B2x	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ
U+1B3x	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ
U+1B4x	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ
U+1B5x	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ
U+1B6x	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ
U+1B7x	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ	ꦲ

Based on Table 1, Balinese *code point* consists of several group allocation for: 1) Various signs (1B00–1B04); 2) Independent vowels (1B05–1B12); 3) Consonants (1B13–1B33); 4) Balinese sign *rerekan* (1B34); 5) Dependent vowel signs (1B35–1B43); 6) Balinese *adeg-adeg* (1B44); 7) Additional consonants (1B45–1B4B); 8) Digits (1B50–1B59); 9) Punctuation (1B5A–1B60 for); 10) Musical symbols for notes (1B61–1B6A); 11) Diacritical marks for musical symbols (1B6B–1B73); and 12) Musical symbols (1B74–1B7C).

Complex rendering was demanded by Balinese script complex behaviours, includes: 1) Reordering and splitting; 2) Various placement and shape of diacritics; 3) Contextual shaping; and 4) Complex ligature construction. Table 2 no. 1–3 shows that complex rendering from several Balinese syllable examples, i.e.: 1) *ba* (U+1B29 Balinese letter *ba*); 2) *be* from *ba* + *pangangge suara e* (U+1B3E Balinese sign *taling*). *Taling* is placed on the left of syllable so that it is appeared as if it is written first and then followed by *ba* [14]. In fact, it is written later. This Balinese complex behaviour was called reordering; and 3) *bo* from *ba* + *pangangge suara o* (U+1B40 Balinese vowel sign *taling tedung*). Separated *taling* and *tedung* each written before and after the syllable. This Balinese script complex behaviour was called reordering and character splitting. Also, *taling tedung* is an example of a character that has several separated glyphs. Table 2 no. 4–7 shows various placement and shape of diacritics, i.e.: 4) *di* from *da* (U+1B24 Balinese letter *da*) + *pangangge suara i* (U+1B36 Balinese vowel sign *ulu*); 5) *dīng* from *da* + *i* + *pangangge tengenan ng* (U+1B02 Balinese sign *cecek*). *Ulu* at *di* was located in the center above *da*, while *ulu* at *dīng* was slightly shifted by *cecek*; 6) *dě* from *da* + *pangangge suara ě* (U+1B42 Balinese vowel sign *pepet*); and 7) *děr* from *da* + *ě* + *pangangge tengenan r* (U+1B03 Balinese sign *surang*). *Pepet* at *dě* was located in the center above *da*, while *pepet* at *děr* not only was slightly shifted by *surang* but also change smaller to make width of *pepet surang* equal to *da* below them. Table 2 no. 8–10 show several forms of glyph that represent *gantungan* of Balinese syllables *ra* (U+1B2D Balinese letter *ra*), also known as *cakra* or *guwung*; 8) *kra* from *ka* (U+1B13 Balinese letter *ka*) + *gantungan ra*; 9) *skra* from *sa* (U+1B32 Balinese letter *sa*) + *ka* + *gantungan ra*; and 10) *krya* from *ka* + *gantungan rya* which is combination of *gantungan ra* (the third *cakra*) + *gantungan ya* (see Table 2 no. 12). The shape of *cakra* glyph on *kra* (the first *cakra*) is narrower than the shape of *cakra* glyph on *skra* (the second *cakra*). Besides, glyph is written below *ka* at the end of the first *cakra*, while glyph is written beside *ka* at the end of the second *cakra*. This Balinese script complex behaviour shows some

characters require glyph selection based on character context. Also, the third *cakra* shows the Balinese script complex behaviour on ligature construction. Table 2 no. 11–12 show those ligatures constructions, for an example *nā* from *na* (U+1B26 Balinese letter *na*) + *pangangge suara ā* (U+1B35 Balinese vowel sign *tedung*).

Table 2. Complex behaviour of Balinese script

No	Latin	Balinese	No	Latin	Balinese	Gantungan
1	ba	ꦧ	7	dér	ꦢꦺꦫ	
2	be	ꦧꦺ	8	kra (using gantungan ra)	ꦏꦫꦫ	ꦫꦫ
3	bo	ꦧꦺꦴ	9	skra (using gantungan ra)	ꦱꦏꦫꦫ	ꦫꦫ
4	di	ꦢꦶ	10	krya (using gantungan rya)	ꦏꦫꦫꦫ	ꦫꦫ
5	ding	ꦢꦶꦁ	11	na + ā = nā	ꦤꦲꦫꦫꦫ	
6	dě	ꦢꦶꦁꦺ	12	ra + ya = rya	ꦫꦫꦫꦫꦫ	

3. RESEARCH METHOD

The accuracy analysis of Latin-to-Balinese script transliteration method was based on Balinese script writing rules and examples (cases) by Sudewa [5] (Table 3), where some of those cases referred to [14] [15]. That analysis was conducted on mobile application that was called Transliterasi Aksara Bali [4]. Several rules can not be tested independently without example, like the appended form of eighteen basic syllables at case 1–18, since provided examples are limited (case 19–25).

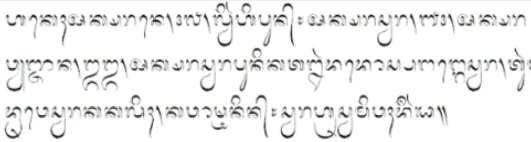
Table 4 shows case sixteen, word boundaries and line break rules (Table 3), with the given sentence and its transliteration.

Table 3. Testing cases

No.	Case	Remark	No.	Case	Remark
1	1-18 19-25	Table 1* basic syllables (no. 1-18) Table 1* examples: word <i>Bakta</i> (bring) ... <i>Krama</i> (member)	9	86-87	Holy symbol <i>Ongkara</i> examples: word <i>Om Swastiasu</i> (May God blesses you) ... <i>Om Śanti, Śanti, Śanti, Om</i> (May peace be everywhere)
2	26-34	Table 2* vowel signs examples: word <i>Kādep</i> (sold) ... <i>Dwi</i> (two)	10	88-89	Table 9* miscellaneous syllables (no. 1-2)
3	35-44 45-51	Table 3* independent vowels (no. 1-10) Table 3* examples: word <i>Aksara</i> (alphabet)... <i>Om</i> (symbol of God)	11	90-99	Table 10* the digits (digit: 0 - 9)
4	52-55 56-57	Table 4* illegal combination of syllable - vowel signs (no. 1-4) Table 4* examples: word <i>Talér</i> (also) ... <i>Kéréng</i> (frequently)	12	100-107	Table 11* punctuations: name <i>Carik</i> (a comma) ... Double Quotes
5	58-62	Table 5* semi vowels examples: word <i>Pak Raman</i> (Mr. Raman) ... <i>Briag</i> (intense)	13	108-113 114-115 116 117-124	Some variation of usages. Combination of independence vowel <i>a</i> <i>kara</i> with vowel signs: vowel <i>i</i> ... <i>õ</i> Pairing of <i>pa kapal</i> with <i>suku</i> or <i>suku</i> <i>ilut</i> : syllable <i>pu</i> ... <i>phu</i> Romanization of the inherent sound: word: <i>Sekala</i> (real) Usage of <i>pangangge aksara</i> : word <i>Samping</i> (side)... <i>Tamblang</i> (a village's name)
6	63-71 72-78	Table 6* <i>aksara swalita</i> (no. 1-9) Table 6* examples: word <i>Gapitri</i> (chain) ... <i>Laghu</i> (low tone in singing)	14	125-146	Table 13* <i>ligatures</i> (No. 1-22)
7	79-83	Table 7* sound killers examples: word <i>Cengceng</i> (a musical instrument) ... <i>Kapal</i> (ship)	15	147-150	Abbreviations example: word Bank <i>Pembangunan Daerah Bali</i> (Development Bank of Bali Province)... <i>Ba Pa Da Bali</i> (BPD Bali)
8	84-85	Table 8* miscellaneous signs examples: word <i>Mang</i> (a holy letter) ... <i>Siddham</i> (perfect)	16	151	Word boundaries and line break rules

*refer to table number at Sudewa [5]

Table 4. Provided sentence and its transliteration

No	Script	Example
1	Latin	<i>Akeh aksarane, 47, luih ipun: aksara suara, 14, aksara wianjana, 33, aksara suara punika talèr dados pangangge suara, tur madhèwe suara kakalih, kawaštjanin: suara hrèswa miwah dirgha. (Many letters, 47, i.e.: vowels, 14, consonants, 33, those vowels also become vowel signs, and have two type of sounds, that were called: sound hrèswa and dirgha)</i>
	Balinese	

4. RESULTS AND ANALYSIS

Experiment for accuracy analysis of Latin-to-Balinese script transliteration method on existing method Transliterasi Aksara Bali version 0.0.2 [4] was conducted on Windows 7 64-bit Operating System, powered by Intel Core M-5Y71 CPU @1.20GHz platform using 8 GB RAM. For the next reference, this method was referred by its abbreviation, i.e. TAB. Since this method using Android mobile platform, the testing using Android Emulator with Nexus 5 platform and Nougat (Android 7.1.1) 32-bit Operating System.

Table 5 shows the accuracy result where column *Case* represents rule or example of Table 3 and column *Result* shows result correct or incorrect (each was marked by check and cross).

Figure 1 on left shows sentence transliteration differences by TAB, while Figure 1 on right shows the modified sentence --onsists of uncommon words in writing-- to get the result more precise.

Figure 1 shows character *ě* and *ř* replacement by *e* and *t*, respectively. There is now way to enter such characters on mobile virtual keyboard. Note that TAB provides additional keyboard for character *ś*. Table 5 shows that TAB has passed over 68% (103 of 151) cases.

Table 5. Testing Results

Case	Result	Case	Result	Case	Result	Case	Result	Case	Result
1	√	32	x	63	x	94	√	125	√
2	√	33	x	64	√	95	√	126	√
3	√	34	√	65	√	96	√	127	√
4	√	35	√	66	x	97	√	128	√
5	√	36	√	67	x	98	√	129	√
6	√	37	√	68	x	99	√	130	√
7	√	38	√	69	√	100	√	131	√
8	√	39	√	70	√	101	√	132	√
9	√	40	x	71	√	102	√	133	√
10	√	41	√	72	x	103	√	134	√
11	√	42	√	73	√	104	√	135	√
12	√	43	√	74	x	105	√	136	√
13	√	44	√	75	x	106	√	137	√
14	√	45	x	76	x	107	√	138	√
15	√	46	x	77	x	108	√	139	√
16	√	47	x	78	√	109	√	140	√
17	√	48	√	79	√	110	√	141	√
18	√	49	x	80	√	111	√	142	√
19	√	50	x	81	√	112	x	143	√
20	√	51	x	82	x	113	x	144	√
21	√	52	x	83	√	114	√	145	√
22	√	53	x	84	x	115	√	146	√
23	√	54	x	85	x	116	x	147	x
24	√	55	x	86	x	117	√	148	x
25	√	56	x	87	x	118	x	149	x
26	x	57	x	88	x	119	x	150	x
27	x	58	√	89	√	120	√	151	x
28	√	59	√	90	√	121	x		
29	√	60	√	91	√	122	x		
30	√	61	√	92	√	123	x		
31	x	62	x	93	√	124	x		

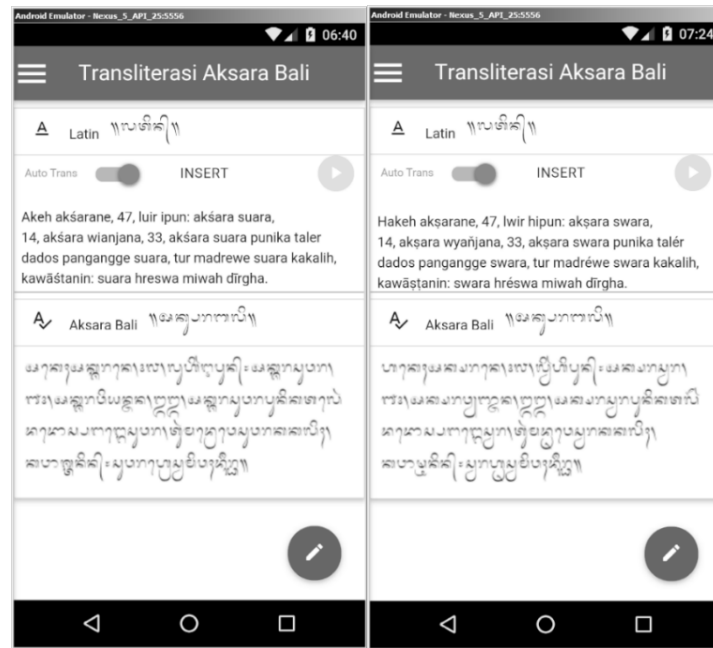


Figure 1. Sentence transliteration limitations (left); and modified sentence transliteration (right)

On the next Section, sixteen analysis for accuracy improvement was discussed based on sixteen case types on Table 3, respectively.

4.1. Basic syllables

Eighteen basic syllables (*akṣara wreṣāstra*) and provided examples (case on Table 5 no.1–25, or 6.1–6.25) were transliterated correctly by TAB.

Dual transliteration of syllable *a*, as a counterpart of syllable *ha*, was shown by case 6.1. TAB transliterated syllable *a* the same as it transliterated independent vowel *a* at case 6.35 (see next Section 4.3). Also, syllable *a* can be transliterated the same as syllable *ha* or independent vowel *a*. It depends on the word. If that word is a special word, for an example word *Akṣara* (letter) at case 6.45 (see next Section 4.3), syllable *a* will be transliterated the same as independent vowel *a*. Otherwise, that word will be transliterated the same as syllable *ha*, for an example word *Angklung* (a musical instrument) at case 6.31 (see next Section 4.2). For TAB, the transliteration algorithm can be improved to handle those special words through word searching on *dictionary* data structure that will give average *time complexity* $O(1)$ regardless of the amount of words save inside [16]. If a special word is found, simply transliterate it by using independent vowel. As a note, the use of *dictionary* data structure was previously done by the authors for biometric data discriminator in [17]-[20]. Unfortunately, there is still no research to know the precise list of those special words that influence the accuracy of developed transliteration system in general.

4.2. Vowel signs

As part of vowels (*akṣara suara*), twelve vowel signs are attached to syllables. They cannot be tested independently, as previously described. Several provided examples (case 6.26–6.34) were transliterated incorrectly, i.e. word *Kādep* (sold), *Jěro* (house), *Angklung* (a musical instrument), *Daiya* (giant), and *Patūt* (should be) at case 6.26, 6.27, 6.31, 6.32, and 6.33, respectively.

On case 6.26, vowel *ā* as *tedong ligature* was used by word *Kādep* (see next Section 4.14). Also, vowel *e* should be written using vowel *ě* (become *Kāděp*) since vowel sign *pepet* was used at transliteration result (like vowel *ě* at word *Jěro* at case 6.27). As said previously, there is a limitation on entering vowel *ě* on mobile application. TAB was incorrectly transliterating vowel *ě* but it provides a replacement vowel *é*. Vowel *e* cannot be a replacement vowel since it has already represented another sound, like vowel *e* at word *Sela* (yam) at case 6.30. Word *Kāděp* and its variations (*Kādep*, *Kaděp*, and *Kadep*) refer to one meaning and should have same transliteration. They represent another kind of special words that there is still no research to know the precise list of them. The transliteration algorithm improvement on them is basically the same like special words at previous Section 4.1. On case 6.31, vowel *A* of word *Angklung* was failed to be transliterated by TAB because of incorrect usage of independent vowel. Case 6.31 is the same as syllable *a* at case 6.1 (see

previous Section 4.1) where word *Angklung* need to be modified become *Hangklung*, which is uncommon in writing but right in transliteration. On case 6.32, TAB was failed to transliterate vowel sign *taleng repa ai* of word *Daitya* but TAB provide a replacement vowel *ê* that also in use for independent vowel *airsania ai* and word *Airlangga* (a Javanese king) at case 6.42 and 6.49, respectively (see next Section 4.3). It related to diphthong *ai* that is pronounced as long vowel *ê*. Not all of words having vowel *ai* should be transliterated like this. Word *Daitya* and its variation (*Dêtya*) refer to one meaning and should have same transliteration. They represent another kind of special words that there is still no research to know the precise list of them. The transliteration algorithm improvement on them is basically the same like special words at previous Section 4.1.

4.3. Independent vowels

As part of vowels (*akṣara suara*), ten independent vowels are used at initial position of the word (case 6.35–6.51). Some of them were transliterated incorrectly by TAB, i.e. vowel *ū*, word *Akṣara* (letter), *Iṣwara* (God's name), *Airlangga* (a Javanese king), *Ong* (a holy letter), and *Om* (symbol of God) at case 6.40, 6.42, 6.45–6.46, and 6.49–6.51, respectively.

On case 6.42 and 6.49, independent vowel *airsania ai* and word *Airlangga* were transliterated with the same mechanism on word *Daitya* at case 6.32 (see previous Section 4.2 that covered the analysis). On case 6.42, vowel *ê* was used for independently written vowel *airsania ai* for correct transliteration. Case 6.44 is basically the same as case 6.42 but on different independently written vowel *au* that should used vowel *ô* (related to the diphthong *au* that is pronounced as the long vowels *ô*) for correct transliteration.

4.4. Syllable - vowel sign combination

Illegal combination of syllable - vowel signs happened on case where syllable *ra* or *la* uses *ě* or *ö* appended to it. Syllable *ra* or *la* must use regular form *ra repa* and *la lenga*, respectively (case 6.52–6.57). All of cases were transliterated incorrectly by TAB, i.e. *rě*, *lě*, *lö*, word *Talěr* (also), and *Kěřěng* (frequently) at case 6.52, 6.53, 6.54, 6.55, 6.56 (also appeared at Table 4), and 6.57, respectively. Major cause is disability on entering vowel *ě* on mobile application, the same as word *Kāděp* (sold) at case 6.26 (see previous Section 4.2). However, TAB provides a replacement vowel *é* and the next analysis was based on that. On case 6.56, *lé* of word *Talěr* was transliterated incorrectly by using illegal combination of syllable *la* and vowel sign *pepet* (to kill previous sound *a* of *la* and replace it with sound *é*), instead of using regular form *la lenga*. Word *Talěr* and its variation (i.e. *Taler*) refer to one meaning and should have same transliteration. They represent another kind of special words that there is still no research to know the precise list of them. The transliteration algorithm improvement on them is basically the same as described on special words at previous Section 4.1. On case 6.57, the condition is basically the same as case 6.56 but on different word with syllable *ré*.

4.5. Semi vowels

Four semi-vowels (*arda suara*) attached to syllable, i.e. *guwung*, *suku kembang*, *gantungan la*, and *nania* for *ra*, *wa (ua)*, *la*, and *ya (ia)*, respectively. These appended forms cannot be tested independently, as described previously. Only case 6.62, from all of provided examples (case 6.58–6.62), was transliterated incorrectly by TAB, i.e. word *Briag* (intense).

On case 6.62, word *Briag* should be transliterated by stacking *cakra* and *nania* together. Using *ya* (rather than *ia*) at word *Briag* gave correct transliteration result. Word *Briag* and its variation (i.e. *Bryag*) refer to one meaning and should have same transliteration. They represent another kind of special words that there is still no research to know the precise list of them. The transliteration algorithm improvement on them is basically the same like special words at previous Section 4.1.

4.6. Akṣara swalalita

In addition to the eighteen basic syllables (see previous Section 4.1), nine *akṣara swalalita* are used for writing Kawi word. Several of them and provided examples were transliterated incorrectly by TAB, i.e. *ṅa*, *ṭa* (syllable *ta tawa*), *ṣa* (syllable *sa saga*), *śa* (syllable *sa sapa*), word *Gaṅitri* (chain), *Partha* (man's name), *Jaṭayu* (a bird), *Bhiśama* (decree), and *Ṣiwa* (God's name) at case 6.63, 6.66–68, 6.72, and 6.74–6.77, respectively.

On case 6.63, and 6.66–68, replacement syllables were provided by TAB, i.e. *ṅa*, *ṭa* (syllable *ta tawa*), *śa* (syllable *sa saga*), and *ṣa* (syllable *sa sapa*), respectively. The use of those replacement solve all of incorrectly transliterated words related to *akṣara swalalita*. Word *Gaṅitri* and its variation (*Ganitri*); *Jaṭayu* and its variation (*Jatayu*); *Bhiśama* and its variation (*Bhisama*); and *Ṣiwa* and its variation (*Siwa*) each refer to one meaning and should have same transliteration. They represent another kind of special words that there

is still no research to know the precise list of them. The transliteration algorithm improvement them is basically the same like special words at previous Section 4.1.

4.7. Sound killers

Four sound killers (*pangangge tengenan*), i.e. *cecek*, *surang*, *bisah*, and *adeg-adeg*, are used to end sound of a syllable and represent consonant *ng*, *r*, *h*, and others, respectively. *adeg-adeg* is the default sound killer that appears after a syllable (other than *nga*, *ra*, and *ha*). All of provided examples (case 6.79–6.83) were transliterated correctly by TAB, except word *Karṇa* (ear) at case 6.82.

On case 6.82, there is no *surang* was used for consonant *r*, the same as word *Airlangga* (a Javanese king) at case 6.49 (see previous Section 4.3), *Talĕr* (also) at case 6.56 (see previous Section 4.4) and Table 4, *Partha* (man's name) at case 6.73 (see previous Section 4.5), and *dīrgha* (long sound vowels) at Table 4. Inconsistency was shown since *surang* was only used for consonant *r* at the end of the word *Talĕr* even though it can appear anywhere.

4.8. Miscellaneous signs

Two miscellaneous signs, i.e. *ulu candra* and *ulu ricem*, are kind of sound killers (see previous Section 4.7) that are used to write Sanskrit words. They are used to end the sound of a syllable and represent consonant *ng* and *m*, respectively (at previous Section 4.7, their counterpart sound killer of *ng* and *m* is *cecek* and *adeg-adeg*, respectively). Their appended form of these syllables cannot be tested independently, as described previously at Research Method Section. All of provided examples were transliterated incorrectly by TAB, i.e. word *Mang* (a holly letter) and *Siddham* (perfect) at case 6.84 and 6.85, respectively.

On case 6.84, a replacement *m̄* was provided for consonant cluster *ng* of word *Mang*. Word *Mang* represent another kind of special words that there is still no research to know the precise list of them. The transliteration algorithm improvement on them is basically the same like special words at previous Section 4.1.

4.9. Holy symbol

When independent vowel *au kara* (see previous Section 4.3) met sound killer *ulu candra* (see previous Section 4.8), the Romanization is *Om*. *Om* was known as *Ongkara*, a holy symbol (*akṣara modre*). This word is used almost everywhere in the text, as it is the symbol of God Himself. All of provided examples (case 6.86–6.87) were transliterated incorrectly by TAB, i.e. phrase *Om Swastiasu* (May God blesses you) and *Om Ṣanti, Ṣanti, Ṣanti, Om* (May peace be everywhere) at case 6.86 and 6.87, respectively.

4.10. Miscellaneous syllables

The existence of two miscellaneous syllables, borrowed from Javanese, in Balinese script is very rare (case 6.88–6.89). Their appended form of these syllables cannot be tested independently, as described previously. On case 6.88, syllable *cha* was transliterated incorrectly by TAB. Syllable *cha* has no regular form. It is always paired with the normal form of syllable *ca*. On case 6.89, syllable *kha* was transliterated correctly by TAB.

4.11. Digits

All of the digits (0–9) were transliterated correctly by TAB (case 6.90–6.99).

4.12. Punctuations

TAB transliterated correctly all of punctuations (case 6.100–6.107). Comma, period, less-than, period-0-period, greater-than, double greater-than, and colon sign was transliterated correctly become *carik*, *carik pareren*, *panten*, *pasalinan*, *pamada*, *carik agung*, and *carik pamungkah* sign, respectively. Double quotes (case 6.107) has the same sign (“”). *Panten* is used at the beginning of a letter, a story, or a verse, while *pasalinan* is at the end of it. *Pamada* is used at the beginning of a religious text, while *carik agung* is at the end of it.

4.13. Some variation of usages

Some variation of usages includes: 1) Incorrect combination of independence vowel *a kara* (see previous Section 4.3) and vowel signs (see previous Section 4.2); 2) Special use of syllable *pa kapal* (see previous Section 4.6) that is never attached to *suku* or *suku ilut* (see previous Section 4.2); 3) Romanization of the inherent sound; and 4) The use of *pangangge akṣara* (see previous Section 4.1).

On third variation of usages, a stand-alone syllable has inherent sound that is always Romanized as *a* and it is common to a Balinese to pronounced an *a* at the end of a word as *ĕ*. On case 5.116, word *sekalĕ* (real) was transliterated incorrectly by TAB since there is transliteration difference with word *sekala*. Word

Sekala and its variations (*Sekalĕ*, *Sĕkala*, and *Sĕkalĕ*) refer to one meaning and should have same transliteration. On this case, they represent another kind of special word that there is still no research to know the precise list of them. The transliteration algorithm improvement on them is basically the same as described on special words at previous Section 4.1. On fourth variation of usages, several provided examples of case 6.117–6.124 were transliterated incorrectly by TAB, i.e. word *Sukĕsma* (thank you), *Kĕsatria* (warrior), *Smerti* (books of Veda), *Utama* (primary), and *Dharma* (religion) at case 6.118–6.119, and 6.121–6.123, respectively. On case 6.119, consonant *ś* of word *Kĕsatria* was transliterated incorrectly by TAB. This case is the same as word *Sukĕsma* at case 6.118. Another aspect, vowel cluster *ia* need to be written *iya* for correct transliteration. Word *Kĕsatria* and its variations (*Ksatria*, *Kĕsatriya*, and *Ksatriya*) refer to one meaning and should have same transliteration. On this case, they represent another kind of special words that there is still no research to know the precise list of them. The transliteration algorithm improvement on them is basically the same as described on special words at previous Section 4.1. On case 6.122, word *Utama* was transliterated incorrectly by TAB. It need to be written as *Utthama* for correct transliteration. Word *Utama* and its variation (*Utthama*) refer to one meaning and should have same transliteration. On this case, they represent another kind of special words that there is still no research to know the precise list of them. The transliteration algorithm improvement on them is basically the same as described on special words at previous Section 4.1.

4.14. Ligatures

Ligature as one pen strokes of two glyphs is desirable but not mandatory form. *Tedung* form ligatures with certain syllables. On case 6.125–6.146, *tedung ligatures* were transliterated correctly by TAB.

4.15. Abbreviations

Three different scheme for abbreviations can be used in Balinese, i.e.: 1) The one endorsed by the government to abbreviate government institutions. The scheme is to follow the way the abbreviation pronounced in Indonesian language; 2) The one used by Simpen [15]. The scheme is to use the first syllable with all the vowel signs attached to it; or if it is an independent vowel, then the independent vowel itself is used; and 3) The one less commonly used, but somehow the shortest one. The scheme is to use only syllable or independent vowel.

On case 6.147–5.150, phrase *Bank Pembangunan Daerah Bali* (Development Bank of Bali Province) and all of its three abbreviation schemes were transliterated incorrectly by TAB. Latin abbreviation of that phrase is *BPD Bali* and its three abbreviation schemes, i.e. *Be Pe De Bali*, *Ba Pe Da Bali*, and *Ba Pa Da Bali* at case 6.148, 6.149, and 6.150, respectively. On case 6.147, word *Bank* and *Pembangunan* each need to be written as *Bang* and *Pĕmbangunan* for correct transliteration. Since consonant *ng* has the same pronunciation with consonant *nk* (foreign sound), it was used for transliteration. Word *Bank* represent another kind of special words that there is still no research to know the precise list of them. The transliteration algorithm improvement on them is basically the same like special words at previous Section 4.1.

4.16. Word boundaries and line break rules

There are no spaces to separate words in Balinese script. In the old time of writing on dried palm leaves, spaces were scarce and there was common practice to break the sentence at any places. For modern writing, several rules applied, i.e.: 1) No line breaks allowed between syllable and any of its signs; and 2) No line breaks allowed just before a colon, comma, or full stop. On case 6.151, several words of sentence were transliterated incorrectly by TAB, i.e. *Akeh* (Many), *akĕsara* (alphabet), *luir* (i.e.), *ipun* (a pronoun to previous word *akĕsara*), *suara* (vowel), *wianjana* (consonant), *talĕr* (also), *dados* (become), *madrĕwe* (have), *kawāĕĕtanin* (called), *hrĕswa* (short vowel), and *dĕrgha* (long vowel).

On word *wianjana*, semi vowel *ia* was transliterated incorrectly except it was written as *ya* (see previous Section 4.5), as shown by Table 4. At another aspect, cluster *nj* was transliterated incorrectly by using gantungan *ja* on syllable *na*, instead on syllable *nga*. This is because there is assimilation combination on syllable *na* into syllable *nga* [21]. TAB provide replacement consonant *ñ* for consonant *n* of cluster *nj* for correct transliteration. Word *wianjana* and its variations (*wyanjana*, *wiañjana*, and *wyañjana*) refer to one meaning and should have same transliteration. They represent another kind of special words that there is still no research to know the precise list of them. The transliteration algorithm improvement on them is basically the same like special words at previous Section 4.1.

5. CONCLUSION AND FUTURE WORK

Accuracy analysis of Latin-to-Balinese script transliteration method on existing method of mobile application, i.e. *Transliterasi Aksara Bali (TAB)*, gave accuracy on TAB up to 68% (103 of 151) cases, based

on certain document of Balinese script writing rules and examples by Sudewa [5]. A Latin-to-Balinese script transliteration method can be improved significantly by taking care of thirteen kind of special words that was identified on the testing cases during the experiment. This research contributes on accuracy analysis and that recommendation for future development of Latin-to-Balinese script transliteration method. Based on authors knowledge, there is still no such analysis study in this research area.

ACKNOWLEDGEMENTS

This research work is supported by Indonesian Ministry of Research, Technology, and High Education through Grant No. 211/UN48.15/LT/2017.

REFERENCES

- [1] Antara, "Satu Hari Berbahasa Bali | The Indonesia News Agency," 08-Dec-2009. [Online]. Available: <http://bali.antaranews.com/berita/2756/satu-hari-berbahasa-bali>. [Accessed: 15-Jun-2017].
- [2] A. Stern, "An Community in Promoting the Balinese Language | BASAbali @ ARMA," 2017. [Online]. Available: <http://basabali.org/>. [Accessed: 15-Jun-2017].
- [3] N. S. Kharusi and A. Salman, "The English Transliteration of Place Names in Oman," *Journal of Academic and Applied Studies*, vol. 1, no. 3, pp. 1-27, 2011.
- [4] Agus Made, "Transliterasi Aksara Bali - Android Apps on Google Play," 2016. [Online]. Available: <https://play.google.com/store/apps/details?id=com.agusmade.aksarabali>. [Accessed: 05-Jul-2017].
- [5] Ida Bagus Adi Sudewa, "The Balinese Alphabet, v0.6 | Babad Bali," 2003. [Online]. Available: <http://www.babadbali.com/aksarabali/alphabet.htm>. [Accessed: 01-Apr-2017].
- [6] M. Eversen and I. M. Suatjana, "N2908: Proposal for Encoding the Balinese Script in the UCS," 2005.
- [7] ISO, "ISO/IEC International Standard ISO/IEC 10646: Information Technology - Universal Coded Character Set (UCS)," Switzerland, 2010.
- [8] A. A. K. O. Sudana, K. S. Wibawa, and I. M. A. D. Tirtha, "Learning Media of Balinese Script Writing Based on Augmented Reality," *Journal of Theoretical and Applied Information Technology*, vol. 90, no. 1, pp. 31-39, 2016.
- [9] I. D. A. M. Sartini, M. W. A. Kesiman, and I. G. M. Darmawiguna, "Pengembangan Text to Digital Image Converter Untuk Dokumen Aksara Bali," in *Seminar Nasional Pendidikan Teknik Informatika (Senapati)*, 2012, pp. 280-296.
- [10] Made Suatjana, "Bali Simbar | Babad Bali," 1999. [Online]. Available: <http://www.babadbali.com/aksarabali/balisimbar.htm>. [Accessed: 01-Apr-2017].
- [11] I. G. N. P. Arimbawa, M. W. A. Kesiman, and I. G. M. Darmawiguna, "Pengembangan Robot Penulis Karakter Aksara Bali Berbasis NXT-G dengan Lego Mindstorm NXT," in *Prosiding Seminar Nasional Pendidikan Teknik Informatika (Senapati)*, 2012, pp. 160-174.
- [12] Yayasan Bali Galang, "Babad Bali," 2000. [Online]. Available: <http://www.babadbali.com/>. [Accessed: 01-Apr-2017].
- [13] The Unicode Consortium, "Official Unicode Consortium Code Chart for Balinese." 2016.
- [14] I. W. Simpen, *Pasang Aksara Bali*. Denpasar: Upada Sastra, 1995.
- [15] I. W. Simpen, *Purwa Aksara*. Denpasar: Upada Sastra, 2002.
- [16] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, "Introduction to Algorithms (3rd ed.) Massachusetts Institute of Technology", Boston: Massachusetts Institute of Technology, 2009.
- [17] G. Indrawan, S. Akbar, and B. Sitohang, "Review of Sequential Access Method for Fingerprint Identification," *TELKOMNIKA (Telecommunication Computing, Electronics and Control)*, vol. 10, no. 2, pp. 199-206, Jun. 2012.
- [18] G. Indrawan, A. S. Nugroho, S. Akbar, and B. Sitohang, "A Multi-Threaded Fingerprint Direct-Access Strategy Using Local-Star-Structure-based Discriminator Features," *TELKOMNIKA (Telecommunication Computing, Electronics and Control)*, vol. 12, no. 5, pp. 4079-4090, 2014.
- [19] G. Indrawan, S. Akbar, and B. Sitohang, "Fingerprint Direct-Access Strategy Using Local-Star-Structure-based Discriminator Features: A Comparison Study," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 4, no. 5, pp. 817-829, 2014.
- [20] G. Indrawan, S. Akbar, and B. Sitohang, "On Analyzing of Fingerprint Direct-Access Strategies," in *International Conference on Data and Software Engineering (ICoDSE)*, 2014.
- [21] DisBud-Bali, *Pedoman Pasang Aksara Bali, Dinas Kebudayaan Provinsi Bali*. Denpasar: Department of Culture, Bali Province, 1997.

BIOGRAPHIES OF AUTHORS

Dr. I Nyoman Jampel, M.Pd., was born in Badung, October 10th, 1959. He worked as Rector at Ganesha University of Education, Bali, Indonesia. He received his doctoral degree in Educational Evaluation and Research from Jakarta State University, Indonesia. He also as Lecturer of Statistics, Methodology of Educational Research and Evaluation in various departments, including Department of Early Childhood Education, Department of Educational Technology, and Department of Guidance and Counseling. He can be reached at jampel@undiksha.ac.id.



Dr. G. Indrawan. Head of Computer Science Department, Graduate Program, Ganesha University of Education, Bali, Indonesia. He received his doctoral degree in Electrical Engineering and Informatics from Bandung Institute of Technology, Indonesia. His research interests include biometrics, pattern recognition, and robotics. He can be reached at gindrawan@undiksha.ac.id.



Dr. I Wayan Widiana, S.Pd., M.Pd., was born in Banjar Sari Tunas Tianyar village, Kubu district, Karangasem regency, Bali Province, on July 5th, 1985. He was received Doctor in Educational Evaluation and Research from Jakarta State University, Indonesia. He worked as Lecturer of Statistics, Methodology of Educational Research and Evaluation in Department of Primary Teacher Education at Ganesha University of Education, Bali, Indonesia. He can be reached at wayanwidiana85@undiksha.ac.id.