

REPORT

OF THE

TRANSLITERATION COMMITTEE.

(Adopted by Council, 8th May, 1894.)

At the General Meeting of this Society held on the 21st April, 1890, a resolution was passed, on the motion of Sir M. Monier-Williams, that this Society should, in co-operation with other Asiatic Societies, urge upon the next Congress the advisability of conferring on the possibility of formulating a uniform and international system of transliteration of Oriental alphabets. At the next meeting of the Council a committee (consisting of Sir M. Monier-Williams as Chairman, Professors Sayce, Bendall, and Rhys Davids, Dr. Rost, Dr. Thornton, Mr. Kay, and Mr. Lyon) was appointed to carry out this resolution. Owing, however, to the continued illness and absence from England of Sir M. Monier-Williams very little progress was made, and the work was not completed in time for the Ninth International Congress, held in London in the autumn of 1892.

In January, 1894, Col. Plunkett, R.E., and Dr. Gaster were added to the Committee. Col. Plunkett was elected Chairman, and frequent meetings have since been held in order to bring the matter before the consideration of the Tenth International Congress of Orientalists, to be held in Geneva in the autumn of 1894.

Dr. Thornton had been kind enough to procure official documents on the subject from India, and the following

letter had been despatched to the Secretaries of various foreign societies:—

ROYAL ASIATIC SOCIETY,
22, ALBEMARLE STREET,
LONDON, W.
21st Nov. 1890.

SIR,—In the present state of the comparative study of Philology, and of the history of religious beliefs and of institutions, it is becoming daily more and more important that the transliteration of Oriental proper names by scholars of different countries should become, as far as is possible, uniform. With a view of ascertaining how far this can be brought about this Society has it in contemplation to take some action at the next Congress of Orientalists. I have the honour, therefore, to ask you whether your Society would co-operate with our own in urging upon the Congress to take the matter into consideration.

The Committee appointed by the Council of this Society to deal with this question is quite aware of the difficulties which surround it. It proposes, therefore, in the first place to ascertain what the amount of divergence between the leading scholars of the various countries interested in Oriental research really is, and to confine the enquiry at present to two alphabets, the Sanskrit and the Arabic. With this object I have been requested to ask you to be so kind as to inform me whether your Society has adopted or recommended any system of transliteration for those alphabets,—if so, what that system is,—and if not, whether your Society, as a Society, could not go so far as to place on record, either in its own Journal or in a communication to this Society, the system which it regards as preferable.

I am, Sir,

Your obedient Servant,

T. W. RHYS DAVIDS,
Secretary.

To the Secretary of the

Society.

From the replies received from the American, French, German, and Italian Societies, it was clear that they were cordially in agreement that some action should be taken, and that while no unanimity of transliteration, especially as regards that of the Arabic alphabet, had been as yet arrived at in any one of the Societies' Journals, they would be willing to co-operate with our Society.

Your Committee therefore set to work to formulate such a scheme as might form the basis for argument at the next Congress.

I. SANSKRIT, PĀLI, ETC.

With regard to the transliteration of Sanskrit, Pāli, and the allied alphabets, their task was a comparatively easy one. A scheme for the transliteration of Sanskrit was adopted, on the proposal of the first President of this Society, Professor Horace Hayman Wilson, by the Bengal Asiatic Society so long ago as 1867. This scheme, with only one or two changes, has been very universally adopted by leading Sanskritists since that time, and it was adopted with very slight modifications by the Indian Government, by the Pāli Text Society, and also by this Society for use in the Catalogue of its Library (published only last year). The Committee would recommend, therefore, the following scheme for use in the Society's Journal. They have appended a few notes on the two or three symbols as to which they venture now to differ, for the reasons stated, from Prof. Wilson.

Indian Alphabets.

Scheme for the transliteration of all the alphabets in use in India, Ceylon, and Further India which are derived from the so-called Pāli Alphabet of Aṣoka's Inscriptions.

Vowels.					
a	ā	i	ī	u	ū
ṛ ¹	ṝ	ḷ	ḹ		
e	ē	ai	o	ō	au

Consonants.

k	kh	g	gh	ṅ ²
c ³	ch	j ³	jh	ñ
ṭ	ṭh	ḍ	ḍh	ṇ
t	th	d	dh	n
p	ph	b	bh	m
y	r	l	v	
ç ⁴	ṣ ⁴	s	h	

Specialities for Sanskrit—

The cerebral l consonant (ळ) is to be represented by ḷ.

Avagraha is to be represented by an apostrophe, thus सो ऽपि sô 'pi. *Visarga* is represented by ḥ, *Jihvāmūliya* by ḥ, *Upadhmanāya* by ḥ. *Anusvāra* is represented by ṁ, thus संसर्ग samsarga, and *anunāsika* by the sign ~ over the letter nasalized, thus अँ ā, आँ ā, and so on. The *udatta* accent is represented by the sign ' and the *scarita* by ^ . Thus, अग्निः agnih, जनिता janitā, क्वā kvā, कन्या kanyā. The *anudatta* accent may be represented by ` . Thus, ते अवर्धत tē dvardhanta.

Specialities for Pāli—

The cerebral l is ḷ. *Anuggahāta* is ṅ.

NOTES.

¹ Wilson has ṛi, rī, ḷi, ḷī for these four letters. But we ought not to have misleading compound symbols to represent the simple symbols of the native alphabets. The only practical objection to ṛ, r̄, ḷ, ḹ is that the ḹ may be preferred for the cerebral ḷ, and that it is impossible to write a long mark over ḹ. But ḹ will do well enough for the few cases in which it is necessary to write the lengthened ḷ. And for the cerebral ḷ, though ḹ would be the most appropriate, ḹ would be sufficient.

As to the long mark being throughout $\bar{\quad}$, and not $\acute{\quad}$ or $\grave{\quad}$ or $\hat{\quad}$, we trust it is scarcely necessary to point out the importance of reserving the European accents as symbols for accents, and adopting as the symbol for length a symbol used in European works on prosody to signify length.

² The representation of the symbol for the sound expressed in English by *ng* (as in 'sing') is open to doubt, as the symbols used in Europe for it are all either double symbols or letters used also as symbols for other sounds. A new symbol might be proposed (such as the η adopted by Prof. Rhys Davids), or n might be used with some mark. If the latter course be chosen \underline{n} or \dot{n} would do. But those who have corrected much transliteration for the Press know too well how often the \dot{n} , unless very carefully written, appears in the proof as i ($\dot{n}g$ as *irg* or *rig*, etc.). An additional reason for preferring \underline{n} to \dot{n} is that the diacritical mark is then under the line, as it is throughout the scheme for consonants. But the new sign η has the great advantage that it would obviate all difficulties and would avoid another diacritical mark; and being really an abbreviated *ng* would be at once easily intelligible. It would also be available for the new character which the Malays introduced into the Arabic alphabet (ξ) to represent this sound, so frequently occurring in their own language. [Since this note was written, the scheme of the Bengal Asiatic Society, in which the η has been adopted, has been received from Calcutta.]

³ Wilson has here, as was natural to an Englishman, *ch*. The sound is really, at least now, a double one (the $tʃ$ of this scheme, *tʃh* in English), but is always expressed in the Indian alphabets by a single symbol. When doubled as a symbol only the first letter, the t , of the compound $tʃ$ or *tʃh* is doubled in pronunciation (*at-tʃa*, *it-tʃa*, etc.). Wilson, for the doubled symbol, had to use *chch*, and when the second letter was aspirated he had to use *chchh* (!). c (with its logical results *ch*, *j*, *jh*) has been widely used already, and is adopted and recommended by Prof. Bühler, Prof.

Whitney, Prof. Cowell, the Bengal Asiatic Society, and the Pāli Text Society. It seems the most practical way of meeting the difficulty.

* There is a slight difference of opinion among scholars as to the representation of the three *s* symbols. The dental *s* (स) should, of course, be represented by the simple *s*. The sound expressed in English by *śh* (श) is pronounced in the same part of the mouth as the *ʃ* of this scheme; it is, therefore, the most systematic to represent it by *ʃ*. And probably *ʒ* or *ç* is the best symbol left for the third letter (ष). Of these two *ʒ* would be the more systematic, but *ç* is more familiar, and has been adopted by various scholars and by the Bengal Asiatic Society. Your Committee, therefore, would suggest *ç*.

In this scheme marks to consonants are placed, whenever possible, *under* the line, and long marks or accents to vowels *over* the line. This is of great practical importance in correcting proofs.

The sign *ś* would contravene this, and when used for transliterating accented words would certainly lead to confusion. The accent (') should be reserved as a symbol for accentuation. If the sign *ʒ* or *ç* be disliked the series *s*, *śh*, *s* might be adopted were it not for the objection, theoretically well founded, which many scholars feel against the *śh*. This last is the solution the Committee might else have recommended as really the least liable to mistake, and as diminishing the number of diacritical marks required.

II. ARABIC.

The transliteration of the Arabic alphabet was found more difficult. Your Committee studied the schemes in use in India, in the Journals of the various Asiatic Societies, and in the Catalogues of Arabic MSS. in the British Museum, in the Society's Library, and elsewhere. They beg to express their very cordial thanks for the assistance they have received in this respect from Dr. Codrington,

from Mr. B. Dé, of the Indian Civil Service, from Prof. Darmesteter, of Paris, and from Prof. Socin, of Leipzig.

As a result of a very careful consideration of each symbol your Committee have drawn up the annexed scheme, to which they have appended suggestions for general rules, and also notes on the isolated letters. The Committee has made every effort to have the whole scheme quite ready for this meeting of Council, but there still remain a very few details in the wording of these notes which require consideration.

Your Committee do not think it advisable for this Society to attempt at present to take any steps with regard to any other alphabets than those referred to in this report.

Your Committee would recommend—

- (1) That this Report be adopted.
- (2) That the draft Report be circulated among the members of the Transliteration Committee with authority to them to settle the final draft of the appendices to the scheme for the transliteration of Arabic symbols.
- (3) That the Report, so settled, be printed in the Society's Journal for July, 1894.
- (4) That a printed copy of the Report be forwarded to the President of the Tenth International Congress, with a request from this Society that the Congress should appoint a Committee, to sit during the Congress, to consider the Report, and to take such steps as they may deem desirable for the establishment of schemes, officially recognised by the Congress, for the transliteration (1) of all the various alphabets used in Sanskrit and Pāli MSS. and (2) of the Arabic alphabet.
- (5) That the Secretary forward printed copies of the Report to the Asiatic Societies who replied to his previous letter, and request those Societies to support officially (by letter to the President of the Congress) the suggestion put forward by this Society.

APPENDIX.

The attached tabular form shows that system of transliterating the Arabic Alphabet which is now proposed to the Society as being the most suitable for use in its Journal, and also the variations found in other systems. In deciding upon it the Committee have been guided by the following considerations:—

- (a) That as the Indian system has been so generally adopted it is undesirable to depart from it without cogent reason.
- (b) That it is desirable to use a single letter or character to represent one character of the original, *i.e.* to avoid (if possible) such combinations as *sh*, *ph*, etc.
- (c) That sufficient diacritical marks should be used to render it easy to transliterate a passage back into the original character.
- (d) That such letters shall be used that in case of diacritical marks and other such symbols being omitted (as must occur in newspaper reports, maps, etc.) the pronunciation of the words will not, to European ears, be materially altered.

NOTES TO THE TABLE.

1. Taking first the consonants as shown in the accompanying table it will be seen that as regards the following letters there is no difference of opinion ب د ه ز ي (as
b d h z y
consonant) ک ل م ن ف ر س ت and the Persian پ.
k l m n f r s t p

In Malay writing ق is used for p. It is proposed, however, that the پ should always be transliterated, and not omitted when silent at the end of a word, as has been customary in India.

2. As regards the following there is no important difference of opinion, the only point that is open to

argument being whether it is best to place the dot below or above the letter ح ص .

h s

3. As regards the g and j sounds represented by ج and the Persian گ there is some difficulty in deciding what is most suitable for both the Arabic and the Hindustani countries. As in the former the ج is in some countries a hard and in others a soft g this letter is the only one which will represent it satisfactorily, while for Persia and India, where it is always soft like j and the hard g is always represented by the گ , it is necessary to use the j. In Malay writing the g is written گ .

4. As regards و , the w alone would suffice, except in Turkish, but, as in this language it is necessary to show the v sound as well, there is no objection to using both as required. This use of alternative Roman letters to represent one character of the original can evidently cause no confusion when retransliterating into the Arabic alphabet.

5. As regards ط ظ and the Hindi ٹ ڈ it would be natural to use the one dot with the Arabic letter which runs through all the languages and to give the second dot to the letter which is brought in for the Hindi words only. But as no question of principle is involved it seems better to keep to the symbols already used in India.

t z

t d

6. As regards ش ث خ and the Persian ژ , which are

s t h d

z

sometimes called h compounds, the compound forms *sh, th,* etc., are cumbersome, and, moreover, do not, except to an Englishman, express the sound required. The mark suggested is the usual cursive method of writing in Arabic the three dots which distinguish some of these letters. The same would apply to the ج , for which c is recommended, but, as the c is not employed in any

other way, if the \angle be omitted and the simple e employed it will not mislead. The intention of the \angle with the e is, of course, to remind readers that the letter has not the s or k sound usual in Europe.

7. As regards ق there seems no doubt that the q is more convenient and expresses better the pronunciation of the original than a k with any diacritical mark.

8. The sound of غ has been represented in very different ways by different nations; in the Irish by gh , in the Welsh by rh , and by others as g . The modern Egyptians use غ to represent a European g , as for instance in writing the name Gordon غوردن , and the Hindoos pronounce it as a hard g . Therefore a ġ , distinguished as in the case of the h compounds, seems most convenient.

9. There is great difficulty as regards the ا , which, according to the Arabs, is a consonant; the Indian system of omitting it whenever it is not a letter of prolongation is not satisfactory, and the German system of representing it by a $'$ seems better. The Committee have, therefore, recommended this.

10. For the ع , which is in Arabic a distinct and well-marked consonant, the Committee would have liked to suggest a regular letter to run with the others, but believe this to be impossible; the letter itself would answer satisfactorily, but there is the insuperable objection that printing it would be impossible without the oriental type. A reversed comma ع is therefore recommended.

11. As regards the Turkish soft sound of k , which is really y , the most convenient way of representing this seems to be the addition of a cedilla, and this, if omitted, will not alter the spelling.

12. For the other Turkish sound of ن (the nasal n) a differentiated n is desirable, and therefore η is recommended. The Malays use the character ن to represent this sound, and it should therefore be transliterated by the η .

13. The equivalents for the four letters peculiar to Pahtū, as adopted in works printed in this country, is somewhat cumbersome. This is of less importance because the literature in this language is not extensive; but the Committee would like to obtain further information as to the systems used in the most recently printed works on the Continent before recommending any change.

14. As regards the taṣdīd, or sign of reduplication of a letter, in the transliteration the letter should be doubled, as is the custom hitherto. When the Arabic article al is used with words beginning with certain consonants the l of the article is in pronunciation changed into the following letter, which thus becomes doubled; there is great difference of opinion whether in transliteration the l should be preserved or whether it should be changed into the initial letter of the following word as pronounced. The Committee consider that the most satisfactory system is to retain the l and to connect it by the sign of equality (=) with the following word.

15. As regards the hamzah; where this occurs with ʾ alif at the beginning of a word the sign for the alif ' will be used, but in the middle of a word it is proposed to put a hyphen (-) in place of the hamzah whatever be the so-called "seat of the hamzah"; for instance

	I write 'aktab اكتب
a head ra-s راس	enclosure ḥā-it حائط
a wolf d'-ib ذئب	wolves dī-āb ذئاب
a chief ra-iš رئيس	peacock ṭā-ūs طاووس

16. As regards the vowels the Committee consider it undesirable to try, as some have done, to indicate all

those variations of the vowel sounds which often depend on their proximity to certain consonants and on dialectical peculiarities. So far as Arabic alone is concerned the representation of the three Arabic vowel marks, *fathah*, *kasrah*, and *dammah*, by *a*, *i*, and *u*, and of the corresponding long vowels by *ā*, *ī*, and *ū*, will suffice; in Hindustani it is necessary to use *e* and *o* in addition to the two latter, and when writing these sounds in the Arabic character a distinction is often shown by omitting the vowel points, i.e. the *kasrah* when *e* is to be used in place of *i* and the *dammah* when *o* is to be used in place of *ū*. If considered necessary there is no objection to using the *e* and *o* also in Arabic and Persian. The vowel sounds in Turkish present greater difficulties, and it seems necessary to use for the *fathah* *e* as well as *a*, and to adopt *ō* and *ū* in addition to the *o* and *ū*.

For the so-called diphthongs *اِي* and *اُو* the equivalents *ai* and *au*, which are generally used, seem the most suitable.

17. The Committee have carefully considered two suggestions made by Mr. Thomson-Lyon. The first of these is, in placing dots to distinguish letters which have to be represented by the same Roman letter, to use the same dots as are used with the Arabic letter, for instance to use for *ش* *â*, for *ذ* *î*, and so on: the advantage of thus having the Oriental form of letter suggested to the European reader is obvious.

The other suggestion is to represent those vowel sounds which are indicated in the original by diacritical marks by minuscules as *a*, *i*, and *u*, and to represent every letter of the original by a capital, so that the consonants will be transliterated by *B*, *G*, *D*, etc., and the long vowels

by A, I, U, etc. By this system the necessity for accents is avoided.

The Committee have not thought it necessary to deal with these suggestions further, as it is hoped that Mr. Thomson-Lyon will read a paper on the subject at the Congress.

[In the annexed table only those letters are given in the other schemes which differ from the proposed scheme.]