THAI, KADAI, AND INDONESIAN: A NEW ALIGNMENT IN SOUTHEASTERN ASIA

By PAUL K. BENEDICT

IN THE present paper the writer presents a general solution to the complex problem of the affinities of the Indonesian languages. The two following premises are basic to the thesis developed here:

1. The true Indonesian substratum on the Asiatic mainland is represented by four scattered languages in southern China, northern Tonkin, and Hainan, all of which constitute a single linguistic stock (Kadai).

2. The recognition of the Kadai stock, which shows numerous points of contact with Thai, opens the way to a new interpretation of the latter as a more distant member of an archaic Thai-Kadai-Indonesian linguistic complex.

Although these suggestions are new and perhaps unexpected, it can be said that they accord with the general picture as reconstructed from historical and cultural data. It is generally agreed that the Indonesian migrations have proceeded from the Asiatic mainland, but the evidence brought forward has been of a generic rather than specific nature, and the area of departure has not been delimited. The linguistic speculation has been notable for range rather than relevancy, and the cultural treatment has in some instances been equally unsound. It is hoped that the argument developed below will provide a number of solid points d'appui from which further ramifications can be anticipated.

The newly recognized Kadai stock comprises the Li dialects of the island of Hainan, the Kelao language of southcentral China, and the Laqua and Lati languages of the China-Tonkin border region. The term "Kadai" has been compounded by the writer from "Dai," one of the forms of the Li term for themselves,* and the kd- prefix found in Laqua kdddü, Kelao kståü "man (homo)." These languages, with the exception of Li, are not generally known to the scientific world, and our available sources are rather meager. The Li dialects have been described by a number of European observers, the most thorough of whom have been Savina and Stübel-Meriggi. Bonifacy has pub-

---

* For some of the earlier speculation, see W. Churchill, The Polynesian Wanderings (Washington, 1911), largely devoted to a criticism of MacDonald’s Semitic theory.

* An outstanding example is furnished by Handy’s derivation of the Polynesian Tangaloa cult from southern China (Polynesian Religion, Bulletin of the Bernice P. Bishop Museum, no. 34, 1927, pp. 312–330).

* Other variants are: B’lai, K’ai, S’ai, Hiai, Lai, Loi, Le, Dli, B’li. The Chinese character employed for this name is pronounced li in North China dialects, lai in Cantonese, and loi in Hoklo.


576
lished word-lists of Kelao, Laqua, and Lati, while additional material on all three languages has been supplied by his compatriot, Lunet de Lajonquière.

A third Kelao source has been furnished by Samuel R. Clarke, the author of a popular account of the little-known tribal groups in southern China.

The Kedai languages have received scant attention from anthropologists and linguists. Li has evoked occasional comment, yet no real analysis has been attempted. The obvious Thai element in the language was noted by Parker over half a century ago, and this theme was further developed by Strzoda.

The less apparent Indonesian affinities were first pointed out by Terrien de Lacouperie, who suggested a relationship with the Indonesian languages of Formosa. In more recent times P. Mus and H. Maspero have further extended this line of thought and have supplied the first concrete bits of evidence. Maspero, a sound and generally conservative scholar, concludes that the Li numerals "certainly" belong to the Indonesian family.

The three mainland languages (Kelao, Laqua, Lati) have attracted still less attention. Bonifacy, who recorded Laqua, noted the analogy between the Laqua and Cham numerals, but this observation seems to have been over-

---


9 W. Strzoda, Die Li auf Hainan und ihre Beziehungen zum asiatischen Kontinent (Zeitschrift für Ethnologie, Bd. 43, 1911), pp. 193-236. Strzoda concludes, however, that "Die meisten Li-Numeralia ... sind Rätsel und lassen sich nirgends unterbringen" (pp. 219-220).

10 The Languages of China before the Chinese (London, 1887). See especially his conclusions on p. 73: "In the numerals, for instance ... similarities exist with those of some tribes of Formosa. But they are remote, and do not come from a direct relationship; they are apparently survivals of a former state of things, previous to their respective migrations, when their various ancestors had relations between themselves on the continent." An English traveller in Hainan, B. C. Henry, had somewhat earlier sought to connect the Li with the Malay on onomastic grounds (B'bai or B'lay = Malay), in his article, The Close of a Journey through Hainan (China Review, Vol. 12, 1883), pp. 109-124, esp. p. 115.

11 Review of Savina, Monographie de Hainan (1929), in Bulletin de l'École Française d'Extrême-Orient, t. 30, 1930 (pp. 436-444). Of his own Cham and Malay comparisons, however, Mus remarks: "Ces rapprochements sporadiques restent jusqu'ici de simple curiosité."


13 Cit. supra, 1908, p. 557. Bonifacy adds the following remark: "Là paraissent s'arrêter les ressemblances entre les deux langues [Laqua and Cham], à supposer même que celles que nous signalons ne soient pas purement fortuites."
looked by Maspero and other scholars. Kelao and Lati have gone almost entirely unnoticed, although W. Schmidt has seen fit to classify the latter as an independent linguistic entity. It was Bonifacy’s observation on Laqua that led to the writer’s discovery of the relationship between Laqua and Li, and thus ultimately to the concept of a single unified Kadai stock.

The Kadai-speaking groups are all of marginal type, as should be expected on the basis of our substratum theory. The Li, who inhabit the mountainous central and south-central parts of Hainan, are under economic pressure from their powerful Ong-Be (Thai-speaking) and Hoklo (Chinese-speaking) neighbors. The Kadai groups on the mainland rank even below the Miao and Lolo, and generally regard themselves as autochthonous. The Laqua, who call themselves Ka Beo, in the upper Rivière Claire valley of northern Tonkin, are described by Lunet de Lajonquière as follows:

*ils se considèrent comme aborigènes et il est certain qu'ils sont venus dans la contrée avant toutes les autres tribus montagnardes* (cit. supra, p. 339).

*C'est une variété* [of economic life] *en complète décadence. La plus grande partie des terres qu'ils cultivaient ont été déjà cédées aux Meo, qui paraissent devoir les absorber* (ibid., p. 341).

Of the Lati, also in the upper Rivière Claire valley, the same writer states simply that "*ils prétendent aborigènes*" (cit. supra, p. 358). Bonifacy places his estimate of the number of Lati at only 450 (76 families).

The Kelao or Lao, who call themselves Thii, range over an extensive area in south-central China and northern Tonkin, but their true home appears to be Kueichou province, whence they have migrated into the northern Tonkin border region (cf. Lunet de Lajonquière, *cit. supra*, p. 356). Clarke, who has given us the fullest available account of the Kelao, stresses the aboriginal nature of the group:

The Keh-lao, however, are now nearly extinct; many of them have married into Chung-chia [Thai] and Old Chinese families. Some writers have spoken of them as extinct. As far as we know, there are now only several hamlets of them in the Anshun prefecture [west-central Kueichou], which altogether do not number more than two or three hundred families. These people claim, and rightly, we believe, to be the real aborigines of that region. In some parts of the province the Miao claim to be the aborigines, but where the Miao and Keh-lao occupy the same district, the Miao allow that the Keh-lao were there before themselves (*cit. supra*, p. 13).

Another missionary writer, Aloys Schotter, also attributes a low rank to the Kelao:

*Le plus bas dans l'échelle sociale c'est peut-être le groupe des Blancs* [White Miao]. *La tribu des Kê-lao est peut-être plus dégradée encore surtout quant aux moeurs*.15

The languages spoken by these primitive groups fall into two major divisions, viz. Li-Laqua and Lati-Kelao, which together constitute the Kadai stock. Dialectal differences can be established both for Li and Kelao, and are of such magnitude that they must fully be taken into account. The numer-

---

15 *Notes ethnographiques sur les tribus du Kuey-tcheou* (Anthropos, Bd. 6, 1911), pp. 318–344; citation from p. 318.
ous Li dialects can be classified under the headings of “Southern Li” and “Northern Li” on the basis of their treatment of original nasal initials. In Northern Li these initials tend to be transformed into the homorganic stops, whereas in Southern Li they are uniformly retained; cf. N. Li ba~pa, S. Li ma “dog” (Thai *hma); N. Li dau~tau, S. Li nau “long” (Thai *nau); N. Li ka, S. Li nga “horse” (Annamite ngūa). The “Central Dai” dialect recorded by Savina and most of the dialects recorded by Jeremiassen and Stübel belong in the Northern Li group, while the “Southern Dai” dialect of Savina, the Yulinkau dialect of Swinhoe and Calder, and the K’iung-Shan dialect of Parker belong in the Southern Li group. Kelao similarly shows a dialectical cleavage between “Northern Kelao” (dialect recorded by Clarke) and “Southern Kelao” (dialect recorded by Bonifacy and Lunet de Lajonquière). The distinctions here, both lexical and phonetic, are even more marked than those that obtain in Li, but conform to no easily recognizable pattern. It is apparent that a full treatment of the linguistic problems of Kadai would require detailed phonetic information on a wide range of dialects for at least four languages, and it is not unlikely that further exploration in the Tonkin-China border area will reveal still other members of this stock. Unfortunately, we lack the materials necessary to implement a complete study of the whole stock, and must content ourselves with a survey of the more salient points.

All four Kadai languages are of monosyllabic, isolating type, with full tonal systems as in Thai. The Kadai word-order, like that of Thai and Indonesian, shows object following verb, and modifying elements (including genitive constructs) following modified elements; thus, Malay mata hari, Li sa ven, Thai *la wan “sun,” lit. “eye (of the) day.” Kadai, like Thai, lacks the affixation apparatus of Indonesian, yet prefixed forms abound in the Lati-Kelao branch of the stock, e.g. Lati prefixed m-in m-tsu “moon,” m-bo “sky,” m-ti “earth,” m-ni “ox,” m-go “cat,” m-so “elephant,” m-si “beak, mouth,” m-tsu “eye,” m-ngū “oil”; Lati prefixed a- in a-na “rain,” a-ia “rat,” a-k’o “monkey,” a-li “tiger,” a-kū “bird,” a-li “fish,” a-k’e “frog,” a-k’u “man,” a-sa “hair,” a-nu “salt.” Kelao has prefixed bu- occasionally corresponding to Lati prefixed m-, as in Kelao bu-to “earth,” bu-tsū “beak” (bu-tsū-lūa “mouth”). Laqua has prefixed kā- in kā-dū “man,” kā-zio kā-pū “boy,” kā-zio ka-mai “girl,” where zio stands for “child” and pū and mài are the sex modifiers.

On the phonetic side, the Kadai languages present a fairly uniform pattern of relatively simple type, though mixed (indeterminate) and front-rounded vowels are uncommonly abundant. Li exhibits the greatest range of initials.

18 None of our records of Kadai languages is satisfactory as regards transcription. Savina employs the cumbersome and inadequate qu’oc-ngū system of Annamite, while Bonifacy makes use of a modified version of the same system. In the present study open vowels are written as short vowels (ə, ɔ), and the “bearded o” (ơ) and “bearded u” (ư) as front rounded vowels (ø, u).
and finals, with Laqua not far behind, while Lati and Kelao have undergone a process of extreme modification, in the course of which almost all final nasals and stops have been eliminated. The phonetic attrition shown by Lati and Kelao has proved to be one of the chief stumbling blocks in our analysis of Kadai phonology. When it is realized that Li and Laqua, the better preserved pair of languages, stand in much the same relationship to Indonesian, some inkling of our difficulties can be gained. The investigation of the phonetic shifts exhibited by the Thai roots in Li, the best recorded of the Kadai languages, has brought to light a number of significant variations, especially as regards initials, which are useful in the study of Thai itself. Thus, the writer has reconstructed a separate phoneme *hr (surd r) for archaic Thai on the basis of the equation Ahom r= Siamese h= Tho t’-; in this series Li significantly has s-, suggesting an original *sr-:

<table>
<thead>
<tr>
<th>Stone</th>
<th>Siamese</th>
<th>Tho</th>
<th>Li</th>
</tr>
</thead>
<tbody>
<tr>
<td>stone</td>
<td>rin</td>
<td>t’in</td>
<td>sien</td>
</tr>
<tr>
<td>loose</td>
<td>rau</td>
<td>t’au</td>
<td>s'au</td>
</tr>
<tr>
<td>break</td>
<td>rak</td>
<td>h’ak</td>
<td>s’ak</td>
</tr>
<tr>
<td>carry</td>
<td>rap</td>
<td>h’ap</td>
<td>t’ap</td>
</tr>
</tbody>
</table>

The variations in initials between S. Li and N. Li are often of unusual type, e.g. S. Li d-= N. Li f-, corresponding to Thai d- (Siamese d-= Shan l- = Khamti n-): 18

<table>
<thead>
<tr>
<th>Earth</th>
<th>Shan</th>
<th>Khamti</th>
<th>S. Li</th>
<th>N. Li</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siamese</td>
<td></td>
<td></td>
<td>Li</td>
<td>N. Li</td>
</tr>
<tr>
<td>fan</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d’n</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d’u</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Aspiration of initial stops is characteristic of Li; cf. Li h’dn, Thai *gu’dn “smoke”; Li ha, Thai *ga “thatching grass”; Li hang, Thai *gang “jaw”; Li h’dn, Thai *k’dn “crow of a cock”; Li h’dn, Thai *k’du “horn,” also “mountain”; and Li k’du, Thai *k’su “old”; Li k’di, Thai *k’oi “fowl.” Li often simplifies the complicated diphthongs and triphthongs of Thai, but note Li medial -ie- = Thai -i-, i.e., as in Li dien~lleen, Thai *ltn “tongue”; Li diep~siep, Thai *dition “raw”; Li liep, Thai *lèp “fingernail,” and Li -du= Thai -au, as in Li bòu, Thai *bau “leaf”; Li t’sòu, Thai *t’sau “heart”; Li t’òu, Thai *t’oi “low.” S. Li retains final -k after short vowels but substitutes a glottal stop 19 after long vowels, while

17 Lati has retained the final stop only in a-liep “claw” (Thai *lèp “fingernail”) and the Annamite loan-word bat “pen.”

18 Thai d- and b- are best reconstructed as lenis surd stop initials, since they belong in the high tonal series along with the regular surd stops (l- and t’, p- and p’-). Li agrees with the majority of Thai languages proper in having b- for Thai b-, as in Li bòu, Thai *bau “leaf.”

19 Represented in Savina’s transcription by the Annamite tone n’ang.
N. Li uniformly retains the final velar stop; cf. Li p'dk, Thai *vāk “hatch”; Li t'ōk, Thai *tōk “fall”; Li fī', Thai *pīk “wing”; Li t'ov, Thai *t'ok “pour”; Li sa', Thai *sak “pestle”; S. Li dru', N. Li jūk, Thai *dūk “bone.”

The morphological and phonological points developed above point to Thai rather than to Indonesian, yet the lexical elements of Kadai bear the unmistakable imprint of the latter stock, along with an equally deep imprint of the former. In brief, the numerals and a scattering of nouns, pronouns, and adjectives show Indonesian affinities, while many of the remaining elements show Thai affinities. On the basis of this distribution, the writer at first regarded Kadai as a composite of Indonesian and Thai, with the former as the more likely substratum. Further analysis of Thai, however, has led to the view presented below; to wit, that Thai, Kadai, and Indonesian together constitute a single linguistic complex. Kadai is the “transitional” member of this triune, though in the main it approaches Thai rather than Indonesian. Both Thai and Kadai have reduced a number of disyllabic roots to monosyllables, have developed complete tonal systems, and have discarded the original morphological apparatus of affixes. Throughout this elaborate linguistic metamorphosis, however, a number of basic lexical landmarks have persisted and it is to these that we shall direct our attention.

The Kadai numerals are of fundamental importance in the present connection, since the Indonesian affinities of the stock are more apparent there than elsewhere. The following table of Kadai numerals, in which reconstructed Indonesian (IN) roots taken from O. Dempwolff’s recent work have been incorporated, serves to illustrate this point.

<table>
<thead>
<tr>
<th></th>
<th>IN</th>
<th>Laqua</th>
<th>S. Li</th>
<th>N. Li</th>
<th>S. Kelao</th>
<th>N. Kelao</th>
<th>Lati</th>
</tr>
</thead>
<tbody>
<tr>
<td>one</td>
<td>*'it'a'</td>
<td>tīā</td>
<td>kū</td>
<td>ū</td>
<td>tsi</td>
<td>si</td>
<td>tām</td>
</tr>
<tr>
<td>two</td>
<td>*'duwa'</td>
<td>bē</td>
<td>dāu</td>
<td>trau</td>
<td>dū</td>
<td>so</td>
<td>fū</td>
</tr>
<tr>
<td>three</td>
<td>*kāhu'</td>
<td>tō</td>
<td>su</td>
<td>su</td>
<td>tō</td>
<td>da</td>
<td>si</td>
</tr>
<tr>
<td>four</td>
<td>*'o(m)pāt</td>
<td>pe</td>
<td>sau</td>
<td>so</td>
<td>pu</td>
<td>bu</td>
<td>pu</td>
</tr>
<tr>
<td>five</td>
<td>*lima'</td>
<td>mō</td>
<td>ma</td>
<td>pa</td>
<td>mlēn</td>
<td>mbu</td>
<td>ng</td>
</tr>
<tr>
<td>six</td>
<td>*'ōnam</td>
<td>nam</td>
<td>nom</td>
<td>tom</td>
<td>tō</td>
<td>nang</td>
<td>nā</td>
</tr>
<tr>
<td>seven</td>
<td>*pītu'</td>
<td>mō tāu</td>
<td>t'ū</td>
<td>t'au</td>
<td>dī</td>
<td>tī</td>
<td>tī</td>
</tr>
<tr>
<td>eight</td>
<td>*wālu'</td>
<td>mō dū</td>
<td>du</td>
<td>au</td>
<td>sīa</td>
<td>vleu</td>
<td>be</td>
</tr>
<tr>
<td>nine</td>
<td>*tīwa'</td>
<td>mō dīa</td>
<td>pōu</td>
<td>fōu</td>
<td>kū</td>
<td>su</td>
<td>lu</td>
</tr>
<tr>
<td>ten</td>
<td>*puluh</td>
<td>pāt</td>
<td>p'ōu</td>
<td>fōu</td>
<td>tsūk</td>
<td>beu</td>
<td>pa</td>
</tr>
</tbody>
</table>

The following variants are worthy of comment: S. Kelao mlēn “5” but

---

20 For the general argument here, see the Thai discussion below.
21 Vergleichende Lautlehre des austronesischen Wortschatzes; Bd. III; Austronesisches Wörterverzeichnis (Beihafte zur Zeitschrift für Eingeborenen Sprachen, Bd. 19, 1930). Forms as cited by Dempwolf, with the exception of y for j, w for v, and r for γ. “Facultative” nasal infixes are enclosed in parentheses.
Some of the leading features of Kadai phonology are illustrated in the above table of numerals. N. Li pa<ma "5," tom<nom "6," and fōi<pōi "9," fuoi<p'ouoi "10" are all regular developments (see the discussion above). Li kūi~ī "1" are probably independent of the IN root, and the analysis of Li su "3," sau~so "4" is not certain. For the latter, Maspero suggests a development comparable with that found in Tarema (Formosa), which has suatto <*sual<*swat<*sbat<*s-pat "4." Li du~au "8" belong to a puzzling series in which S. Li initial d- corresponds to N. Li initial h- or vocalic anlaut, e.g. dai~hai "iron," dai~hiai "a Li," duoitt~ui "fat" (n.), dūi~ūi "thin." These forms seem to have been derived from roots with labial+liquid initial cluster; cf. the variant form b'hai "a Li," and the frequent correspondences with Thai initial I-, as in S. Li da, Thai *rīa "boat"; S. Li dāt, Thai *rīt "squeeze"; S. Li dūi, Thai *rīan "house" (N. Li plong). We can infer a bifurcate development of the type *walu'>*wlu'>pluoi or *blu'>*duoi (N. Li). Li p'uoit~fuoi "10" attest to a pair of shifts, viz. final -h>-t, as in IN "darah, Li dat~laf "blood," and medial -u->uo-, as in Thai *nung, Li ūng "mosquito." The development here has been of the type *puluh>*p'luoi>p'out.*

Laqua parallels Li in the developments mō du>*walu"8" and pāt>*puluh "10." Laqua tīd<*ti'a"1," lōu>*tolu"3," and mō bei>*biwai "9" reveal IN affinities not apparent in Li. The Laqua vocalic shift a<e~ā is found in the forms be "2," pe "4," and mō "5"; cf. Laqua pō, IN *batu' "stone"; Laqua pe, IN *bapa' "father"; Laqua(te), IN *matu' "eye"; Laqua ne Li na~ia, Thai *na "rice-field"; and, medially, Laqua dōn, Li dān "100"; Laqua nen, Li ēn "moon." The appearance of mō "5" in the Laqua numerals "6" to "8" is suggestive of a quinary system; cf. the S. Kelao numerals cited by Lunet de Lajonquière: sū-u "2," lo-u "3," pu-u "4," nlē-u "5," but tāi-ni "6," dž-ni "7," suo-ni "8," ku-ni "9." Laqua and Li have a common root for "100" (dōn ~dan), which is independent of the IN root (*ratut').

The Kelao and Lati numerals are, in general, further removed from the IN system as reflected in Laqua and Li. Notable, however, are S. Kelao mīlēn (Bonifacy)~nīlē (Lajonquière)<IN *lima' "5," and N. Kelao velu<IN *walu' "8," which show retention of the liquid phoneme l. Kelao pu~bu, Lati pu "4" reflect an a>u vocalic shift, which is especially characteristic of Lati; in the table below, the contrast with the Laqua e~ō vocalism is made clear:

---

22 The IN system is decimal rather than quinary, yet the Formosa languages show irregular features suggestive of the latter, e.g., Sek-hwan has 5+1=6, 5+2=7, etc., and Tsul-hwan and Bu-hwan have 3×2=6, 4×2=8; vide T. L. Bullock, *Formosa Dialects and their connection with the Malay* (China Review, Vol. 3, 1875, pp. 38-46).
The variability reflected in the Kadai numerals appears also in other aspects of the vocabulary. Scarceley any roots prevail everywhere, and there are a number of confusing "partial equations," yet many significant features emerge. One of the most notable of these features is the regularity shown in the roots for "dog," "pig," and "horse," the first two with Thai affinities, the last with Annamite:

<table>
<thead>
<tr>
<th></th>
<th>Laqua</th>
<th>S. Li</th>
<th>N. Li</th>
<th>S. Kelao</th>
<th>N. Kelao</th>
<th>Lati</th>
</tr>
</thead>
<tbody>
<tr>
<td>dog</td>
<td>*hma</td>
<td>mā</td>
<td>ma</td>
<td>pa</td>
<td>xmā</td>
<td>mu</td>
</tr>
<tr>
<td></td>
<td>(Thai)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pig</td>
<td>*hmu</td>
<td>mu</td>
<td>mau</td>
<td>pau</td>
<td>xmūd</td>
<td>ma</td>
</tr>
<tr>
<td></td>
<td>(Thai)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>me</td>
</tr>
<tr>
<td>horse</td>
<td>*ngūa</td>
<td>rre</td>
<td>nga</td>
<td>ka</td>
<td>ngūd</td>
<td>niau</td>
</tr>
<tr>
<td></td>
<td>(Ann.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ngo</td>
</tr>
</tbody>
</table>

Note the N. Kelao and Lati shift u<a in mu<*hma "dog," and N. Li p-<m-, k-<ng- (vide supra). These loan-words, if such they be, must be of some antiquity, in view of the selective nature of the distribution (there is no trace of the prominent Thai-Chinese root *ma "horse"), as well as the noteworthy equation S. Kelao xm-=Thai hm-, the latter a reconstructed phoneme (surd m) not found in any of the modern Thai languages. In the same general class belongs the correspondence between Laqua k’di, Li and N. Lelao k’ai, Lati ka “fowl,” and the Thai root *kdi; contrast the earlier stratum reflected in the series IN *manuk “fowl, bird," Laqua nuk, S. Kelao nie “bird,” and Thai *nōk “bird.”

The following group of comparisons, arranged roughly according to natural lexical divisions, is intended to serve as an index of the Kadai-Indonesian relationship:

1. Laqua vuon (Lajonquière mo ven) “sun,” Li ven “day,” sa ven “sun” ("eye of the day"), S. Kelao du vuđ “sun,” IN *wari “day, sun” (IN medial -r-<Laqua and Li -n). Cf. also N. Kelao veļ “sky,” which shows a contrasting type of development (*wari’<*wli<veļ).

2. S. Li (pa) pūn “rain,” IN *s(m)bun “atmospheric precipitate” (Tagalog ’ambon “fine rain”).

3. Li nom~nam, IN *danum “water.”

4. Laqua pāi, Li pei~fei, S. Kelao p’i, N. Kelao bai, Lati pie, IN *’apuy “fire.” For the Li development (*’apuy>*api>*pei), cf. Li ngei, IN *tangit’ “weep”; Li nei, IN *’ini “this.”
5. Laqua *pung, IN *bunga’ “flower.”
6. Laqua kā-dāa, Li ċu, IN *tawu’ “man (homo).”
7. Laqua pe, Li fa-ba, S. Kelao ē-ba, Lati pu, IN *bapa’ “father.” For the vocalism, see the analysis above.
8. Laqua *ru (Lajonquière), S. Li dau, N. Li fo~o, IN *ulu’ “head.” The Li development has been *ulu’>*ulu>du~o, exactly paralleling IN *walu’>Li du~au “8.”
9. Laqua ḍam, S. Kelao lō so, N. Kelao ma sang, Lati a-sa, IN *d’a(m)but “hair.” The original palatal initial has everywhere been assibilized: *d’a(m)but > *d’am > ḍam and sang~sa~so. For the Laqua initial ṭ-, cf. Laqua ċa, IN *hud’an “rain.”
10. Laqua te, Li sa, N. Kelao dau, Lati m-tšu, IN *mata’ “eye.” Li appears to have developed a sibilant initial through aspiration (*mata’>*m-t’a>sa); see the discussion below of the Thai root ṭa. S. Kelao perhaps retains the root in the compound du vud “sun,” paralleling Li sa ven, Malay mata hari (“eye of the day”), yet this dialect also has du die “moon,” du dê “star.” The picture is further complicated by the evidence from Lati, which has m-tšu “eye,” m-tšu “month” (on different tones as recorded by Bonifacy), but m-tšua “moon” and simply tšua “star.”
11. Laqua rō, Li yāi~t’ai, N. Kelao rau, Latu lu, IN *talinga’ “ear.” For Laqua rō < *talinga’, cf. Laqua rre, Annamite ngūa “horse” (Laqua -d< -a; see the analysis above). The Li forms point to an original *nāi or *nūi with palatalized nasal initial, whence S. Li yāi (through further palatalization) and N. Li t’ai (n->t- is the regular N. Li shift). This reconstruction is supported by two outside comparisons, one with Thai (Dioi) and the other with IN, as shown in the table below (Central Li from Savina, Shaved Head Li and White Sand Li from Stübel):

<table>
<thead>
<tr>
<th></th>
<th>S. Li</th>
<th>C. Li</th>
<th>Shaved Head</th>
<th>White Sand</th>
</tr>
</thead>
<tbody>
<tr>
<td>ear</td>
<td>*talinga’ (IN)</td>
<td>yāi</td>
<td>t’ai</td>
<td>t’ai</td>
</tr>
<tr>
<td>finger</td>
<td>*niang (Dioi)</td>
<td>yeng</td>
<td>{tėng}</td>
<td>{t’eng}</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>snake</td>
<td>—</td>
<td>ya</td>
<td>t’a</td>
<td>t’a</td>
</tr>
<tr>
<td>yellow</td>
<td>*kuning (IN)</td>
<td>{yēng}</td>
<td>—</td>
<td>{t’eng}</td>
</tr>
<tr>
<td></td>
<td></td>
<td>hiēng</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

12. S. Li (hāi) p’en, S. Kelao du pio, N. Kelao bang, IN *tōn “tooth.”
13. Li k’ok, N. Kelao k’au, IN *kakī “foot.”
14. Lati tšu, IN *t’u’u “breast.”
15. Li dat~tlat, IN *darah “blood.” For the final, cf. Li p’uot~fuot, IN *puluh “10.”
16. Laqua nen “fat” (n.), S. Kelao nu, Lati m-ngā “oil,” S. Kelao nu xmüă, Lati m-ngā me “fat” (“oil of pig”), IN *miṅak “oil”~*məṅak “fat.”
17. Laqua küön, N. Li k’an, S. Kelao kā mōn~mōn kā (mōn perhaps for mo “rice”), N. Kelao ka, Lati k’o, IN *ka’~*ka’om~*ka’i “eat.” Note also Laqua ngōm “drink,” IN *pangan “eat” (cf. Lati k’o “eat,” also “drink”).

18. Li sop~sap, Lati (ngua) so, IN *rabi “night.” Cf. also Javanese strāp “twilight,” and the equation Li s= Thai hr- analyzed above.

19. Li negi, IN *tangit “weep.”

20. Laqua tie, IN *matay~*patay “die.” For the vocalism, cf. Laqua te, IN *mata “eye.”

21. Li sop-sap, Lati (ngua) so, IN *rabi “night.” Cf. also Javanese strāp “twilight,” and the equation Li s= Thai hr- analyzed above.

22. Laqua tie, IN *matay~*patay “die.” For the vocalism, cf. Laqua te, IN *mata “eye.”

23. Laqua nin, S. Li yēng~hieng, N. Li t’ēng, S. Kelao t’é ni, N. Kelao nyi, Lati o-hni (recorded as an hi), IN *kuning “yellow.” Li *neng or *nieng>yēng ~t’ēng (vide supra); *nieng>*kuning, with medial diphthong as in No. 21.

24. Li tik~tik, IN *ni(n)’i~*i(n)’i “small.”

25. Li nei, IN *’in “this.”

In addition to the above, a number of significant correspondences exist within the Kadai stock itself, thus serving to tie the group together. The more important of these lexical agreements are listed below:

1. Laqua nen, Li nān “moon.” Possibly related to IN *bulan “moon”; thus, *bulan>*wulan>*dan (paralleling *wulu>*du “8,” *ulu>*du “head”) >nān (through assimilation to the final nasal). Cf. Li nuk, IN *bölük “monkey.”


3. Laqua mōn dōng, Li pā dang om, S. Kelao zū dōng “thunder.”

4. Laqua hōng, S. Kelao ngā-ye “water” (but zōng ngū̀ “drink”); N. Kelao u, Lati i “water”; cf. N. Kelao du, Lati m-ti “earth.” S. Kelao retains the element u in the compounds u ngē uđ “tears” (Lati i m-ṭsu), i lā-ṭu “milk” (Lati i t’su). A possible comparison exists with IN *wawar “water.”

5. Laqua dōm, Li som “fruit”; S. Kelao mā, Lati mì “fruit.”


8. Laqua mā~māi, Li mei, S. Kelao mu (vē), Lati mia “female, mother.” Cf. Thai *me “mother.”

10. Laqua mun, Li müöm~ţūm “beard” (cf. the treatment of nasal finals in the foregoing example). Cf. Thai *mâm “beard” (only in the northern Thai speeches: Dioi mum, Tho kang mun, Nung mom).

11. Li ClientRect[1]’a mòu, N. Kelao mau “hand.” Cf. Thai *mù “hand.”

12. Laqua nie, S. Li yeng, N. Li l’eng~lêng “finger” (Li *nieng, vide supra). Cf. Dioi (Thai) niêng “finger.”


14. Laqua ì, S. Kelao ìtu “urine.”

15. Laqua nung, Li ñau, S. Kelao ñu, N. Kelao nyo, Lati a-tnu “salt.” Cf. IN *'uyah “salt.”

16. Laqua yeu, S. Kelao ha, N. Kelao a, Lati ho “meat, flesh.” Li has the puzzling forms mam~am.

17. Laqua mâi, Li mau~pau “year.”

18. N. and S. Kelao eu, Lati eu “go.”

19. S. Li müön, N. Li püön~pôu, S. Kelao xm, N. Kelao mu “come.” Cf. Malay mari, Cham mòrai~mai “come,” indicating the development IN medial ~ Li -n, as in *wari’ > sen “day.”

20. Laqua neng, Li dêng~t’êng “red.” Cf. Thai deng “red.”


We have, finally, to consider the nature of the affinity of Kadai and Indonesian with the Thai group of languages, spoken over a wide area in southern China, Siam, French Indochina, Burma, and Assam. The Thai family includes Ahom, Khamti, and Shan, in the west; Siamese and Lao, in the south; White Tai and Black Tai, in the east; Nung and Tho, in the northeast; and Dioi, in the north. Despite the geographical extent of this group, the several languages are closely interrelated, and thus rather detailed reconstructions of the parent speech can be made. The earliest systematic investigation in this field was Maspero’s study of the Thai tonal system.²⁴ This study was supplemented by several brief articles by G. Coedès and J. Burnay,²⁶ but no comprehensive review of the problem appeared until almost a quarter of a century later, when K. Wulff published his monumental work on Chinese and Thai.²⁸ The writer has further extended the analysis undertaken by Wulff and has filled in cer-

²⁶ The most important of these are: Note sur les tons et les initials du vieux siamois à l’époque de Sukhodaya (Journal of the Siam Society, Vol. 21, 1927), pp. 103-117); [v] et [x] et leur origine (Ibid.), pp. 119-126.
tain lacunae in that scholar's work, so that our present knowledge of Thai phonology may be regarded as reasonably complete.

As regards the affiliations of Thai, the generally accepted view has been that Chinese and Thai constitute a single "Eastern" division of the Sino-Tibetan or Indo-Chinese stock, in opposition to the Tibeto-Burman or "Western" division. It is this view that has been developed by Maspero, Wulff, and, most recently, R. Shafer (largely on the basis of Wulff's work), and that has given rise to attempts at direct Siamese-Tibetan comparisons, such as those of O. Schrader. The writer must plead guilty on the same charge, though in modified form. Almost alone among students of the Thai languages, Coed and Burnay have evinced a healthy skepticism of the dogma of a Chinese-Thai relationship. Conrady, a pioneer in Far Eastern linguistics, sought to connect Indo-Chinese, including Thai, with the Austro stock (Mon-Khmer, Khasi, Munda, et al.) established by Schmidt, in terms of a "common substratum" (gemeinsame Unterschicht). Wulff, apparently under the influence of Conrady, has attempted to demonstrate the existence of infixes in Siamese, which he compares with those characteristic of the Austro languages. Of Conrady's proposed Indo-Chinese-Austro grouping, Wulff makes the following assertion:

The similarity of the formations [infixes] in both languages [Siamese, Javanese] rests not on chance, since the relationship of Malayo-Polynesian with Indo-Chinese, which Conrady sought to show with insufficient means, is certain (sicher). (cit. supra, p. 17, note 1)

27 Notably in Ahom, where Wulff failed to make use of the most important lexicon on that language (Borua, Ahom-Assamese-English Dictionary, Calcutta, 1920); Tho, completely neglected by Wulff but for which we have two utilizable sources, viz. E. Digué, Étude de la langue Thô (Paris, 1910), and Fr. Th. Gordaliza, Estudio sobre el dialecto Thô de la región de Lang-sôn (Anthropos, Bd. 3, 1908), pp. 512-532; White Tai, for which an excellent source has recently appeared, viz. G. Minot, Dictionnaire tày blanc-français (Bull. de l'École Française d'Extrême-Orient, t. 40, fasc. 1, 1940, pp. 1-237.)

28 The writer has in preparation a comparative dictionary of the Thai languages, based in large part on materials collected by the Sino-Tibetan Philology Project of the Works Progress Administration, sponsored by Prof. A. L. Kroeber of the University of California during the years 1935-40. The writer here wishes to record his indebtedness to Prof. Kroeber for having made possible this investigation of Far Eastern languages, of which the present paper may be regarded as a by-product.


Maspero, in a review of Wulff's work,\textsuperscript{32} has convincingly dismantled the thesis of Thai infixation, and with it much of the Conrady-Wulff hypothesis. A similar hypothesis has been brought forward by J. Przyluski in the well-known account in \textit{Les Langues du Monde}.\textsuperscript{34} Przyluski suggests that Thai is transitional between Sino-Tibetan (Chinese and Tibeto-Burman) and Austro-Asiatic, yet offers no support for this view, other than a few comparisons of demonstrative pronouns in Siamese, Annamite, Khasi, and Palaung.

The writer's conclusions differ significantly from any of the above. The thesis presented here holds that Thai has a truly genetic linkage with Kadai and Indonesian rather than with Chinese and Tibeto-Burman (Sino-Tibetan), but has undergone extensive modification under Chinese influence. A similar view was propounded many years ago by Gustav Schlegel, in a highly unsystematic and unscientific fashion.\textsuperscript{36} Schlegel was unaware of the existence of the Kadai group, but pointed out many analogies with Malay, and in general seems have been on the right track. The writer has developed the present hypothesis entirely independently of Schlegel, and largely as a by-product of his own synthesis of the Kadai stock.

The writer has long been aware of the fact that the lexical resemblances between Chinese and Thai are of a restricted range which fails to support the generally held view of a genetic relationship between the two languages. A careful analysis of the material assembled by Wulff, in addition to his own supplementary material, has made this fact still clearer. The primary lexical agreements lie in the numerals, especially from "3" to "10" and "100," a few words for parts of the body, certain animal names, and a number of terms for cultural objects and the like. Let us examine these loosely defined categories in some detail.

The Thai numerals from "3" to "10" and "100" are in fairly close agreement with the Chinese: Thai *Sam, Ch. *säm "3"; Thai and Ch. *si "4"; Thai *ha, Ch. *nguo "5"; Thai *hrôk, Ch. *liuk "6"; Thai *tîet, Ch. *tîiêt "7"; Thai *pet, Ch. *pwat "8"; Thai kôu, Ch. kiu "9"; Thai *sîp, Ch. *siêp "10"; Thai *pak, Ch. *p阿k "100." Thai ordinarily agrees with Chinese as opposed to Tibeto-Burman, yet shows interesting variations in the direction of the latter, e.g. Thai *ha, TB *l-nga "5," with h-<ng- as in Thai *han, Ch. *nian, TB *nian "goose"; Thai *hrôk, TB *d-rug "6." The Chinese forms for "7," "9," and "10" illustrate the diphthongization characteristic of that language, the -du = -iau equation being particularly well attested: Thai *kôu, Ch. *k'iau "hill"; Thai *k'ôu, Ch. *kiou "pigeon"; Thai *kôu, Ch. *g'iau "old"; Thai *gôu, Ch. *g'iau "owl" (Tibeto-Burman has final -u in this series). On the basis of the above phonetic evidence, the borrowing of this numeral system must be

\textsuperscript{34} Edited by A. Meillet and M. Cohen (Paris, 1924), pp. 361-384 (\textit{Le Sino Tibétain}).
\textsuperscript{36} \textit{Siamese Studies} (T'oung Pao, t. 2, n. s., 1902, Supplement).
assigned to an early period antedating the $r \rightarrow l$, $-a \rightarrow -uo$ shifts in Chinese. The Thai forms are still irregular, however, with *pet "8" rather than *pud, and *st$p "10" rather than *si$p, and cannot be reconciled on any scheme of genetic relationship (for the latter, cf. Thai *suk, Ch. *tiuk "ripe."

The regular Thai numerals for “1” and “2” are *hnu$ng and *song, respectively, which appear to be remnants of the original Thai numeral system. The corresponding Chinese terms, significantly enough, appear only in the combinations *si$p t “11” (Ch. *"iu “1”), and *ni st$p “20” (Ch. *"i < *ni “2”). In addition, a basic root *d$au “20” is preserved in Lao and the western Thai languages, and *roi “100” is found in Siamese, Lao, and some of the eastern Thai languages. The evidence from the numeral system, therefore, cannot be held to speak in favor of the theory of a genetic Thai-Chinese relationship.

The common roots for parts of the body are as follows: Thai *kien “arm,” Ch. *kien “shoulder”; Thai *veng “leg,” Ch. *yieng “shin, shank”; Thai *fa, Ch. *pa “palm of the hand”; Thai *e$u, Ch. *i$au “waist, loins.” In this group belong also Thai *hnu$ng, Ch. *nu$ng “pus”; Thai *nya, Ch. *nie “urine, urinate.” Thai *nga “tusk, ivory,” Ch. *nga “molar tooth,” must be considered in relation to the loan-word for “elephant” (infra), the regular Thai roots for “tooth” being *k$ria and *o$an. Similarly, Thai *pi$u “cuticle, epidermis” is connected with Ch. *piu “skin, hide,” but the regular Thai root, *hnu$ng, “skin, hide,” has no Chinese correspondence. Basic roots for parts of the body such as “eye,” “nose,” etc. are significantly lacking in this list.

The group of common roots for animal names is equally enlightening in its exclusiveness. Here we find Thai and Ch. *ma “horse,” and the associated roots: Thai *an, Ch. *"an “saddle”; Thai *k’i, Ch. *g’yie > k’i “ride (a horse).” These correspondences strongly indicate that the Thai peoples borrowed the horse-complex directly from the Chinese. This group also includes Thai *d$ang, Ch. *ti$u “elephant,” Thai *ngua, Ch. *giu “bull, ox, cow,” Thai *d (restricted distribution), Ch. *i$u “hare”; Thai *k$i, Ch. *kie “fowl,” Thai *p’u$ng, Ch. *p’iu$ng (equivalent to *p’i$u “bee”; and, from the above dis-

36 Cf. the penetrating study by Coedès and Burnay, Notes d’étymologie Tai, No. 1: Le nom de nombre “Cent” (Journal of the Siam Society, Vol. 20, 1926), pp. 49–52. Coedès and Burnay identify *roi “100” with the root *pak “to string.” They further conclude that *pak “100” is common Thai because of the concordance of tones, yet admit the possibility of its having been borrowed from Chinese by the parent Thai speech.

37 Cf. the associated correspondence between Thai *nga “tusk, ivory” and Ch. *nga “molar tooth.” The root for “elephant” has a restricted extension in Tibeto-Burmese (Burmese Is’ang).

38 Ch. *ni$u should have produced Thai *ngua rather than *nga. The latter corresponds rather to the root *nga “bull, ox, cow” of the Kachin-Nung-Burmese division of Tibeto-Burmese.

39 For the finals, cf. Thai *g$ai “who, which,” Ch. *hei “why, how, what”; Thai *k’$i, Ch. *k’$i “to open.” The Thai root for “egg” (*k’r$ai) is independent of the Chinese roots (*lu$u, *d’u$u).

40 The interesting Chinese root *mit “honey,” an ancient loan-word from Indo-European (Sanskrit madhu; Old Slavic met a; Tocharian mit, whence Ch. *mit through diphthongization;
cussion, Thai *han, Ch. *ngan “goose (wild)”;
Thai *gdu, Ch. *g’iu “owl”; perhaps also Thai *nguk “crocodile, dragon, siren,” Ch. *nguk “crocodile.” Significantly lacking are roots for “dog,” “fish,” “bird,” “snake,” and the like.

The fourth and last of the groups mentioned above includes Thai *ngön, Ch. *ngienn “silver” (Tibetan dngul); Thai *gram (restricted distribution), Ch. *lom <*glam “indigo” (Tibetan rams);42 Thai *tšid, Ch. *t’sie “paper”; Thai *hmük, Ch. *mok “ink”; Thai *bái, Ch. b’ai “cards (for playing).” Here also may be placed Thai gêm, Ch. *iom <*giäm “salt,” and Thai *gödn, Ch. xiuan “smoke”; it should be noted that the regular Thai root for “salt” is *klúa rather than *gêm. It is apparent that no great importance can be attached to this group of roots.

The above lists of the principal Thai-Chinese correspondences have been carefully drawn up, and should give an accurate picture of this aspect of the problem. There are, to be sure, additional correspondences, some of which have been cited above, but these hardly affect the picture as a whole. Below, by way of contrast, are listed the basic Thai-Indonesian correspondences on which our conclusions have been built. That these are truly basic correspondences as compared with those between Thai and Chinese is sufficiently clear even after a cursory inspection of the material.

1. Thai *wän “day,” ta wän “sun” (“eye of the day”); IN *wari “day, sun”; Laqua vuon, Li ven “day, sun.”
2. Thai *blüdn “moon”; IN *bulan; Laqua nen, Li nän.
3. Thai *dau “star”; IN *a(n)daw-*ha(v)gaw “sun”; Li tshêm drau “star.”
4. Thai *fôn “rain” (“fine rain,” as opposed to *hra “heavy rain, shower”); IN *a(m)bun “atmospheric precipitate” (Tagalog ’ambon “fine rain”); Li (pa) pûn “rain.”
5. Thai *nöm “water”; IN *danum: Li nom~nam.
6. Thai *vai “fire”; In* ’apuy; Laqua pâi, Li pêi~fei.
7. Thai *na “rice-field”; IN *bana “low-lying land, flooded land”; Laqua ne, Li na~ta “rice-field.”
8. Thai *nök “bird”; IN *manuk “fowl, bird”; Laqua nuk “bird.”
9. Thai *rông “nest”; IN *t’alang (Toba-Batak, Javanese, Malay, Dayak sarang).

Greek μεθο “wine”; English mead), is not found in Thai, which makes use of the periphrasis “bee-water” (Siamese and Shan nâm p’t’ung), or even equates “bee” with “honey” (Ahom, Tho) or with “wax” (Lao).

41 In view of the correspondences for “fowl” and “goose,” it is somewhat surprising to find distinct roots for “duck” (Thai *pté, Ch. *ap).

42 The prototype must have been *ram-s rather than *gram-s, since the latter could have yielded only *grams in Tibetan. On this line of reasoning, Chinese *lom <*glm includes a prefixed g- element, and the Thai borrowing can thus be dated as posterior to this prefixation, but anterior to the subsequent *gräm >*glm >*lom development in Chinese (completed ca. 500 B. C.).
10. Thai *rud “boat”; IN *pa³ahu (Malay porahu “prau”); Li da.
11. Thai *tu “door” (often in composition with *pak “mouth, opening”); IN *pinu; Laqua tu.
12. Thai *hru³ “head”; IN *ulu³~hulu³; Laqua ru, Li du³o.
13. Thai *ta “eye”; IN *mata³; Laqua te, Li sa.
14. Thai *dang “nose”; IN *ug³’ung~*ig³’ung (Malay hidong, Cham idung); Laqua lang.
15. Thai *vun “tooth”; IN *i³pon; Li (hai) p³en.
16. Thai *pot “lungs”; IN *put³’uh “heart” (Tagalog puso³, Toba-Batak pusu³ “heart,” Javanese pusuh “lungs”).
17. Thai *g³ng or *gr³ng “body”; IN *daging “body, flesh.”
18. Thai *tuk “bone”; IN *ta³(n)duk “horn”; Li dru³~f³uk “bone.”
19. Thai *l³u³t “blood”; IN *darah; Li dat³lat.
20. Thai *m³n “fat, oil”; IN *mi³nak “oil”~*m³nak “fat”; Laqua nen “fat,” Lati m-ng³ “fat, oil.”
21. Thai *pu³, “grandfather”; IN *s³(m)pu³ “grandfather, grandchild” (reciprocal term).
22. Thai *d³m “black”; IN *i³(n)tem “black,” *d³tem~*t³tem “dark”; Laqua d³m, Li d³m “black.”
23. Thai *s³m “sour”; IN *a³tem.
24. Thai *b³t “blind”; IN *b³ta³.
25. Thai *tai “die”; IN *mat³ay~*pat³ay; Laqua tie.
26. Thai *d³p “raw, green, alive” (Ahom has the doublet forms dip “living, to be alive,” lip “unripe”); IN *h³d³p “live”; Li die³~f³e³ “raw.”
27. Thai *kin “eat”; IN *ko³~*ko³³~*ko³³; Laqua k³³n, Li k³n.
28. Thai *t³t “flatus ventris”; IN *s³(n)t³³~*u³(n)t³³~*ko³(n)t³³; Li t³³ot.
29. Thai *k³ “I” (pejorative); IN *ak³³; Laqua k’d³u, Li hau.
30. Thai *ni “this”; IN *s³n³i³; Li nei.

In the above set of correspondences, the most obtrusive single feature is the development of monosyllabic roots in Thai from the disyllabic roots typical of Indonesian. It must not be assumed that all the roots involved were originally disyllabic, since in some instances monosyllabic roots can be postulated for proto-IN itself, e.g. *tem~*tem “black, dark” (No. 22), *tay “die” (No. 25), *ka³ “eat” (No. 27), *tut “flatus ventris” (No. 28), and cf. *danum “water” with *inum “drink,” apparently from a root *num.4 The real criterion here lies in comparison with Thai and Kadai, as well as with the more remotely related Mon-Khmer languages, e.g. IN *mata³, Thai *ta~*m-ta (vide infra), Mon-Khmer *mat (Annamite mát) “eye,” where *mata is the only feasible reconstruction for the parent stock. The material assembled in this paper lends itself

to the view that the majority of Thai-Kadai-Indonesian roots were disyllabic rather than monosyllabic, and that Thai and Kadai have undergone extensive phonetic reduction. The writer has elsewhere called attention to a parallel reduction in the aberrant Cham dialect spoken on the island of Hainan, and to similar phenomena in the standard Cham speech of the mainland, e.g. Cham *bulan~lan* "moon," *apuēi~puēi* "fire" (through aphaeresis). This aspect of Thai-Kadai phonology, therefore, calls for no especial demonstration.

The varying types of phonetic development shown by the Thai roots under consideration are in part explicable in terms of stress variations. Kadai offers an excellent instance of this in the bifurcate development shown by S. Li *du*, N. Li *au* "8," from IN *walu*, where we must reconstruct as follows: *walū* > *wlu* > *du*, *walū* > *wau* > *au*. Similarly, for Thai we must postulate shifts of the type: *danūm* > *nām* "water," *pintū* > *ti* "door," *matā* > *ta* "eye," but *pālūh* > *pol* "lungs," *mokā* > *mān* "fat, oil," *būla* > *bot* "blind," etc. The stress seems normally to have been on the last syllable, but certainly not always so.

The finals of this group of roots present fewer problems than the initials. Among the vocalic finals, we have Thai *-a* = IN *-a* (Nos. 7 and 13); Thai *-u* = IN *-u* (Nos. 11, 21, and 29); Thai *-i* = IN *-i* (No. 30); Thai *-au* = IN *-aw* (No. 3); Thai *-ai* = IN *-aw* (No. 25). Thai *vāi*, IN *apuēi* "fire" suggest a simple -āi = -ēuy equation, perhaps via an intermediate form -uei (cf. Cham *apuēi*), yet Thai has both -ui and -uei series, the latter of some importance. Two independent bits of evidence indicate that Thai *vāi* was developed from a root *vi*, probably via an intermediate form *vėi*, thus paralleling the development shown by Thai *kāi*, Ch. *kēi* "fowl" (see note No. 39). Firstly, Dioi and a group of related dialects have the form *fi* rather than the regular *fāi* (contrast Dioi *kāi* "fowl"); secondly, Li has *pei~fei* (cf. the discussion above), rather than *pāi~fāi* (contrast Li *kāi*, Thai *kāi* "fowl"); Li *lai*, Thai *kōi* "far"). The lone possible analogy here is furnished by Thai *hāi*, Li *ngei*, IN *tangit* "weep," with Thai *h* < *ng* as discussed above.

Nos. 10 and 12 must be considered in relation to each other. These two comparisons are, admittedly, uncertain ones, but the parallelism between them, extending even into Kadai (Li), has led to their inclusion in our list of correspondences:

<table>
<thead>
<tr>
<th></th>
<th>Indonesian</th>
<th>Thai</th>
<th>Kadai (Li)</th>
</tr>
</thead>
<tbody>
<tr>
<td>boat</td>
<td>*pajahu</td>
<td>*rūd</td>
<td>*da</td>
</tr>
<tr>
<td>head</td>
<td>*hulu</td>
<td>*hrūd</td>
<td>*du</td>
</tr>
</tbody>
</table>

---


45 The form *fi* "fire" is cited for the Tai Yoi, Kon Yai, and To-jen dialects in W. C. Dodd, *The Tai Race* (Cedar Rapids, Iowa, 1923), word-lists on pp. xiv–xxi.
The aspiration in these roots seems to have played a role in the \( l \rightarrow r \) shift. For the final of Thai *rūdā "boat," a possible parallel exists in IN *'alū́' "dog," Thai *sū́dā "tiger."

The consonantal finals are regular for the most part. Final -\( r \), which is lacking in Thai, is replaced by -\( n \) (*warī' \( \rightarrow *\)wān "sun, day"), as in loan-words from Khmer or Pali, e.g. Siamese kānun < Khmer k'nur "jack-fruit." In this connection, cf. Siamese and Lao pun, IN *'apur \( \sim *\)kapur "lime" (probably a loan-word in these southern Thai speeches). Final -\( h \), also foreign to the phonemic system of Thai, is represented by -\( l \), as in Kadai (*darah \( \rightarrow *\)lū́d "blood"). Final stops and nasals are preserved in Thai, with the exception of -\( t' \) > -\( t \) (No. 16). The most likely instance of interchange of nasal finals is furnished by IN *rūmāh, Li dùön, Thai *rūdn "house," perhaps via the forms *ruam > *ruan.

Short medial vowels are predominant in the Thai roots under consideration, and must be regarded as characteristic of these basic roots as a group. IN medial -\( a \) is represented by -\( ô \) (Nos. 1, 2, and 9); cf. also IN *balakang, Thai *hōnōng "back" (n.). Thai ordinarily has -\( ô \) for IN medial -\( a \) (Nos. 15, 20, and 22), yet has -\( ô \) in one instance (No. 23); for the latter, cf. IN *tīrem, Siamese hōi iŏm "oyster," undoubtedly a loan-word in Siamese (in composition with the Thai root *hōi "shellfish"). IN medial -\( u \) is represented by -\( o \) (Nos. 5 and 14) and -\( ò \) (Nos. 4, 8, and 28), as well as by -\( u \) (No. 18). After the labial stop initials \( p \) and \( b \), Thai has -\( o \) rather than -\( ò \) for IN medial -\( u \) (Nos. 16 and 24). An additional equation is furnished by Nos. 17 and 26, yielding IN medial -\( i \) = Thai -\( i \). The medial vowel of Thai *kīn "eat" (No. 27) cannot be satisfactorily explained on the basis of our present knowledge, though the contrast with the -\( a \) vocalism of Li is matched by Thai *din, Li dāb ~ jan "earth," perhaps related to IN *tanah \( \sim *\)tanah "earth, land."

The treatment of initial consonants in Thai presents a number of interesting features. The first of these to come to the writer's attention is the peculiar aspiration of the Thai roots for "eye" (No. 13) and "die" (No. 25) in the Tho-Nung group of dialects.\(^{47}\)

<table>
<thead>
<tr>
<th></th>
<th>IN</th>
<th>Laqua</th>
<th>Siamese</th>
<th>Tho</th>
<th>Nung</th>
</tr>
</thead>
<tbody>
<tr>
<td>eye</td>
<td>*mata'</td>
<td>te</td>
<td>tā</td>
<td>t'a</td>
<td>t'a~ha</td>
</tr>
<tr>
<td>die</td>
<td>*malay</td>
<td>tie</td>
<td>tai</td>
<td>t'ai</td>
<td>t'ai~hai</td>
</tr>
</tbody>
</table>

\(^{46}\) Note Thai medial -\( ô \) < -\( u \) only before final nasals. Thai medial -\( ô \) < -\( u \) seems to be the normal development before final stops; cf. Thai *hrōk, Chinese *liuk "6"; Thai *mōi, Malay, Javanese, Karo sēmu "ant" (cited in Brandstetter, cit. supra, p. 37). For Thai medial -\( u \) < -\( u \) before final velar stop, cf. Thai *luk "anything round," IN *kliuk \( \sim *\)paluk "bend, curve."

\(^{47}\) Our sources for Tho and Nung are in agreement on this point, and there can be no doubt as to the reality of the phenomenon in question. An additional check is furnished by the form t'a "eye" cited for a dialect of Tho-Nung type by Yu Wēn, *A Vocabulary of a Non-Chinese Tribe inhabiting the Tipingfu Area of Kwangsi Province, with Chinese Transliterations and Notes* (Academia Sinica, Bulletin of the Institute of History and Philology, Vol. 6, pt. 4, 1936), pp. 505-552 (in Chinese).
With one partial exception, these are the only roots so treated in Tho and Nung, hence this phenomenon cannot be explained in terms of Thai itself. On the basis of Indonesian, however, we can postulate a development of the type: *mat岩石 > *m-la > *m-t’a > t’a-ha; *matay > *m-tay > *m-t’ay > t’ai-hai, with secondary aspiration after the nasal prefix. The Li form sa “eye” can be explained along the same lines. The remarkable parallelism shown in the treatment of these two roots constitutes perhaps our most significant single piece of evidence for a Thai-Indonesian linkage.

The reconstruction of initial bl- for Thai, as in the root *blidn, represents a new advance in Thai phonology. As ordinarily reconstructed, Thai has initial bl-, pl-, p’l, and br-, pr-, pr’, but neither bl- nor br-. The typical Thai initial d-series shows the following equation: Siamese and Lao d= Ahom, White Tai, Tho, Nung, Dioi d= Shan and Black Tai l= Khamti n- (vide supra). Three roots, however, diverge from this equation in the direction of the initial b-series, and in one of these roots bl- is actually preserved in the archaic Ahom language, hence we can safely reconstruct all three roots with initial bl-:

a. Siamese and Lao dok “flower,” but Ahom blok, Tho biok, Nung beok


Reconstructions: *blok “flower,” *blūn “moon,” *bli “bile.” Note that initial bl-, which is of labial type, is best preserved before the labial vowel o, and worst preserved before the front vowel i; also that Black and White Tai preserve b- in all three roots. Initial br- cannot be reconstructed for Thai, and may be represented simply by b-; cf. Thai bōm “ripen fruits,” IN *polom “ripen fruits artificially” (Toba-Batak porom, Malay poram), perhaps via a form *porom.

Some interesting equations appear among the stop consonants, especially

---

48 Tho and Nung t’en “wasp” correspond to the general Thai root *t’en, but the doublet form in initial h- is lacking in Nung. The regular development with unaspirated initial t- is observed in a long series of Thai roots, including *tōp “liver,” lām “low,” tōng “glue,” tōu “turtle,” lāt “cut,” lēm “full,” līn “foot,” tōk “fall,” lōm “mud,” lōm “tree trunk,” lōt “flatus ventris,” and lōd “animal.”

49 A good parallel here is furnished by Tibetan, which has aspirated all initial surd stop or affricate consonants after prefix m-, e.g., Tibetan mūl’in “liver,” corresponding to Tibeto-Burman *m-sin.

50 For Thai *blok “flower,” cf. the subsidiary IN root represented by Bisaya bolak, Tagalog bulaklak “flower,” which Brandstetter (cit. supra, p. 22) derives from a root *lak “to unfold.” No IN comparison has been uncovered for Thai *bli “bile.”

51 This comparison is semantically too specific to be trusted, and we should expect Thai *pōm rather than *bōm. IN l> r as in *polahu > *rūd “boat,” *palang > *rāng “nest”; IN a> 5 as in *at’om > *sōm “sour.”
in the labial series. Thai ordinarily has \( t < t \) (Nos. 11, 13, 25, and 28), and \( d < d \) (Nos. 3, 18, and 26), while the correspondence shown in Thai *\( \dd ng \) IN *\( \dd g 'ung \) "nose" (No. 14) must be considered in connection with the IN doublet forms *\( a(n)daw \sim ha(y)g 'aw \) "sun" (No. 3). Thai *\( \dd m \) "black" must therefore be equated directly with IN *\( d \dd m \sim \dd m \) "dark" rather than with *\( \dd i(n) \dd m \) "black" (No. 22). The palatal stop (\( t' \)) of IN is represented in Thai by \( s \)- as an initial (No. 23), but by \( -t \) as a final (No. 16). The velar correspondences are regular: Thai k = IN k (Nos. 27 and 29); Thai g = IN g (No. 17). In the labial series, however, we find two types of correspondences, viz. Thai *\( \dd \) = IN initial \( b \)- (Nos. 2 and 24), Thai \( p = \) IN initial \( p \)- (No. 16), but Thai \( f = \) IN medial -\( b \)- (No. 4), and Thai \( v = \) IN medial -\( p \)- (Nos. 6 and 15). Thai *\( \dd m \), IN *\( \dd m \) "grandfather" (No. 21) would seem to run counter to this scheme, but in this instance IN has the doublet roots *\( \dd m \) "forefather, sir" and *\( \dd m \) "sir," the latter evidently the basic etymon from which Thai *\( \dd m \) was derived.

This explanation of Thai \( f \) and \( v \) as secondary phonemes derived from medial labial stops clears up one of the most abstruse aspects of Thai phonology. Li (southern dialect) and the Kadai languages in general have preserved the labial stop in these roots:

<table>
<thead>
<tr>
<th>Indonesian</th>
<th>Li</th>
<th>Thai</th>
</tr>
</thead>
<tbody>
<tr>
<td>rain</td>
<td>( *\dd m )( \dd n )</td>
<td>( \dd n )</td>
</tr>
<tr>
<td>fire</td>
<td>( *\dd )( \dd y )</td>
<td>( \dd i )</td>
</tr>
<tr>
<td>tooth</td>
<td>( *\dd m )</td>
<td>( \dd n )</td>
</tr>
</tbody>
</table>

It is a striking fact that, in the above set of comparisons, initial \( b \) and \( f \) appear to the exclusion of the sonant stops \( b \) and \( d \). An examination of the stock of Thai roots assembled by the writer shows an overwhelming predominance of basic roots with initial \( b \) and \( d \), some of the most important of which are listed below:

Initial \( b \)-: *\( \dd on \) "arum," *\( \dd li \) "bile," *\( \dd n \) "bridegroom," *\( \dd \) "butterfly," *\( \dd k \) "carry (on shoulders)," *\( \dd t \) "cloud" (v.), *\( \dd h \) "cowrie shell," *\( \dd u \) "crucible," *\( \dd o \) "pit, well, mine," *\( \dd l o k \) "flower," *\( \dd i n \) "fly" (v.), *\( \dd i t \) "fish-hook," *\( \dd a i \) "leaf," *\( \dd u \) "light (not heavy)"), *\( \dd u \) "lotus, water-lily," *\( \dd a \) "mad," *\( \dd a \) "shoulder," *\( \dd o k \) "speak," *\( \dd o k \) "tube," *\( \dd u e i \) "cocoanut spoon," *\( \dd n \) "village," *\( \dd a i \) "wound," *\( \dd o n g \) "hole," *\( \dd o t \) "blind," *\( \dd u \dd n \) "moon."

Initial \( d \)-: *\( \dd u \dd \) "boil" (v.), *\( \dd e k \) "child," *\( \dd n \) "earth," *\( \dd d \) "extinguish," *\( \dd u \dd \) "fig," *\( \dd m \) "forest," *\( \dd i \) "good," *\( \dd a m \) "handle" (n.), *\( \dd on \) "high, hill," *\( \dd o n g \) "kind" (n.), *\( \dd a i \) "ladder," *\( \dd u \) "look," *\( \dd o i \) "mountain," *\( \dd o n g \) "parents of in-laws," *\( \dd o n g \) "red," *\( \dd m \) "shield" (n.), *\( \dd o m \) "smell" (v.), *\( \dd u t \) "suck,"

---

\(^{52}\) Dempwolff's reconstruction of \( t' \) rather than \( s \) for IN is open to criticism; cf. the review by A. Capell, in Bull. of the School of Oriental Studies, Vol. 9, 1938, pp. 459-462. Thus, IN *\( \dd a m \) "sour" is represented by Tagalog *\( \dd a s i m \), Toba-Batak *\( \dd a s m \), Javanese *\( \dd e m \), Malay *\( \dd a s m \), Dayak *\( \dd e m \), all with initial \( s \).

In contrast to this impressive array, the sets of roots with initial b and d seem restricted indeed. With initial b- we find *bē "goat" (but *bē in Lao and White Tai), *be "raft," *bu "mountain" (but Siamese has b'u, as in loan-words) *bān "seed, kind," *brāk "tomorrow," *bra "large knife," *bi "fat, big," *bi "elder sibling" (perhaps etymologically connected with the foregoing), *bo "father," while with initial d- we find *dong "belly," *drai "sand, gravel," *dak "leech," *điān "true, correct," *đang "road," and *do "weave." The contrast is so marked that one is tempted to conclude that roots with b- and d- belong to the older Thai-Kadai-Indonesian stratum, and roots with b- and d- to one or more younger superimposed strata, including Chinese loan-words such as *bāi<Ch. *b'āi "cards" (vide supra). The existence of the roots *bi "elder sibling" and *bo "father," with initial b-, does not constitute a conclusive argument against this view, inasmuch as the Thai kinship nomenclature as a whole appears to have no intimate connection with the Indonesian. The presence of initial b- or d- in a given root may even be used as supporting evidence for a proposed Indonesian comparison, e.g. Thai *ba, IN *bara' "shoulder" (*bara'>*bāa>*dā, contrasting with the development shown in *war'>*wān "day, sun"); Thai *dēk "child," IN *dīkh~*s(n)tīk~*iṭik "small" Li tik~tok "small."

Still another problem is presented by Thai *liūāt<IN *darah "blood" (No. 19), apparently via a form *dlat (dl- is not retained in Thai). A possible analogy here is furnished by the Thai root for "tongue," which the Li dialects treat in a parallel manner:

<table>
<thead>
<tr>
<th>Indonesian</th>
<th>S. Li</th>
<th>N. Li Shaved Head</th>
<th>Thai</th>
</tr>
</thead>
<tbody>
<tr>
<td>blood</td>
<td>*darah</td>
<td>dat</td>
<td>tlat</td>
</tr>
<tr>
<td>tongue</td>
<td>*dilah</td>
<td>dien</td>
<td>tlien</td>
</tr>
</tbody>
</table>

Yet Thai has *pla "fish," corresponding to S. Li da, N. Li tla, Shaved Head Li sla, with initials as in the above series, hence no certain conclusions can be drawn.

The above discussion does not exhaust the possibilities of the complex Thai-Indonesian field, and it is possible that a more searching analysis of Indonesian material will yield further comparisons, yet it is believed that most of the important lexical correspondences have been uncovered. The writer has eliminated from the discussion certain obvious loan-words in Siamese, e.g. muāng "mango"<IN *mānga. Attention should be called, however, to the noteworthy agreement between Thai *nga and IN *tōnga' "sesame." The Thai root *nga is widely extended in that stock (Siamese, Lao, Shan, Ahom, White Tai, Nung), hence cannot be regarded as a recent loan from Indonesian.
It is apparent that our judgment must be based almost entirely on lexical rather than morphological analogies, inasmuch as the rather elaborate affixation system of Indonesian is not represented in Thai. We must remember, however, that the reduction of disyllabic or trisyllabic roots to monosyllabic forms, as in Thai, necessarily involves the loss or incorporation of affixed elements. Thus, if a root *tay "die" be reconstructed for proto-IN on the basis of the doublet roots *matay-*patay, and the elements ma- and pa- be regarded as prefixes, the purely phonetic development *matay>*m-tay>*tai, paralleling *mata>*m-ta>*ta "eye" (vide supra), necessarily entails the loss of this prefixed element. In other instances, the affix may have been incorporated in the derived form; cf. IN *miṅak "oil">*māṅak "fat," Lati m-ngā, Thai *māṅ "fat, oil" (No. 20), and IN *ka~*ka’on~*ka’i, Thai *kīn "eat" (No. 27).

Of some interest in this connection are the traces of prefixes preserved in Siamese. Siamese prefix kā-, by far the most prominent of the lot, is found with a few words for parts of the body (kā-duk, kā-diau "bone," kā-do "male genitals," kā-bō~bō "stomach"), and with some animal names (kā-tai "hare," kā-tōk "sparrow," kā-te "tupaya"), but is characteristically associated with curious derived forms, e.g. bōng "stick"~kā-bōng "cudgel," dōng "oscillating"~kā-dōng "distorted, twisted," lān "push"~kā-lān "touch lightly," tūn "mole"~kā-tūn "kind of large rat." No great importance can be attached to this prefix, yet one possible IN correspondence has been uncovered, viz. Siamese kā-duk "bone," IN *ta(n)duk "horn"; cf. the kā~tā- interchange in kā-bōng~tā-bōng "mussel."

The problem of the development of tones in Thai cannot satisfactorily be handled until good material on Kadai tones is made available. As reconstructed, the Thai tonal system includes two series of tones, one connected with roots having surd initials, the other connected with roots having sonant initials (a similar system is found in Annamite and Chinese). Each of these series, furthermore, includes three tonemes, the original values of which are uncertain. It is probably significant that almost all the Thai roots having IN correspondences are associated with a single toneme, represented in Siamese by the mid-level tone (with sonant and unaspirated surd stop initials) or the high-rising tone (with other surd initials). The only exceptional roots here are *nām "water," *ni "this," *pu "grandfather," and *sām "sour.

Aside from the rudimentary prefixes found in Siamese, the Thai stock closely conforms to the classical type of monosyllabic, isolating languages. Maspero has successfully refuted Wulff's thesis of infixation in Siamese (see note No. 33), hence no comparison with Indonesian infixes can be made. As pointed out above, Thai agrees with Indonesian and Kadai, and sharply diverges from Chinese, in placing modifying elements after rather than before modified elements. This significant agreement in syntax contributes no little support to our Thai-Kadai-Indonesian hypothesis. Attention must also be called to the traces of a distinction between inclusive and exclusive forms for
the 1st pers. pl. pronoun in Thai, as represented by the exclusive form *tu “we” in Khamti, Lao, and archaic Siamese. This distinction is paralleled in Indonesian in the forms *kita “we” (inclusive), *kami “we” (exclusive).

The Thai-Kadai-Indonesian hypothesis, as outlined in the present paper, bears far-reaching implications for the history of the peoples of southeastern Asia and Oceania. If we accept the view that these three linguistic stocks are genetically related, we must place the center of their dispersion somewhere in the South China area, the present home of the Kadai tribes as well as the early home of the Thai peoples. On the basis of this distribution we can conclude, with a high degree of probability, that the proto-IN-speaking peoples migrated from the South China coast, perhaps via the island of Hainan, to Formosa on the north, the Philippines on the east, and Annam, Borneo, Java, Sumatra, and the Malay Peninsula on the south. The Cham and Malay linguistic areas, in southern Annam and the Malay Peninsula, respectively, surely are to be regarded as Indonesian enclaves on the Asiatic mainland, not as possible points of departure for the Indonesian migrations.

In still broader perspective, Thai-Kadai-Indonesian appears in its true light as the northern division of Schmidt’s Austric superstock. The archaic cleavage between Thai-Kadai-Indonesian on the one hand, and Mon-Khmer on the other, must have come about in the South China-Indochina area, with subsequent localizations of these two divisions in the north and south, respectively. The anomalous position of Malay at the present day, south of the main body of Mon-Khmer speeches, can be explained only on the basis of a sea-borne migration from the islands of Indonesia. Thai and Kadai in the north, Cham in the east, and Malay in the south, show a peripheral distribution with respect to the Mon-Khmer languages. As suggested above, Cham and Malay fall into their place in this picture as intrusive Indonesian languages overlying a Mon-Khmer substratum.

Annamite, too, takes its proper place as the northeasternmost member of

---

63 Cf. the discussion in G. Coedès, Nouvelles notes critiques sur l'inscription de Râma Khambeng (Journal of the Siam Society, Vol. 17, 1923), pp. 113–120.

64 The general Thai movement southward into Indochina appears to have begun on a large scale only toward the close of the first millennium A. D. The first group of Siamese inscriptions, the Sukhodaya, are from the 13th to 16th centuries, and the famous Râma Khambeng inscription, the earliest monument of the Siamese language, is dated no earlier than 1292; cf. G. Coedès, Notes critiques sur l’inscription de Râma Khambeng (Journal of the Siam Society, Vol. 12, 1918), pp. 1–27, and Recueil des Inscriptions du Siam; Première Partie: Inscriptions de Sukhodaya (Bangkok, 1924).

65 The writer accepts Schmidt’s postulation of an Austric superstock including Mon-Khmer and Austronesian, even though this relationship has not yet been thoroughly demonstrated. In the present instance, the Austric hypothesis is useful in interpreting certain roots which Thai and Mon-Khmer have in common, notably Thai *yo, Mon-Khmer *go (Annamite ko) “neck.” Cf. the Thai-Khmer comparisons listed in Wulff, cit. supra, pp. 68–70, and the Dioi-Khmer and Dioi-Bahnar comparisons in D. Doutreligne, Contributions à l'étude des populations Dioy du Lang Long (Anthropos, Bd. 26, 1931), pp. 35–53.
the Mon-Khmer stock. Annamite stands in relation to Mon-Khmer somewhat as Thai stands in relation to Indonesian. Like Thai, it has suffered much phonetic attrition, has developed a complete tonal system, and has lost its morphological apparatus of affixes. These changes must be attributed to Thai influence, in view of the not inconsiderable body of Thai roots in the language. The overwhelming majority of basic roots, however, are of Mon-Khmer rather than Thai origin. On the analogy of our analysis of Thai, there can be no question as to the genetic nature of the Mon-Khmer-Annamite relationship.  

With Thai, Kadai, and Annamite in their proper settings, the linguistic picture of southeastern Asia assumes definitive shape for the first time. There remains only one linguistic problem of major importance, viz. the affinities of the Miao-Yao stocks of languages, spoken throughout much of central and southern China and northern Siam and Indochina. Our material on these languages is scanty and generally poor, and almost no comparative work has been done on the group. Miao and Yao are well differentiated divisions of a single stock, and each appears in a number of dialectical varieties, with Miao showing the greater variation. Pateng, spoken in the Rivière Claire section of Tonkin, is a subsidiary member of the stock. Miao-Yao resembles Thai-Kadai and Annamite in its monosyllabism and tonality, and further investigation may reveal a relationship with Proto-Austric or with one of its later divisions. A final judgment here must await the reconstruction of Mon-Khmer and the assembling of more material on the Kadai languages.

The proposed classification of Southeast Asiatic languages is as follows:

```
Proto-Austric
  /   \
 /     \  
Thai   Mon-Khmer
|       |
Kadai  Annamite

Sino-Tibetan
  /   \
 /     \  
Chinese   Tibeto-Burman
  /     \  
Karen

?Miao-Yao
```

56 H. Maspero, Études sur la phonétique historique de la langue annamite (Bull. de l'École Française d'Extrême-Orient, t. 12, 1912), pp. 1-126, was so impressed by the monosyllables and tones of Annamite that he postulated a genetic kinship with Thai, even in the face of the dominant Mon-Khmer lexical element. Przyluski, in Les Langues du Monde (cit. supra), rightly breaks with Maspero on this point and classifies Annamite with Mon-Khmer.


On the ethnological side, the Kadai group offers the most promise for future investigation. At present, our material on this group is confined to the scraps of information gathered by Bonifacy and Lunet de Lajonquière, together with Stübel's fairly extensive study of the Li tribes (see sources cited above). Stübel points out a number of Indonesian and Micronesian parallels, e.g. in weaving (cit. supra, p. 293) and basketry (id., p. 294), and expresses his astonishment at the general cultural similarity to the tribes of Formosa (id., p. 296). It may be that the Li retain certain Indonesian culture traits that have been discarded by the Sinicized Kadai tribes of the mainland. It is fairly evident, however, that the general ethnological picture of Thai-Kadai-Indonesian has been destroyed beyond repair, and that our linguistic thesis must stand or fall on its own merits.69

WASHINGTON, D. C.

69 The writer has not had access to the most recent comparative study on the Li, viz. Chung-see Liu, Preliminary Study of the Origins of the Tribes of Hainan Island (Meridio-Occidentale Sinense, Vol. 1, No. 1, 1940), pp. 1–23.