

## TRANSCRIPTION AND EXPLANATION OF THE SIAMESE ALPHABET

By F. OTTO SCHRADER

[*Preliminary note.* This article, dealing with the Siamese phonetic system, partly in comparison to Tibetan as the most archaic of Indo-Chinese languages, was written in the Internment Camp of Ahmednagar where the presence of quite a number of half-breeds using Siamese as their mother-tongue was an inducement to study that language. There is no reference in it to Lepsius' attempt at reconstruction (in his "Standard Alphabet") nor to F. Müller's (see Wiener Zeitschrift für die Kunde des Morgenlandes, 1893), because at the time of writing I knew nothing of these, while later I found them superseded by my own more extensive researches. The first penning underwent but one notable change, viz., by my discussing at greater length than originally my equation: Siamese "low *h*" = Tibetan "small *a*" = *spiritus asper* (wherein -- however maintaining the transcription *h* -- my friend H. Koerber has followed me in his publication "Morphologie des Tibetischen" now ready for the press). With Prof. Conrady's classical work "Eine indochinesische Causativ-Denominativ-Bildung" (also unknown to me at the time of writing), which I refer to subsequently several times, I am at variance in some points such as my derivation of the surd mediæ; with his transcription, however, only indirectly in that he follows Lepsius and (for the vowels) Pallegoix. — *This article uses from the outset the system of transcription dealt with in it.*]

In Siamese, as in other languages, the pronunciation came in conflict with tradition; for, Siamese is written even now just as it was pronounced half a millennium ago, though -- by sound-shift and other mutations -- it has changed to such a degree as to have the number of its consonants and semivowels dwindled down from 44 to 20.

In view of this fact it seems strange that Bishop Pallegoix's system of transcribing *modern Siamese* should have been again resorted to in a book<sup>1</sup> claimed by its author to be the first scientific introduction into the Siamese, while he, in his own words, is "fully aware of the unscientific character of this transcription." For, if it is not quite unimportant even for a practical learning of the Siamese to know why there are six different *khs* used in the script only, why four of these never occur in words with rising tone, etc., — it is, of course, quite indispensable for scientific studies not only to understand these differences but also to be constantly aware of them.

To find for the Siamese a scientific transliteration unobjectionable in every detail is, indeed, not possible yet, as a number of linguistic problems (esp. for the diphthongs) have to be solved before. But linguistics, especially comparative linguistics, cannot work without a system of transcription which essentially reproduces that pronunciation on which the script is based; and this essay is to prove that such a system may be created even now and without overmuch difficulty.

## 1. CONSONANTS.

For the larger part of the consonants the method of transliteration results quite naturally from the Siamese classification of consonants as "high", "middle", and "low" letters. For, this classification — as is clear from the spelling, as compared with the actual pronunciation, of Sanskrit and Pāli words used in Siamese — is but an expression of the historical fact that in the tonal system now in use 1) all *originally* (*i. e.*, at the time of the introduction of the alphabet) *sonant* consonants (including nasals and semivowels) appear as "low" letters; 2) all *originally non-aspirate and surd* explosives (including the spiritus lenis) are "middle"; and 3) all the *remaining surd* explosives and other sounds are "high" letters. The class of "high" letters, therefore, comprises all (originally and now) *surd sibilants, h-sounds*<sup>2</sup> and aspirate surds, as also *f*.

By taking notice of this scheme and by transcribing all sounds borrowed from Sanskrit or Pāli or common to these and Siamese in

the way long since usual in Indology we obtain the following transcription of consonants.<sup>3</sup>

A. *Foreign Consonants, i. e.* such which do not occur in purely Siamese words. These appear to be the following believed to have been introduced into Siamese from Sanskrit or Pāli (the three divisions are those of the tonal system: high, middle, low):

ṭ	ṣ	ṣ	ḍ	ḍ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ
ṭh	ṣ	ṣ	ḍ	ḍ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ	ḡ
(th)	(s)	(s)	(d)	(Ṭ)	(kh)	(ch)	(y)	(th)	(th)	(n)	(th)	(ph)	(l)

But of these fourteen sounds five only, *viz.* *ṭh, ṣ, ṣ, ḍ, ḍ*, prove to be really confined to Sanskrit and Pāli words, while the remaining nine, though more frequent in such words than in others, occur also in Further-Indian and Chinese loan-words and *ḡ* apparently even in a few purely Siamese words.

[The following, among others, must presumably be considered as non-Indogermanic loan-words in Siamese: *ṣṭ* neck (term of respect), *ṣṭik* war, *ṣṭok* enemy;<sup>2</sup> *ḡḡ* bush, twig (Cambodj.), *ḡḡ* to stare (~ Annam. *xem* to see) and (in spite of Pāli *ḡḡ* burnt) *ḡḡ* to burn; *ḡḡ*, two (Chin.), *ḡḡ* Peguan, *ḡḡ* unhappy, miserable (~ Annam. *ḡḡ* light, to contempt); *ḡḡ* old; *ḡḡ* at, near, *ḡḡ* in day-time (~ Annam. *ḡḡ* day, sun); *ḡḡ* donkey<sup>3</sup> (~ Annam. *ḡḡ* mule), *ḡḡ*, *ḡḡ* to deceive (Annam. *ḡḡ* to be mistaken, a mistake). Again, it being by no means likely that the Siamese should have preserved in a small number of words an original sonant aspirate, the few non-Aryan words containing *ḡḡ, ḡḡ, ḡḡ*, which appear, indeed, to be Indo-Chinese, may also for the present be considered as originally foreign to Siamese.<sup>4</sup> But there is no reason why *ḡḡ* in the group *ḡḡ-* should not be purely Siamese. For, although *ḡḡ*, "big, older" reminds one of Annam.

<sup>1</sup> The bracketed letters refer (in terms of our system) to the present pronunciation wherever this differs from the original one. For Ṭ, etc., see below the discussion of *voiceles medix*.

<sup>2</sup> These words must have been introduced into Siamese after the shifting, in the latter, of ṣ to s (see below).

<sup>3</sup> Beside which we have, with ordinary *ḡḡ*, *ḡḡ* mule, which may be purely Siamese.

<sup>4</sup> For these and their correspondences see Conrady, *Caus.-Denom.-Bildung*, p. 119. As a sixth we may add *ḡḡ* to weave (also *ḡḡ*) = Tibetan *ḡḡ-pa*. The pronoun of the second person *ḡḡ, ḡḡ* (cf. Burm. *ḡḡ*, Tib. *de*) is restricted to court-parlance and thus may be also of foreign origin.

<sup>1</sup> Frankfurter, *Elements of Siamese Grammar*, Leipzig 1900.

<sup>2</sup> "High ḡ" (inclusive of "leading ḡ") and "leading ḡ" (cf. below p. 66 end.).

*ngài* "Your Honour",<sup>1</sup> still *hñiy* "woman" is evidently connected with Tib. *gñen-za* "wife", *hñā*, "grass" with Tib. *rya-* "hair (of animals)", etc., so that, to all appearances, *ñ* is a genuine Indo-Chinese sound preserved in Tibetan.]

We, therefore, cancel *ñ* from the list of foreign consonants, and enter it among the original sounds.

To the foreign consonants belongs also *ṭ*, a "low" letter now spoken *kh* and suitably transcribed by an underlined *g*, its sign resembling that of *g* *ṭ*. It occurs in a single word only, viz. *gṭ* "neck", which is said to be Annamite.<sup>2</sup>

There is, finally, the letter *ṣ* which I should tentatively class among the foreign consonants as long as we know nothing of its origin. It occurs in just a few Pāli words (in the place of *kh* < *kṣ*, *kṣy*) and in about a dozen non-Aryan words and is now written and also pronounced as *ṣ* *kh*. It may, therefore, be transcribed by *kh*. Part of the words containing it are undoubtedly Indo-Chinese, e. g. *khāṣ*, "spinning top" (~ Tib. *khāl-ba*, *b-kal* "to spin"), *khān* (*khñ*, *khān*) "robust, strong" (~ Burm. *khvan ā*, pronounced *khon ā*, "strength"); but *khṣ* "moon", according to Michell, comes from Cambodjan.

B. *Purely Siamese Consonants* (partly identical with those of Sanskrit, etc.). The following list includes also those of dubious transcription to be discussed hereafter.

1. "High letters" (akṣor<sup>3</sup> sūṅ):

ṣ	ṭ	ṇ	Ṃ	ṃ	ṣ	ṭ
kh	ch	th	ph	f	s	h

These have not changed their pronunciation.

<sup>1</sup> For the equivalence or shifting resp. of *ñ* and *ng* (*y*) in Annamite compare e. g. *nhānh* = *ngānh* "twig, branch".

<sup>2</sup> Compare, with ordinary *g*, *gṭ* "neck of animals", which seems to be the original Siamese word for "neck" in general beside which as terms of respect the above *gṭ* as well as *ṣṭ* were introduced. There seems to be no reason why these words should not be Indo-Chinese (cf. Tib. *m-gu-bar*, *m-gur*, *m-gul-pa*), while, on account of its initial and of its normal tone, *gṭ* cannot be derived from the modern Annamite word for "neck" viz. *có* (i. e. *kō* with rising tone). Compare also the Āhom word *kho* and Shān *khō*, both = "neck".

<sup>3</sup> Pronounced 'akṣōu; compare note 2 on page 59.

2. "Middle letters" (akṣor klāṅ):

ṅ	ṅ	ṅ	ṅ	ṅ	ṅ	ṅ
k	c	t	ṭ	p	p	'
	(j)	(d)		(b)		

3. "Low letters" (akṣor tā):

ṅ	ṅ	ṅ	ṅ	ṅ	ṅ	ṅ	ṅ	ṅ	ṅ	ṅ	ṅ	ṅ	ṅ	ṅ
g	ṅ	j	z	ñ	d	n	b	w	m	y	r	l	v	'
(kh)		(ch)	(s)	(y)	(th)		(ph)	(f)						(h)

Let us now examine, as far as necessary, the particulars of this table.

The transcription of the *palatal explosives* by *ch*, *c*, *j* (and *jh*, see table A) may be objected to on account of their being compounds. But transcriptions like Wershoven's *tch*, *dj* are also unsatisfactory for the one among other reasons that it may well be questioned whether these sounds are not, or at least originally were, *simple* palatals like the Sanskrit sounds with which they are identified by the Siamese themselves.<sup>1</sup> For simplicity's sake too the transcription offering itself should be kept.

Excepting the *ch* just dealt with, the transcription of all "high letters" is clear.

The "middle letters" require throughout an explanation.

Here we have first of all to note the remarkable phenomenon of voiced and voiceless media<sup>2</sup> existing side by side in the actual pronunciation — the more remarkable for being confined to the dentals (with the cacuminals<sup>3</sup>) and the labials<sup>4</sup>. That both sounds — the

<sup>1</sup> It is true that the Sanskrit palatals are now-a-days often pronounced as compound sounds — so always in Europe, where English *church*, *jew*, etc. are compared, and in Kashmere where they have turned into dental affricates — but in South India, at any rate, they are still the simple sounds which comparative philology and Indian metrics proves them to be.

<sup>2</sup> The latter are said to differ from the former not only by lacking the voice but also by a somewhat stronger pressure of the evading air (Wershoven, *Lehr- und Lesebuch der Siamesischen Sprache*, p. 4). They are pretty frequent also in Europe, e. g. in Middle German dialects, in the place of, but not beside, the voiced media.

<sup>3</sup> Cf. *t*, *t'* of table A. These cacuminals not natural to the Siamese have most probably always been pronounced by them as dentals.

<sup>4</sup> For these sounds and their pronunciation in the neighbouring *Mon* see P. W. Schmidt, *Grundzüge einer Lautlehre der Mon-Khmer-Sprachen* (in *Denkschriften der Wiener Akademie der Wissenschaften*, Phil.-hist. Klasse, 1906), esp. §§ 137 and 143. Of

one written as surd but pronounced as sonant (*t, t, p*) and the voiceless one (*r, r, r*) — were tenues before, may be inferred from their having been used for writing the Sanskrit tenues. For, we find *e. g.* *tāpa* "heat", *tojk* "mason's chisel", *pori* "all around", *pitar* "father"; but also: *ta-thāgor* "the Exalted one" (Buddha), *toykit* "chiseling, carving (of stone)", *pra-bha* "brightness", *pita* "father". This, however, may be accounted for by the fact that, after the tenues (*t, t, p*) had become voiced mediæ in pronunciation, while the "low" sounds *d, d, b* had turned into the "high" *th, ph*, nothing but the voiceless mediæ (*r, r, r*) appeared to be left to the Siamese for writing the tenues in newly introduced Sanskrit and Pāli words. That these mediæ also had shifted *within Siamese* from original tenues, cannot be substantiated; but there is reason for supposing that their origin in *pre-Siamese* time is connected with certain consonantal groups containing a tenuis.

[For, as is evident from the following instances, the Siamese voiceless mediæ correspond with certain Tibetan consonantal groups or their first constituent. This is the third of three cases of correspondence: 1) *Tibetan tenuis = Siamese tenuis*: Tib. *tol-ba* = Siam. *tol* to reach, arrive; *g-tiy* bottom, depth = *tiy*, deeply (adv.); *d-pal* brightness, prosperity ~ *pol* to sacrifice, wish prosperity; *d-pon* master, superintendent ~ *pon* to order, superintend. 2) *Voiced mediæ on both sides*: Tib. *day* = Siam. *day*, with, also; *r-dap-ba* = *dop* to fold; *b-de* bliss, durability, *b-den* true ~ *dē*, genuine, true; *bun-loy* whirling up and down ~ *buŋ*, to cast, fling; *d-baŋ-po* powerful ~ *bēŋ*, strong, strength. 3) *Voiceless mediæ in Siamese*: Tib. *s-tau* seat, stool ~ Siam. *tau*, small bench; *s-ti-ba* to rest ~ *tiy* quiet, silent; *'dam* = *Tom* mud; *draŋ* = *Troy* straight; and with labial (for *sp > f*, etc., see below the discussion of *w*): Tib. *phra* thin, fine ~ Siam. *pro* fragile, *'phro-ba* and *'brab-pa* to strew (*s-pro-ba* to spread) ~ *prōi* and *prāi* to strew; *'bri-ba* to decrease, grow less (*phri-ba*, *'phri-ba*, *s-brid-pa* to diminish) ~ *prī*, receding water, *prē* sloping ground; *'byed-pa* = *pōt* to open. — If in a few cases we do not find the expected voiceless mediæ (comp., *e. g.*, Tib. *s-tab-s* = Siam. *tab* opportunity, *'brab-ba* ~

the few words containing these sounds (transcribed by Schmidt as *β, d*) some correspond with the Siamese, *e. g.* *βan* "to plait" = Siam. *pan*, "to turn, twist", while others are purely Mon-Khmer (not Indo-Chinese), *e. g.* *βā* "two" = *barea, bar*, etc. of the Mundā languages.

*blat, brat, brāk* to be separated), there need not be a direct correspondence here. For, we must bear in mind that Siamese has been influenced by many a dialect related to it, such as Khāmī, Shān, Āhom, etc.; also that here as in Tibetan numerous adaptations have taken place (comp., *e. g.*, Siam. *tū*, *rū*, *thū*, *dū*, blunt, stupid ~ Tib. *g-ti-mug* stupidity, *thub-s-po* dark, *di-hi* drunk); further the rank growth of prefixes in Tibetan; and, finally, the loss, on one or the other side, of words with an original prefix.]

The transcription of these sounds must keep account of the fact that in the Siamese script the voiceless mediæ is but a modified tenuis ( $\left( \begin{array}{c} \text{ᨾ} \\ \text{ᨿ} \end{array} \right) < \left( \begin{array}{c} \text{ᨾ} \\ \text{ᨿ} \end{array} \right) \right)$ ,<sup>1</sup> which can be indicated by underlining, but better, the case being different from *k* and *g*, by small capitals not overtopping the smaller ones of the minuscules, *i. e.*, *r, r, r*.

With these pairs of sounds does not go a guttural pair, as should be expected, but instead of it we have the single ᨾ, in which are united, as is shown by Tibetan<sup>2</sup>, the original<sup>3</sup> guttural tenuis and a later non-āspirate and voiceless guttural which, in analogy with the discussed pairs of sounds, would have to be pronounced as voiceless mediæ (*x*). But the two have become one both in speech and script, and the one sound, as it is spoken now, neither corresponds with our *k* nor with our voiced *g*, nor is it exactly, as appears to me and others, what might be expected here in analogy with the *r* and *p*. One might, therefore, feel tempted to render it in some special way, but *k* may be kept just as well.<sup>4</sup>

<sup>1</sup> This means that from the Siamese point of view these mediæ are more akin to the tenues than to the voiced mediæ. Compare with this Sütterlin's calling the tenues "strong voiceless explosives" and the voiceless mediæ "weak voiceless explosives" (Die Lehre von der Lautbildung, p. 136).

<sup>2</sup> For instance, there being in Siamese neither *pr* nor *phr* but only *pr*, we should expect to find for Tib. *khrog* "to pour down the throat, to swallow" a Siamese word with voiceless guttural mediæ as initial; in point of fact, however, we find *krōk* "to swallow" with the very *k* which occurs as original tenuis *e. g.* in the emphatic particle Tib. *ka* = Siam. *kō*, in Tib. *b-kab-pa* "to cover" ~ Siam. *kāp* "sheath, shell", etc.

<sup>3</sup> Whether in an earlier stage, before the differentiations of individual languages, the Indo-Chinese tenues sprang from mediæ, is of no account for this research.

<sup>4</sup> The sound appears to me to be a voiceless mediæ which, unlike *r* and *p*, is not produced at exactly the same place as the corresponding voiced mediæ. Wershoven puts it together with *r* and *p*, while Frankfurter sees in it, as also in *r* and *p*, a tenuis, saying that the Siamese had introduced ᨾ (see above p. 48) in order to obtain a pair of gutturals corresponding with *t-r* and *p-r*. But ᨾ is a "low" letter, not to speak

There is also no palatal pair corresponding with *t*, *t'* and *p*, *p'*, and I doubt whether there ever was one.

One more of the "middle letters" remains to be mentioned, viz., *ṅ* which as "middle letter" (for it has some more functions to be dealt with in the chapter on vowels) denotes the cracking start of the voice in words beginning with a vowel, i. e., the *spiritus lenis* familiar to us from Ancient Greek<sup>1</sup>, because in other European languages, such as English, it is not written, though spoken. This being reckoned as an individual letter by the Siamese, it must also be transcribed as such (ṅ), which is not very cumbrous and is, as will be seen, advisable for other reasons too.

Among the "low letters" there are only three the transcription of which requires a discussion.

We must, first of all, justify our use of *z* for *ṅ*. Siamese grammarians count this letter among the palatals (compare its place in the alphabet), so that, if they were right, we ought to transcribe it by *ḷ* (= *j* in French *journal*, *s* in Engl. *tension*)<sup>2</sup> and not by *z* (*s* in *rose*, etc.). That they are not right, but have — as is easily understood — mistaken the nature of the sound, becomes evident as soon as we inquire into the fate, in Siamese, of the two old couples of sounds preserved as the Tibetan sibilants *ś*, *ṣ* (palatal) and *s*, *z* (dental). For, the first thing this inquiry reveals is the fact that for the Tibetan *ś* we have *s* in Siamese;<sup>3</sup> in other words: the pre-Siamese *ś* has been absorbed in Siamese by *s*. This suggests the possibility of a parallel, may be coinciding, process for *ḷ* and *z*, a supposition apparently endorsed by the remark with which Sarat Chandra Das in his Tibetan-English Dictionary introduces the *ṅ* (*z*): "In olden times and in the frontier-provinces to the present day it was and is sounded like the

of other objections. Frankfurter contradicts himself by explaining *T*, *T'*, *P* on p. 7 of his book as "original sonants", but on p. 19 as "originally surd".

<sup>1</sup> Where, however, it may have better accorded with its name.

<sup>2</sup> It is advisable — for consistency's sake — to characterize all palatal consonants, except where no diacriticon is needed, by the acute accent (also all cacuminals by a point under the letter) and to make use of *v* only in urgent cases. Still *ṅ* may also be kept in the sense of *ñ + j*, i. e., for distinguishing the compound from the simple palatal nasal (it is used in this article without regard to this difference).

<sup>3</sup> Compare, e. g., Tib. *śom-pa* "to prepare, make ready" with Siam. *som* do., Tib. *śin-tu* "very, in a high degree" with Siam. *sin<sub>2</sub>* "all, every, end, to come to an end", Tib. *g-sag-pa* "difference, to distinguish" with Siam. *sph* "to divide, classify".

English *z*; but in Lhasa and Tsang it is now pronounced like the English *s*, but always low-toned." In Siamese too the actual pronunciation is *s* which, however, cannot be directly derived from *ḷ*. But in Central Tibet (acc. to S. Ch. D.) *ṅ* (*ḷ*) is now pronounced something like *ṅ* (*ḷ*); and thus, indeed — all *ḷ* having become *s* in Siamese — the possibility is given of pre- or early-Siamese *ḷ* having in the long run developed to *s* (*ḷ* > *ś* > *s*)<sup>1</sup>. And this hypothesis, viz., that both *ḷ* and *ś* have become *s* in Siamese, seems to be confirmed by the facts,<sup>2</sup> so that the few cases in which Tib. *ḷ* seems to correspond with Siam. *z*<sup>3</sup> presumably are no direct correspondences. This much is certain that the only Tibetan sound corresponding with Siam. *z* in pretty many cases is just Tib. *z* (not *ḷ*)<sup>4</sup>.

For the letter *W*, now called "low *f*" (as distinguished from the "high" original *f*), there remains as transcription only *w*, because *ṅ* is considered with the Siamese as equivalent to Sanskrit *ṅ* and must, therefore, be transcribed (as in Indology) by *v*. It is true that in this way the phonetic relation of the two is inverted in so far as *ṅ* is bilabial like the English *w*; still the use of *w* as labio-dental sonant fricative (as in German) has the practical advantage that the graphically simpler sign *v* is used to denote the simpler and far more common sound. This *w* as well as its "high" counterpart, the *f*, both alien to Tibetan, have, in my opinion, formed like the voiceless medials, viz., *w* from *śb* (*'b*) and *f* from *śp* (*'p*)<sup>5</sup>.

<sup>1</sup> I. e., a low-toned consonant to a high-toned one. But we must remember that, when this change took place (if it is a fact at all), the tone-system was still non-existent or only in an embryonic state.

<sup>2</sup> Compare Tib. *b-ti* "four" = Siam. *si<sub>1</sub>*; *śib-pa* "exact", *śib-ter* "close examination, comparison" ~ *śiip* *sq* "to investigate, examine"; *śiig-s* "fire" (term of respect) ~ *śuk* "bright, shining". Many more examples will hardly be found, as most of the few Siamese roots with initial *ś* seem to be lost.

<sup>3</sup> Thus, apparently, in Tib. *śud-pa* "grown lean, to fade away" ~ Siam. *śut* "to shrink, sinking"; *śu-ba* "to ask for, beg" ~ *zō* "beggar".

<sup>4</sup> Compare Tib. *g-zim-pa* "to sleep, fall asleep" ~ Siam. *zim* "nearly sleeping, sleepy"; *zab-s* "thick", *b-zab-s* "abundance, lots" ~ *zap* *sqn<sub>2</sub>* "to heap, too much"; *zam* "line, continuation, sequel" ~ *zam* "often, to repeat"; *b-zo-bu* "to make, manufacture", *b-zom-pa* "diligence" ~ *zpm<sub>2</sub>* "to prepare, make ready, practise"; *g-zu-bo* "straight, honest" ~ *zū<sub>1</sub>* do.; *g-sig* "atom, particle" ~ *zē* "part, piece, half"; *zol* "sly, false" ~ *zou* "naughty, roguish"; etc.

<sup>5</sup> Compare Tib. *ś-bo-ba* "to swell (of the belly, etc.)", *'bo-ba* "to swell, heave" ~ Siam. *wō<sub>2</sub>* "swelling of the belly"; *ś-bud-pa* "bellows", *'bud-pa* "to blow" ~ *wut* *wōt* "puffing noise (as of an automobile)"; *ś-ba* "secret, hidden", ~ *wō<sub>2</sub>* "sky (in the sense

The last letter of the alphabet is the low *h*, ཨ. Its position in the alphabet, where it comes after the ཨ or *spiritus lenis*, and its graphic similarity to the latter show that it is considered as related to it, *i. e.*, as another kind of *spiritus*. For this reason and a rather strange one<sup>1</sup> Michell (in the preface, well worth reading, to his Siamese-English Dictionary, Bangkok 1892) declared this *h* to be the *spiritus asper*, and a comparison with Tibetan seems to confirm his view. For, it shows the identity not only (which is obvious) of Tib. ཨ with Siam. ཨ,<sup>2</sup> but also of Tib. ཨ with Siam. ཨ (both of which latter we render by '):

Tib. 'ug-pa "owl" ~ Siam. 'āk do.; T. 'o-dod "lamentation" ~ S. 'ō "sound of weeping"; T. 'u-cag "chimney" ~ S. (Chin.) 'ue "fire"; T. 'os-pa "to be worthy; fitting" ~ S. (Chin.) 'ē<sub>2</sub> "good". — To this we might add a number of words expressing various kinds of noise.

This being the most enigmatical letter of the Tibetan alphabet and, moreover, of very frequent occurrence in that language, we shall now try to obtain some more information about it.

To draw any conclusions, as to its original sound, from the Chinese loan-words in Siamese<sup>4</sup> is evidently impossible, as there is but one *h* in Chinese which, like the Indian *h* (*intū* "Hindoo", but *hima* "snow"), is rendered in Siamese partly through the "low", partly through the "high" *h*.<sup>5</sup>

of cover of the world), wāṅ 1) "straw" (= "husk"), 2) "indistinct, dull" (= "veiled"); 'boṅ-ba "roundness" ~ wāṅ-n class prefix for eggs; — and with *p*: Tib. *s-pu* "hair" ~ Siam. *fpi* "down-feather, fine, thin", *fun* "fine dust"; *s-pyi* "general, concerning all" ~ *fī* Pāk "eloquence" (lit.: "swell-mouth"); *s-puṅ-s* "many" ~ *fün* "multitude", class prefix for crowds, herds, etc.; *s-paṅ-s-pa* "to repair to, transfer to" ~ *fak fāi*, "to give oneself up to", *fāk* "to send"; *s-paṅ leb* "shield", *s-paṅ khebs* "apron" ~ *fāṅ* "to hide, to screen"; etc.

<sup>1</sup> Because the Siamese write the word *hôtel* with "low *h*"! This is, of course, done because the word, if written with "high *h*", would not have the normal tone.

<sup>2</sup> Comp., *e. g.*, Tib. 'anan "little man, dwarf" with Siam. 'anā "within bounds".

<sup>3</sup> In the Tibetan alphabet ཨ follows *s* and precedes the semivowels; these are followed by the voiceless sibilants, then *h* and, as the last letter of the alphabet, the *spiritus lenis*. This place of ཨ is probably due to its nature of a sonant fricative.

<sup>4</sup> In addition to the two words mentioned above Cartwright's Dictionary terms as Chinese: 'kē "quick, impetuous", huōṅ "wind", and huōṅ<sub>2</sub> "to fight". For the first comp. Siam. *yiē* "fast" and *hik* "hold" (Chin. *hik* "already"?), for the second Chin. *fün* "wind".

<sup>5</sup> This difference, if not a mere inconsistency (cf. p. 58 n. 3, p. 62 n. 1), may have something to do with the tonal system; cf. p. 54 n. 1.

Nor do we get much farther by Siamese alone, for here also the language, as it is now spoken, knows only one *h*, the difference between the "high" and the "low" *h* consisting merely in their being connected with differently toned vowels. Yet it remains noteworthy that here "low *h*" is not only graphically related to the *spiritus lenis* but is also connected, through a change of sounds similar to that in the dental series (cf. *h<sub>2</sub>* etc., above p. 51), with both the *spiritus lenis* and the "high *h*", for which reason these three (and, fourthly, *l* which, however, has hardly anything to do with them) form one class of sounds in indigenous Siamese grammar. Thus we find in Siamese: 'ōk "to bellow at; roaring of the tiger" besides *hōk*, the sound of vomiting, and 'uak<sub>2</sub> (ཨ᳚) "to feel sick, to vomit"; further 'at "to sneeze" besides 'at "to be short of breath". Compare also Tib. 'ud "to boast, boasting" ~ Siam. 'uat "to boast"; Tib. 'ab-pa "to bark" ~ Siam. hao, the same; etc.

It is, after all, from Tibetan that we learn most about its "small *a*" (*a chun*), as the letter is called here, presumably for graphical reasons only:

The Tibetan ཨ—a "[sub-]feminine", *i. e.*, [very] "low" letter—is, according to Jäschke's detailed description,<sup>1</sup> spoken now as follows, "changing like the shades of a chameleon": 1) in all parts of the country when intervocalic, and in Central Tibet in other positions too (but not before surd aspirates, where it is mute) as the "vowel in itself" (see below); 2) in Western Tibet, when initial before a vowel, as the *spiritus lenis*, while as initial before a consonant it is mute; 3) in Khams (Eastern Tibet), when initial before a vowel, as a voiced fricative guttural (γ); when initial before a consonant, as the nasal of the consonantal class concerned; to which may be added that in compounds the nasal pronunciation is rather frequent also in remaining Tibet (*e. g.*, *dge-* 'dun > *gendun*).

Now, that this ཨ originally was a kind of *h*, as which it had already been transcribed by Csoma de Kőrös, may be proved by Jäschke's own words, in spite of his opposition. He sees in the original ཨ the voiced fricative corresponding with the (explosive) *spiritus lenis*, *i. e.* the "vowel in itself", the "purely vocalic sound freed from all

<sup>1</sup> In the Monatsberichte der Königl. Preuß. Akademie der Wissenschaften, Berlin 1868, p. 158 ff., and in the Introduction to his dictionaries.

consonantic nature" "well-know when occurring within a word, as it appears wherever in non-compound words one vowel follows another, as in *deo*, Tib. རྩོམ་". It is, therefore, *not* a question of the "vowel in itself" (such a thing does not exist), *not* one of some weak vocalic element preceding every initial media,<sup>1</sup> but what we have to do with here is the "voiced *h*" we all know from its intervocalic occurrence in French *fléau*, *là-haut*, German *Alkohol*, *sie hat*, etc.<sup>2</sup> This agrees perfectly with what an eager scholar of Tibetan<sup>3</sup> writes to me from Calcutta about the pre-vocal རྩོ (as distinguished from the pre-consonant རྩོ which he declares to be completely mute): "The Tibetan initial རྩོ is not a vowel or semi-vowel but a soft *h*. I do not know enough about phonetics to know whether such a letter exists, but I should say that the initial རྩོ stands to ཨ<sup>4</sup> as *g* to *k*, or *d* to *t* . . . The European, and even the Tibetan, easily pronounces it too softly, so that it becomes inaudible, or too hard when it turns into *h*." The latter case has happened in Siamese, just as in Sanskrit where *h*, long since voiceless in pronunciation, has yet retained the value of the voiced sound which comparative philology proves it to have been.

To transcribe this voiced *h* by the *spiritus asper* is all the more justified as its secondary derivation from *y* may still be observed in Tibetan, e. g., in *kyan* > *yan* > 'ay "also",<sup>5</sup> which is the very process by which the Indo-European semivowel *y* became in Greek the *spiritus asper*, as in *δs* = Skt. *yas*. Vice versa, the vulgar *y* in the place of ' , e. g. in *yog* < 'og "below, down",<sup>6</sup> tallies with the intervocalic "weak" *y* of Jaina Prakrit.<sup>7</sup>

From all this it seems to be evident that the antevocal རྩོ of

<sup>1</sup> Compare the rendering of the mediae as nasal + tenuis in Modern Greek, and the like (J. Kirste, Die altindischen Platten, Vienna Acad., Phil.-Hist. class, vol. 160, first essay, p. 8). Jäschke explains the sound to be „die mit summendem Anlaut hörbare Anfüllung des am Cuntural-, Labial- oder Dentalpunkte geschlossenen Luftweges" as found in the initial of English *go*, *do*, *be*, *jew*, while the German mediae are nothing to him but "soft tenues", (Handwörterbuch, preface p. 1v).

<sup>2</sup> Jespersen, Elementarbuch der Phonetik (Leipzig-Berlin 1912) 6, 52f.

<sup>3</sup> My friend and former colleague Johan van Manen, now secretary of the Asiatic Society, Calcutta.

<sup>4</sup> The "high" *h* of Tibetan.

<sup>5</sup> All three forms are in use, the third after vowels.

<sup>6</sup> Also pronounced *nos* (*wog*). Compare also (?) *yoy-bu* besides 'oy-bu "to come" (~ Chin. *wōy* "to go").

<sup>7</sup> Pischel, Grammatik der Prakrit-Sprachen, § 187.

Tibetan and, consequently, every Siamese རྩོ is to be looked at as a *spiritus asper*. Even before consonants (i. e., as a "prefix") རྩོ need not be explained as a vowel or original nasal (*a* < *m*), as done by Professor Conrady.<sup>1</sup> For, spirituses and h-like sounds before consonants are nothing unheard of in the Indo-Chinese languages (cf. the Burmese *ha thō*, the Siamese now mute *hō nā* and 'ō *nā*, etc.); and the facts believed to prove the nasal origin of རྩོ show merely that in pronunciation (not spelling) ante-consonant རྩོ has locally (and in compounds rather universally) turned into a nasal, while in another part of the country it has also arisen from or replaced *m* (which seems to have led to a confusion, also in spelling, of the two prefixes). The latter process is quite naturally explained by the very frequent prefix རྩོ getting the better of the less frequent རྩོ; while, as to the former (which Jäschke accounted for "by neglecting the perfect closing of the mouth channel"), the spirant γ, as which the antevocal རྩོ is spoken in Khams, shows the way, how the *spiritus asper* may have developed into the nasal.

## 2. VOWELS.

To the transcription of vowels, the historical principle is, if ever, anyway not applicable now. Vowels and diphthongs must, therefore, be rendered in accordance with the actual pronunciation, except in the two cases (*ā*, i. e., nasalized *ā*, now spoken *am*, and *āi*, now *ai*; see below) in which two originally distinct letters, still written separately, are now pronounced as one. Such distinctions as the script makes for the same vowel according to its position as initial, medial, or final, as part of an open or of a closed syllable, as the first letter of a monosyllabic or of a polysyllabic word,<sup>2</sup> need not be maintained by transcription. These differences may be historic, but are sufficiently clear from the position of the vowel.

In the Siamese vocabulary the long vowels are more frequent

<sup>1</sup> Eine Indochinesische Causativ-Denominativ-Bildung, p. 23. Conrady bases his hypothesis (*a* < *m*) on the pronunciation of Khams, on the Central-Tibetan pronunciation of the prefix *m* as *a* (as he prefers to write Jäschke's "vowel in itself" which he accepts), and on a few comparisons with words from related languages which, in his own words, are "neither exhaustive nor perfectly safe". He will find it difficult to answer the question why in Khams the two prefixes are kept distinct, *a*, but not *m*, assimilating with a following consonant (*nda* against *ndol-pa*, etc., see Jäschke's tables).

<sup>2</sup> All these cases will be mentioned in their places farther on.

than the short ones, wherefore it has become habitual since Bishop Pallegoix to mark only the short syllables, not the lengths,—a method contrary to general philological practice and therefore not advisable for a transcription that is to serve comparative philology. We, therefore, shall indicate length and abnormal shortness, but leave normal shortness unmarked.

In Sanskrit, nominal stems ending in a vowel (except *r*, etc.) — short or long—may have, for inflection, the so-called *visarga*<sup>1</sup>: an h-like sound, transcribed by *h*, representing original Indo-European *s*. Now, in Siamese, the final *ah* of Sanskrit words such as *māsah* “the month” (nominative) was, and partly still is, pronounced as a short *a* (or as *a* with following *spiritus lenis*): this gave rise to the conception of the *visarga* as a mere symbol of shortness for final vowels (i. e., vowels ending a word or a syllable), and it was, therefore, used to designate such short finals as do not exist (or have no sign of their own) in the Indian script (Sanskrit and Pāli),<sup>2</sup> i. e., for all except *i* (and the Siamese *ü* directly derived therefrom) and *u*. In this manner were obtained: 1) the short *a* just spoken of as final of dissyllabic Sanskrit words<sup>3</sup>; 2) the short *a* as final of monosyllabic words or other than last syllables, e. g., in words taken from Sanskrit such as  $\text{प्रा} \text{ra}$  and  $\text{प्रा-रि} \text{ra-ri}$ , or in the purely Siamese  $\text{pra}$ ,  $\text{ra}$ , etc.; 3) the final short vowels  $\text{[ + Kons. + ] } \text{z}$ ,  $\text{[ + Kons. + ] } \text{z}$ ,  $\text{[ + Kons. + ] } \text{z}$ ,  $\text{[ + Kon. + ] } \text{z}$ , i. e., *e*, *ē*, *o*, *ō*, these being nothing but  $\text{[ + Kons. + ] } \text{z}$ ,  $\text{[ + Kons. + ] } \text{z}$ ,  $\text{[ + Kons. + ] } \text{z}$ ,  $\text{[ + Kons. + ] } \text{z}$  with the sign of shortness added to them;

<sup>1</sup> From *vi* +  $\sqrt{\text{syj}}$  “to throw off, abandon”.

<sup>2</sup> Sanskrit possesses short *a*, *i*, *u*, but after consonants *a* is not expressed (but merely understood) in writing; Pāli has two more short vowels, *e* and *o*, but they are written as long ones.

<sup>3</sup> It is written now in some words with, in others without  $\text{z}$  (cf.  $\text{द्वि} \text{dvi}$  “deer,”  $\text{पुत्र} \text{putra}$  “enemy”, spoken *mīṣa*, *bora*); often appears even as  $\bar{a}$  (thus in the name *Rāmā* = Skt. *Rāmāh*); interchanges with Pāli  $\bar{o}$  (*bālā*, *bālā*, *bālō* “roguish” < Skt. *bālāh* “young, child”); and partly has grown mute (as in  $\text{पुत्र} \text{putra}$  pronounced *phut*, beside the Pāli  $\text{पुत्र} \text{putra}$  *bulldho*;  $\text{प्राण} \text{praṇa}$  pronounced *kān*, “time season”, beside  $\text{प्राण} \text{praṇa}$  *kālu* “tense (in grammar)”; etc.). To transcribe the *visarga*, which is unnecessary in purely Siamese words, *h* might be used, as a mistaking it for Siamese  $\text{h}$  is excluded.

<sup>4</sup> Which without  $\text{z}$  would sound *prā*, or — as a purely Siamese word — *prī*.

and 4) the final short vowel  $\text{[ + Kons. + ] } \text{z}$  =  $\bar{o}$ , which means in reality the shortening not of  $\text{[ + Kons. + ] } \text{z}$  =  $\bar{o}$  but of a diphthong (viz. *ao*, see below), resorted to because  $\text{[ + Kons. + ] } \text{z}$  seemed indispensable for the privative *a* of Sanskrit (see below).

All these short vowels except *o* (see below) occur also as *medials*, i. e., in closed syllables; but only *i*, *u* and *ü*, (like *i*, *ü*) have here the same sign they have as finals. Medial *a*,  $\bar{a}$  are written  $\text{[ + Kons. + ] } \text{z}$  or  $\text{[ + Kons. + ] } \text{z}$  respectively ( $\text{[ + Kons. + ] } \text{z}$   $\bar{a}$  occurring also as final), medial *ē*,  $\bar{e}$  by means of a combination of the *e*-sign ( $\text{[ + Kons. + ] } \text{z}$ ) with that of *i* or  $\bar{i}$  ( $\text{[ + Kons. + ] } \text{z}$ ,  $\text{[ + Kons. + ] } \text{z}$ ).

*O*, in the sense of the short equivalent of *o*, (not  $\bar{o}$ ), though occurring as final, does not exist as a medial in Siamese. The letter *o*, therefore, may be used for transcribing a frequently occurring similar *o*-sound, occurring in closed, and, in Sanskrit words, also in open syllables, which I propose to call the “inherent *o*”: a short vowel midway between broad and narrow *o* (i. e., between  $\rho$  and final *o*)<sup>1</sup>, not written in Siamese (cf. non-initial *a* in Sanskrit and Pāli) but spoken after every consonant that has no vowel-sign attached to it<sup>2</sup>.

Medial *e*,  $\bar{e}$ ,  $\bar{o}$ , are frequent in nasal diphthongs, but have no signs different from those of  $\bar{e}$ ,  $\bar{e}$ ,  $\bar{o}$ . They are expressed through the long vowel together with either the tone or the symbol of (medial) short sounds called *lekhe pēt* ( $\text{[ + Kons. + ] } \text{z}$ ), viz., in the following way<sup>3</sup>:

In words with a nasal as final and with falling or low tone<sup>4</sup> the

<sup>1</sup> And always distinguished with perfect safety from both of these, though it is said to be nearer the second than the first.

<sup>2</sup> Not, however, as final in naming the letters of the alphabet, where no vowel is written but  $\bar{e}$  is spoken. Inherent *o* is spoken for the Sanskrit inherent *a* in older loan-words (as in  $\text{पुत्र} \text{putra}$  besides later  $\text{पुत्र} \text{pu-tri}$ , but not before final *u* < *r* < *ra* where, again,  $\bar{e}$  is spoken (e. g. in  $\text{अक्षर} \text{akṣara}$ , spoken *akṣōn*). The correspondence, in Tibetan, of Siamese inherent *o* is also *o*, e. g., in Tib. *zom* “feeble, decrepit” = Siam. *zom* “weakly, frail” (other instances on p. 50, p. 52 n. 3, p. 53 n. 4). Where *a* appears to, correspond, as in *d-pal*, *r-dab-pa* (above p. 50), etc., we have in reality no direct correspondence (because Siam. medial *a* = Tib. *a*, e. g., in *brāl-ba* ~ *brāl*, etc., see above pp. 50, 53 n. 4 and 5 etc.), but Tibetan or earlier ablaut (in which sense *d-pal* may probably be compared with *phol* “blister, abscess”; *r-dap-ba* with *dob-dob* and *thob-chi*; *dam* with *thom-pa*, etc.).

<sup>3</sup> Compare the paragraph on the tonal system.

<sup>4</sup> Not “with rising or falling tone”, as Wershoven teaches (Lehr- und Lesebuch der Siam. Sprache, p. 15, note). Shortness of the vowel, rising tone, and final nasal are found together only in the few, mostly onomatopoeic words with middle initial and



vowels  $\bar{e}$ ,  $\bar{e}$ ,  $\bar{o}$ , -- written as such by the Siamese -- appear, in pronunciation, more or less short or shortened: unmistakably short is the vowel in the (non-accented) words with *lĕkh pĕt*, and as a rule also in the words with low initial and first accent; in the remaining words the vowel is usually half-long, frequently, however, (at least in individual pronunciation) also short<sup>1</sup>. To draw here a sharp dividing-line will be hardly possible even with the help of phonetic instruments; and, as we cannot tell at present and by Indo-Chinese linguistics alone may never be able to find out which short vowels have been originally such and which are merely the result of shortening,<sup>2</sup> the best method at the time being seems to be to write as short vowels also the half-long ones. Consequently, the long vowels  $\bar{e}$ ,  $\bar{e}$ ,  $\bar{o}$  of Siamese are to be transcribed as short ones in words with a nasal as final and with 1. high, middle, or -- if raised by prefixed *h* (*hĕ nĕ*) -- even low initial and first or second accent; and 2. low initial and first accent (= *hĕ nĕ* and second)<sup>3</sup>. On the other hand, even though the word may end in a nasal, if there is a different accent from those mentioned or no accent at all, the vowel is to be read long, unless, indeed, shortness be directly indicated by *lĕkh pĕt*, which occurs only in non-accented words, including such also as do not end in a nasal<sup>4</sup>.

Thus,  $\text{เล็น } len_1$  "to play" and  $\text{เล็น } len$  "louse"; also  $\text{เล็ญ } khĕŋ_1$  "to view",  $\text{เล็ญ } khĕŋ_2$  "shin", and  $\text{เล็ญ } khĕŋ$  "hard, solid"; and likewise  $\text{ย็ญ } yĕŋ$

fourth accent (*keŋ<sub>1</sub> reŋ<sub>1</sub> tom<sub>1</sub> com<sub>1</sub>*). The use of the *lĕkh pĕt* in non-accentuated words is a proof of their vowel being as a rule long.

<sup>1</sup> E. g., in *sen<sub>2</sub>*, *khem<sub>2</sub>*, *khĕm<sub>2</sub>*, *pĕm<sub>2</sub>*, *khĕŋ<sub>1</sub>*, *hĕmĕŋ<sub>1</sub>*, *yĕŋ<sub>1</sub>*, etc. For restoring the quantity, the nasal of such words is often somewhat prolonged.

<sup>2</sup> A case of original shortness of the vowel seems to me to be  $\text{เล็ญ } sen_2$  "sinew" (= Barm. *say*; spoken *sin*) -- comp. O. H. Germ. *senarwa*, O. Engl. *sinu*, Skt. *sināmi* and Lett. *seni* "to bind" (Osmanli *sinir* "nerve"), and, with complete loss of the vowel, Skt. *smāyu* "sinew", O. II. Germ. *smuor*, etc.; possibly also Greek *τένω* "sinew", Lat. *tenus* "rope", etc.; while we may have original length (*i*?) in the word  $\text{เล็ญ } en$ , meaning also "sinew" -- comp. Greek *ίνε* "sinew", IIung. *én*, *in* do., Tamil *īnai* "nerve of a leaf".

<sup>3</sup> No shortening takes place after *spiritus lenis* (*en<sub>2</sub>*, *ōn<sub>2</sub>*, *ōm<sub>2</sub>*) except where specially marked (*'en* "sinew", = *sen<sub>2</sub>*).

<sup>4</sup> In much used words, such as *hen* "to see", *pen* "to be", the *lĕkh pĕt* is usually omitted.

*yĕŋ*, "to go on tip-toes",  $\text{ย็ญ } yĕŋ$ , "on the stealth"; and  $\text{ห็ญ } hyĕŋ$  "to be afraid" have, indeed, all of them a different tone, but their vowel is the same, or very nearly the same, as to shortness. On the other hand we have, with the corresponding long vowels<sup>2</sup>, e. g.  $\text{เลน } len$  "mud, mire",  $\text{แขวง } khvĕŋ$  "distrikt"  $\text{ยัญ } yĕŋ$  "stag, deer",  $\text{เว้น } vĕn_2$  "except",  $\text{เก็ญ } kĕŋ_3$  "patch",  $\text{ค็ญ } gĕn_2$  "hammer".

That other vowels too, not only  $\bar{e}$ ,  $\bar{e}$ ,  $\bar{o}$ , should have been influenced by a nasal following them, seems probable. As a matter of fact, when e. g.  $\text{ถ้ำ } tĕm$  is pronounced, not the long *a* of  $\text{ถ้ำ } tĕm$  is heard, though, indeed, also not the short *a* of  $\text{ถ้ำ } tĕm$ ; and even though the latter would be heard (which may be the case in individual or dialectic speech), there can here be no doubt (unless we are concerned with pre-Siamese conditions) as to the *original length* of the vowel. For, in writing these other vowels, as we have seen, the distinction of length and shortness is always maintained and, consequently, *lĕkh pĕt* never required; and, considering the shortening tendency of the nasals, it seems perfectly excluded that e. g. an original  $\text{ถ้ำ } tĕm$  should have come to be written  $\text{ถ้ำ } tĕm$ . There is, therefore, no reason for abandoning the Siamese writing of these sounds.

*Initial* vowels in the strict sense of the term are unknown in the Siamese script, since every vowel is preceded, if not by a consonant in our sense, at least by the *spiritus lenis* which for the Siamese is a consonant.<sup>1</sup> As initial of a closed syllable without vowel-sign this spiritus ( $\text{๑}$ ) is understood to contain the same inherent *o* with which in like conditions every other consonant is spoken (therefore, e. g.,  $\text{๑} = 'ok$ ). But if a vowel-sign is added, the vowel expressed by it, not *o*, is spoken (e. g.,  $\text{๑} = 'ak$ ,  $\text{๑} = 'ok$ ,  $\text{๑} = 'ik$ ). If, however, initial *a* is a syllable by itself, it is pronounced *ā*, no matter whether it is written  $\text{๑}$  or  $\text{๑}$ : compare  $\text{๑} = 'alajjī$  "shameless" with  $\text{๑} = 'alobho$  "not greedy". The difference is as intimated by the tran-

<sup>1</sup> The discussion of the mute  $\text{๑}$  of this and the mute  $\text{๑}$  of the next word I have reserved for a special article.

<sup>2</sup> Of accented words very few in addition to those with low initial and second accent come into consideration here.

<sup>3</sup> With  $\bar{e}$ , etc., the shortening is, if at all, audible to a very well practised ear only.

<sup>4</sup> A purely vocalic initial, as e. g. in French *ami*, can therefore not be written in Siamese, if not by its "low" *h* (?).

scription: 'a in the second word breaks off with a spiritus-lenis-like cracking noise and has no influence on the tonal value of the following syllable, while in the first word the ๒ determines as "middle" letter the tone of the next syllable.<sup>1</sup>

As to *under-normal shortness*, the opinion<sup>2</sup> that this is indicated by the *visarga* (ः), is surely wrong. The error may be due to the fact that in many Sanskrit words final *a* or *ah* has grown mute. But *visarga* is by no means confined to Sanskrit words, and there is no reason why the *a* e. g. in កា *ka* should be shorter than the *u* in កូ *ku*. The one under-normal short vowel existing in Siamese is the *sva-rabhaktic vowel* intervening between a consonant, other than a guttural or labial explosive, and *l* or *r*. As long as we do not know for certain that this vowel is purely euphonic<sup>3</sup> in every case of its occurrence, we should not follow the Siamese script in leaving it unexpressed; and, though its somewhat varying pronunciation<sup>4</sup> might perhaps best be rendered by *ɔ* (inverted *e*), we may choose *ᵃ*, as it is thus understood by the Siamese.

The simple vowels, then, in the transcription I propose, are the following:

*a-sounds*: *ā*, *ã*, *a*, *ᵃ*. Of these four, *ā* is long *a* (e. g., in កា *Tā*). *ã* nasalized *ā*, now pronounced *ãm* (as in ណាំ *nam*); *a* the common short *a*, the several Siamese writings of which are due to its position only and need not therefore be reproduced by the transcription; and, fourthly *ᵃ* the *sva-rabhaktic vowel* (e. g., in ឥស្យាម *s'yām* "Siam").

*ī*, *i*; *ū*, *u*. Of these, *i* and *u* are open (broad) in closed syllables, close (narrow) in open ones; *ī*, *ū* always close, but *ū* less so than *ī*.

<sup>1</sup> This distinction is perfectly arbitrary, not justifiable in any way by the Sanskrit originals, otherwise too the rendering of Indian words in Siamese being most capricious. In this connection may be specially mentioned the amusing transcription of Sanskrit *an-eka* "not-one" = "many" by 'a-nēk.

<sup>2</sup> Frankfurter, *Elements of Siamese Grammar*, p. 8.

<sup>3</sup> As would appear from ឥស្យាម *s'yām* "Siam" said to be = Skt. *śyāma* "black". Comp. also Siam. *khayōk* "to limp, lame" ~ Tib. *kyog-po*, *gyog-pa* "crooked, bent, curved"; Siam. *khayuk* "to move by fits and starts" ~ Tib. *khuyug-pa* "to run, move swiftly"; Siam. *kh'net* "weary" ~ Tib. *g-ñid* "sleep". According to Prof. Conrady's view (*Eine indo-chinesische Causativ-Denominativ-Bildung*, p. 53) there would be no *sva-rabhakti* here, but an original vowel preserved in Siamese but dropped in Tibetan.

<sup>4</sup> *Sayām*, *seyām*, *siyām*; but only *sanuk*, and ឥណ្ឌ almost like *syop*.

Examples: តិ *tī* good, ដិត *mit* knife, ដំ *mi* not, កំ *kin* gold; កូ *ku* snake, កូក *luk* child, រូ *ru* to wash, កូត *khut* to dig.

*ē*, *e* are the narrow *e*-sounds in Italian *nero*, *veniamo*, German *see*, *theater*. Examples: ដេ *dē* to pour out, ឆេក *chek* very tired, គេ *re* to kick, តេក *tek* child, រេង *rey*, to hasten.

*ē̄*, *ē* are the broad *e*-sounds, e. g., in French *merc*, *net*. Examples: ឆេ *chē* and, ឆេង *chēng* strong, ឆេង *chēng* to dig out, ឆេង *chēng*, class prefix for flat objects.

*ō*, *o* are the narrow *o*-sounds of French *chose*, Engl. *admonition*, of which *o* occurs only as final, i. e., in open syllables. Wherever we use the sign *o* as medial, it means the "inherent *o*" which is always medial except in the particle ក៏ "also" (with high tone) which we transcribe by *kō*. Examples: ខេ *xō* beggar, យេក *yok* wavering, ខេង *gōng*, arched, ខេង *lo* to throw away; កេង *gōng* solid, តេង *sot* fresh.

*ō̄*, *ō* are broad as in French *port*, *portable*. In words consisting of 'ō + consonant, this *ō* is graphically identical with the initial 'a of dissyllabic Indian words: compare ឧក *ōk* "breast" with ឧណ្ណ *'anu* "atom". Initial or solitary 'ō is ឧ, i. e., spiritus lenis + *ō*; post-consonant *ō* is = ឧ. The latter is also pronounced but not written when a consonant is called by name (e. g., ឧណ្ណ = *hō nam*). Examples: ឧក *yō* to flatter, ឧក *ōt* lungs, កេង *kō* island, ឧណ្ណ *dōu*, piece.

*ī̄*, *ū̄* is a pair of sounds peculiar to Siamese: not wholly identical with the French or German *ii*-sounds nor with Annamite "u barbu".<sup>1</sup> The long sound is written *ī̄* in closed, *ī̄* + ឧ in open syllables<sup>2</sup>; the short sound always by *ī̄*. As transcription Wershoven chooses *y*, which in our system (as in Indology) stands for the palatal semi-vowel (ឃ). Examples: ឃ *thū̄* to ask, ឃ *mūt* dark, night, ឃ *hū̄* fie! ឃ *pū̄ng* swamp, marsh.

*ō̄*, *ō̄* is the broad *ō̄* in French *saur*, *œuf*, written *ō̄* [+ cons. + ] ឧ and *ō̄* [+ cons. + ] ឧ in open syllables, *ō̄* and *ō̄* in closed syllables. This pair appears in the Siamese alphabet among the diphthongs, pro-

<sup>1</sup> The latter is broad and similar in sound to *ō̄*, *ō̄*; but Siam. *ī̄* is something between the broad *ō̄* (see below) and the German narrow *ii*, while Siam. *ū̄* in closed syllables sounds almost as in German *hütte* (in open syllables it is hardly found).

<sup>2</sup> Which may or may not point to a slight (original) difference of pronunciation.

bably in accordance with its origin.<sup>1</sup> Examples: เกล็ด *klē*, friend, เฒ่า *pōy*, to stare, look at, เเขย *gō* shy, coy, เงิน *yōn* silver, money.

The Sanskritic *r*, *ṛ*, *l*, *ḷ* ought to be confined to Sanskrit words, but they are mixed up by the Siamese with their *r + ü*, *r + h*, *l + ü*, *l + h*.

The transcription of *diphthongs*<sup>2</sup> is a difficult topic: most of them contain a semivowel which ought, perhaps, from the historical viewpoint, to be rendered as such; this, however, seems for the present impracticable.<sup>3</sup>

The twenty-five diphthongs<sup>4</sup> may be described as follows:

1. *Diphthongs without a semivowel*: *āi*, *āi* (now both pronounced *āi*)<sup>5</sup>; *ao*; *ha*<sup>6</sup>. Examples: ไ้ *rāi* to go, ใบ *pai* leaf; เกา *kao* to scratch; เกล็ด *klua* salt, เมือง *mūay* town.
2. *y-diphthongs*<sup>7</sup>: *āi*, *ai*; *ōi*; *ōi*; *ōi*<sup>8</sup>; *ui*; *üi*.—*ue*<sup>9</sup>.—*ia*, *ie*<sup>10</sup>.—*hē*<sup>11</sup>.

<sup>1</sup> Comp. Siam. *kāt* "to be born" with Tib. *skye-ba* do., *s-kyed-pa* "to generate", Siam. *pōt* "to open" with Tib. *byed-pa* do. Even the short vowels *e*, *ɛ*, *o*, *ɔ* seem to be looked at as diphthongs, but this is apparently due to their graphic symbols only.

<sup>2</sup> On nasal diphthongs see above p. 59 seq. Other combinations of vowel + sonant besides these and those to be dealt with do not exist in Siamese, final *l* and *r* having become *n*, at least in pronunciation.

<sup>3</sup> It will probably be found that, as in Chinese, many a diphthong has arisen from vowel + final consonant. Compare also Tibetan, where *e. g. nas, tes*, are now spoken in Northern Ladak as *nai, lei*, while, on the other hand, in the pronunciation of Central Tibet *chos, las* have—undoubtedly through the intermediate stage *choi, lai*—turned into *chō, lē* (comp. Jäschke, Tibetan Grammar, pp. 6 and 13, also our last but one note).

<sup>4</sup> In Wershoven's list (see his grammar, p. 28) as well as in Cartwright's (Siamese-English Dictionary, Introduction, pp. 22–23) four more diphthongs are mentioned (*hā, hā, hā, hā*) which, however, exist only in theory or as exclamations (as which also occurs *hā* added by Pallegoix), while, on the other hand, those two real diphthongs are ignored which in writing have a shortened long vowel as first part.—The Siamese do not look upon *āi, ai, ao, wa* as diphthongs, while they reckon as such six other sounds which now, at any rate, are none (see above p. 63 seq. and 58 seq.: *ō, e, ɛ, o, ā, ɔ*).

<sup>5</sup> The connecting hyphen is needed to distinguish these sounds from the *y*-diphthongs *āi, ai* which the Siamese too consider as diphthongs.

<sup>6</sup> I cannot endorse Wershoven's observation that *ha* is confined to open syllables, while in closed ones *hā* is spoken.

<sup>7</sup> I thus call all diphthongs written with *ไ้* in the end.

<sup>8</sup> Written with, or, more often, without *ไ้*.

<sup>9</sup> Understood as *hē* by Wershoven. But this *h* is not different from that of the *y*-diphthongs *ua, wa*; it is a short though narrow *u*.

<sup>10</sup> The former in open, the latter in closed syllables, with but one graphic symbol.

<sup>11</sup> Wershoven's rendering *hē* is inexact, and so is his *hō* for *hō* (see *y*-diphthongs). In both diphthongs the second vowel too is audibly long.

Examples: ตาย *rāi* to die, นัย *nai* kind, sort; โภย *kōi* to pick up; ฝอย *fōi* down-feather, ฝอย *phōi* to drop off, to sleep; เนย (เนอຍ) *nōi* butter; ทุย *khui* dusty; ฮัย 'hī gee up! (excl. to oxen or buffaloes).—นวย *nue* fickle, inconstant.—เมีย *mīa* wife, เรียน *rīen* to learn.—เหนื่อย *nūē* to be tired.

3. *v-diphthongs*: *āo*; *eo*, *eo*; *āo*; *iū*, *iū*; *ua*.—*uo*.—*iō*.<sup>1</sup> Examples: ยาว *yāo* long; เอว 'eo waist, เรว *reō* quick; แนว *nāo* row, line; ฮว 'hu, hurrah! นัว *chiu* a little, gently; หัว *hua* head.—สวน *suon* garden.—เวยว *cīo* to roast.

Whether there existed *triphthongs* in Siamese (may *hē* and *iō* have been such?) remains to be seen. In such words as ไทย "Siamese" (lit. "free"), วิเวย "monk's code", etc., the mute *y* must be written: *dāiy, vināiy*, etc.<sup>2</sup>

### 3. TONAL SYSTEM

In the Siamese script the tones are expressed *indirectly* as follows:

1. in words ending in a long vowel, diphthong, or nasal by *four accents* called "first" to "fourth" each of which, together with the *original* initial ("high", "middle", or "low"), is capable of expressing some or one resp. of the *five tones*; and 2. in words ending in a short vowel or a consonant (other than a nasal) by the mere initial which produces, if "high" or "middle", the low tone, or, if "low", either with a long vowel—the low tone, or with a short vowel—a *sixth tone* described by Wershoven (loc. cit. p. 25) as an uniform high tone.

Now, the original initials being *throughout* preserved only in writing, but *only partly* in modern speech, it is evident that a transcription based on the latter cannot express the tones in the way described, but is forced to express them *directly*, i. e., by accents each of which always denotes one and the same tone. Thus the *acutus* was chosen for the rising tone, the *gravis* for the falling tone,

<sup>1</sup> Written *เวยว*. Comp. *เวย* < *hē*.

<sup>2</sup> There seems to be a real triphthong in the word ไ้ *suāi* "buffalo" which would perhaps be more correctly transcribed as *guāi*.

the circumflex (˘) for the low tone, and the dot (point) for what was called *tonus gravis* by Pallegoix (Cartwright's "acute tone", Wershoven's "eingehender Ton"), the normal tone remaining unmarked, and the sixth tone being treated either as normal or as *t. gravis*.

This second method of accentuation is less practical than the first in that it increases not inconsiderably the number of marked words.<sup>1</sup> Moreover, it is less exact than the first, even if a sign for the sixth tone were added. For, these five or six signs characterize the Siamese tonality just in the main points only, but are far from exhausting it.<sup>2</sup> The watchful observer is said to be able to distinguish and read from the Siamese script no less than fourteen different inflexions. This is hardly to be wondered at, considering the multiplicity of factors (initial, final, quantity, etc.) co-operating in the Siamese tonal system. It seems, therefore, the most advisable, for a scientific transcription, to keep as close as possible to Siamese writing also with regard to the tonal system, *i. e.*, to render the accents written by the Siamese themselves. This may be done by using for them the numerals 1 to 4 (˘ = 1, ˙ = 2, ˘˘ = 3, ˙˙ = 4) to be placed at the end of the words, *viz.*, below them, so as to avoid their being mistaken, in comparative linguistic research, for the Chinese accents written also after the words, but at the upper end, in the form of numerals.

To write the accents and not the tones is advisable also because in that case it is beyond doubt that the two enigmatical prefixed letters known as "leading *h*" and "leading *'*" (*h̄p n̄ā* and *'p̄ n̄ā*), even though these be mere tonal signs (as which they are now understood), must be actually rendered in transcription, and not suppressed, as is generally done.

<sup>1</sup> *E. g.*, in Frankfurter's first Reading exercise (*loc. cit.* p. 23) from seven to fifteen, in the second from sixteen to twenty-three.

<sup>2</sup> Michell (*Siamese-English Dictionary*, Preface, p. XV) speaks appropriately of "the Siamese language with its infinitely varied modulations".