THE POSITION OF THE MUNDA LANGUAGES WITHIN
THE AUSTROASIATIC LANGUAGE FAMILY

By HEINZ-JÜRGEN PINNOW

1. The Austroasiatic controversy, which arose in 1930 when Wilhelm
Hevesy 1 violently attacked Wilhelm Schmidt, 2 the real discoverer of the Austro-
asiatic family, still remains unsettled. Schmidt, it will be recalled, had grouped
the Munda languages, Khasi, and the Nicobar, Palaung-Wa, Mon-Khmer,
Malacca, and Cham languages together under the designation 'austroasiatisch-
Sprachstamm'. Some scholars, among them Sten Konow, 3 C. O. Blagden,
F. B. J. Kuiper, 5 R. Shafer, 6 J. H. Greenberg, 7 and A. J. Shevelenko, 8 in the main
agreed with Schmidt and favoured the recognition of the existence of an Austro-
asiatic family including the Munda languages but excluding the Cham languages.
Others, in particular Georges Maspero, 9 and more recently also A. I. Blinov,
rejected this contention. The majority of the critics, especially T. A. Sebeok
and M. B. Emeneau, 10 have deferred judgment in the belief that the problem can
be solved only through further investigation.

How unsettled the classification of the languages of South East Asia still
is can be seen in Richard Salzner's Sprachenatlas des Indopazifischen Raumes,
in which the Munda and Malacca languages are excluded from the Austroasian
family but which follows Schmidt in once again designating as Austroasiatic
Cham languages, which have long been recognized as Indonesian. The whole
question becomes even more complicated by the fact that the linguistic relations
ships of Vietnamese and of Műông are disputed. 14 Vietnamese, Műông, and the
Malacca and Cham languages are, however, all clearly mixed languages: they
all reveal at least an Austroasiatic substratum, which, however, is so small

1 BSOS, 6, 1930, pp. 187 ff.
2 Die Mon-Khmer-Völker, Braunschweig, 1906.
4 W. W. Skeat and C. O. Blagden, Pagan races of the Malay Peninsula, London, 1912,
7 ‘Historical linguistics and unwritten languages’, in Anthropology today, ed. A. L. Kroebel,
Chicago, 1953, pp. 265 ff.
8 Sov. V., 1958, pp. 101 ff.
10 Sov. V., 1956, 153 ff.
11 Language, 18, 1942, pp. 206 ff.
R. Shafer, BEFEO, 40, 1940, pp. 439 ff., HJAS, 6, 1942, p. 399, Đàn Việt Nam, No. 1, 1956,
pp. 1 ff.
Vietnamese, Müöng, and the Cham languages that we may unhesitatingly regard Vietnamese and Müöng as Thai languages and the Cham languages as Indonesian. In the Malacca languages, on the other hand, the Austroasiatic elements are so numerous that the languages are still to be regarded as Austroasiatic, although all of them (particularly Jakud) are gradually being assimilated into the Indonesian group.

The Munda languages, which are undoubtedly not mixed languages, do indeed differ greatly from the pattern of the other Austroasiatic languages (for which the designation ‘Khmer-Nicobar group’ seems suitable), but the points of correspondence of the Munda languages with this group are so numerous that these similarities can by no means be coincidental. Borrowing may also be ruled out, for the common features lie especially in the basic vocabulary, in word-formation (where the infix-formation is particularly characteristic), and in certain morphological and syntactical peculiarities. The typological differences between the two groups can in large part be explained as the results of influence by the various neighbouring languages, with which individual Austroasiatic sub-families have formed so-called ‘language leagues’ (Sprachbünde). A further discussion of this subject is to be found in my Versuch einer historischen Lautlehre der Kharia-Sprache ¹ and in an article in the Indo-Iranian Journal, 1960.²

2. Schmidt’s view that the Munda languages form a sub-family of the Austroasiatic group proves to be well founded and correct, but it still remains to be ascertained exactly what the relationships of the various members of the family (Munda, Nicobar, Khasi, Palaung-Wa, Mon-Khmer, and the Malacca group) to one another are, and a suitable classification of the Austroasiatic family must be worked out. Furthermore, the classification of the Munda languages themselves must be established. The unity of all the Munda languages (not including Nahali, which is not Munda) is apparent not only from a comparison of the phonological systems in these languages but even more from their highly developed inflectional systems, in which they differ from all other Austroasiatic languages. Recent investigation ³ has shown that the verbal inflection of all Munda languages is traceable to a Proto-Munda inflectional system, which was later expanded in the north and considerably reduced in the south. From this evidence and on the basis of lexical differences the Munda languages may be divided into a Northern group with the sub-groups Kurku and Kherwari (Santali, Mundari, Korwa, etc., belong to the latter branch), and a Southern group, which is further subdivided into a Central group (including Kharia and Juang), and a South-Eastern group (including Sora, Pareng, Gutob, and Remo). R. Shafer (in Word, 1960, p. 425)

¹ Wiesbaden, 1959.
has also divided Munda into Northern and Southern groups. The relative of Kherwari and Kurku is much closer than that of Central and South Eastern Munda, which must have been separated much earlier than Kherwari and Kurku.

This classification, the reasons for which are given in more detail in the investigation of the Munda verb mentioned above, differs greatly from the one classification of Schmidt, who grouped the languages solely on the distribution of k and h (from Proto-Munda *q). He concluded that there is an Eastern group (= Kherwari; with h), a Western group (Kurku, Kharia, Juang; with k), and a supposedly ‘mixed group’ (= South-Eastern Munda; with a loss of Proto-Munda *q). As this classification was based on a single argument, it could not do justice to the facts.¹

3. W. Schmidt likewise attempted a general classification of the Austroasiatic languages,² and again on the basis of a single, if important, factor. On the basis of the comparison of numerals and some other important words he came to the conclusion not only that the Munda languages are Austroasiatic but that they together with the Mon-Khmer languages (in the narrower sense of the term) form one of four Austroasiatic sub-families,³ the others being the older Malay group (including Semang and Sakai), the central group (including Khasi, Nicoleño, and Palaung-Wa), and a so-called south-east mixed group (i.e. the Cham languages). Schmidt classifies Besisi and Jakud as a younger Malacca group within the Mon-Khmer languages. His classification of the ‘South-East mixed group’ is not appropriate, as these languages are clearly Indonesian. Because of its one-sidedness, Schmidt’s classification is as a whole questionable and must be examined. However, in the following we will limit ourselves primarily to a reconsideration of the position of the Munda languages.

4. There are two ways of solving the problem of the classification of the Munda languages within the Austroasiatic family: first, according to structural synchronic considerations; secondly, according to the historical development of the languages. The latter classification is undoubtedly the more important one. Reliable information can be obtained only through such intensive historical-linguistic investigation of the various sub-groups as R. Shafer, for example has done in the field of the Palaungic languages.⁴ The work of W. Schmidt pioneering as it may be, does not suffice to clarify the complicated situation. As long as the eastern Austroasiatic languages have not been subjected to a comprehensive historical examination in which the history of the individual group is established, the classification of these languages on the basis of their common and differing features will remain provisional. In the comparison of cer

¹ Die Sprachfamilien und Sprachenkreise der Erde, Heidelberg, 1926, p. 140.
³ See n. 2 above.
⁴ See p. 140, n. 6.
individual correspondences in vocabulary, morphology, and phonology the
greatest caution is necessary, for all too often one finds completely independent
parallel developments that do not at all justify grouping the languages with such
common features together. Reasonable assurance for the correctness of the
provisional classification is offered only by a correspondence of several major
factors, in which connection morphological features deserve particular attention.

5. Features particularly to be regarded in classification are the following:

5.1. The position of parts of the sentence, especially in the relationship of
the determiner to the thing determined. From the structural viewpoint three
groups can here be clearly distinguished: (a) the eastern languages (with the
exception of Nicobar), in which the determinans consistently follows the
determinandum; (b) the Nicobar group, in which this rule is only partially
observed, in so far as the 'adjective', without an affix, precedes the 'substantive',
e.g. Nancowry ăă yuŋ lopaa kalĩŋ 'the two good foreigners' (literally 'two
persons good foreigners') or ăă yuŋ kalĩŋ tə-loopaa; in the latter case the adjective
following the noun has the prefix tə-. In Car Nicobarese an adjective with the
prefix tō- can also precede, and the practical result of this is that the whole
determinans precedes the determinandum, e.g. tō-tamii coonŋ 'steamer' (literally
'with fire provided boat'; tamii-yō 'fire'). Lastly (c) the Munda languages,
in which the determinans usually precedes, but in which the reverse holds true
in a number of cases. Thus in compounds, especially in Sora, the 'genitive'
often follows, e.g. ə-jeelu-boonŋ-an 'buffalo meat' (literally 'meat-buffalo-the').
The postposition of the personal pronoun as a possessive suffix is also frequently
employed, e.g. Juang bui-nɔm 'thy mother' (literally 'mother-of-thou').

Historical linguistic investigation shows the following: at the earliest period
for which we can make any statements the order determinans—determinandum
was predominant. Later began a period in which the determinans was in all
Austroasiatic languages largely but not exclusively placed after the determinandum.
The eastern languages with the exception of Nicobarese continued this tendency,
so that there now exist only a few remnants of old 'preposition', such as the
'preposition' of the numerals and numeral classifiers in Khasi and Bahnar.
The Munda languages returned to the customary 'preposition' of the deter-
minans—presumably under the influence of Dravidian and Aryan languages.
Nicobarese took a similar course—perhaps influenced by Andamanese languages.
The only difference from Munda is that in Nicobarese the number of formations
with 'preposition' of the determinans is comparatively small. There is no reason
to assume that Munda and Nicobar in this case passed through a common stage
of development; however, a common tendency may have been present.

5.2. The use of prefixes and suffixes is very closely connected with the position
of the determinans. The classification given above is also applicable here, but

\[1\] III, 4, 1960, pp. 97 ff.
Nicobarese and Munda are in this regard even more closely connected. The three groups are: (a) the eastern languages (with the exception of Nicobarese) which use prefixes almost exclusively. Remnants of old suffixes appear in Lao -do in kon-do ‘child’ = Kharia kon-du? (cf. Mon kon); Khmer lō-k ‘raise’, lō-ŋ ‘rise’, lō-s ‘more’, from lō ‘up’; cf. Kharia tob-lun ‘up’.

(b) The Nicobar group, which uses a large number of prefixes and suffixes. It is a striking fact that some affixes can sometimes be used both as prefix and suffix: e.g. Car ma, maa, affix indicating a nomen agentis, in ma-haŋh ‘sleeping’ (haŋh ‘to go to sleep’), but ma-maa ‘an eater of food’ (ma ‘to eat’). Further, Car ha, Nancowry hə (ha), affix indicating transitive or causitive voice (= Khmer pə-, Bahnar pə-, Sora ab-); e.g. Car ha-kwən ‘to cause to carry’ (kwən ‘to carry’), Nancowry fət (fat) ‘to break by striking’, fət-hə (fat-ha) ‘to cause to break’.

(c) The Munda languages, which are in this respect like the Nicobar languages. Both prefixes and suffixes occur both as productive and unproductive affixes. Among the productive suffixes are almost all the inflectional suffixes, e.g. Sanyam Mundari, Kharia -a? indicating the genitive. Unproductive suffixes occur, e.g. example in Kurku ka-ku ‘fish’ (Sre ka), Kharia sin’-ko-e < *sim’-ko-e ‘for’ (Mundari sim id., Sre sim ‘bird’). The suffix -ku, -ko occurs as a prefix in Wani k-sem, Mon ga-cem ‘bird’. Many of the suffixes in the Nicobar languages, particularly the directional affixes, are not connected with Munda affixes; others, however, a close connection can easily be supposed, e.g. Santali, Mundari -o?, verbal suffix of the middle or passive voice, Car -ə, Nancowry -ə (a). Car mūk ‘to see’, passive mūk-ə, Nancowry hərək-hətə ‘to burn’, past hərək-ə; further Mundari -ian, Santali -en, Proto-Munda *-ian, suffix of durative (with intransitive verbs), Nancowry -yen-de (-yen-de), suffix of the continuative, in top ‘to drink’, t-enm-op-yen-de ‘to continue drinking’. In some cases Munda suffixes occur as prefixes or as particles in Nicobarese, e.g. Juang - Kharia -tai, -a-tai, ablative suffix, Nancowry təi, in Kharia am-a-tai = Nancowry təi me ‘from thee’ (Kharia am ‘thou’ < *a-me; cf. Santali me ‘thou’). Further Mundari, Santali -le-d, -le-n, suffix for the non-resultative perfect, Sora perfect suffix, Nancowry leet; e.g. Mundari abu-native ‘I have washed’, Nancowry cū lae leet orii ‘I have beaten’ (from orii ‘to beat’; cf. Juang rim ‘hit’). Corresponding prefixes in Munda and Nicobarese are to be found, for example, in Mundari ta-, Nancowry to- (ta-), prefix indicating possessive pronoun, Mundari ta-ino, Nancowry to-ciò ‘my’, e.g. Mundari apu-ta-ino ‘my father’, Nancowry diu to-ciò ‘my canoe’.

It may be regarded as certain that a large number of both prefixes and suffixes were employed in the Austroasiatic languages at an earlier period. The fact

---

once again neither Munda nor Nicobarese lost their suffix-formations but rather extended the use of such formations through the erection of various new suffixes is evidence not for an actual common development but rather for a common conservative tendency.

5.3. A treatment of the numerous infixes, most of which are undoubtedly old, would be superfluous here, for they are encountered in all the groups and are hence meaningless for the classification.¹

5.4. The repetition of words that logically need to occur only once in a sentence is to be found both in Munda and in Khasi, e.g. Mundari am gapa-m sen-re-do, ka-m tebai-a 'if thou goest tomorrow, thou wilt not arrive' (literally 'thou tomorrow-thou go-if, not-thou arrive-will'), Khasi ka-wei ka konthei ka la iap 'the one woman died' (literally 'she-one she woman she died'). We have here to do with a kind of agreement which of course has not reached the extent to be observed, for example, in some African classificatory languages. Whether this agreement is old or whether it may earlier have been present in other Austroasiatic languages cannot yet be determined. A direct relationship of the two languages in respect of these forms cannot be shown, but as a common feature this point is not uninteresting. A classification from a structural point of view is then: (a) the eastern languages excepting Khasi; (b) Khasi and Munda.

5.5. On the basis of their syntactical framework the Austroasiatic languages may be classified into two groups: (a) the largely co-ordinating and analytic Khmer-Nicobar languages, and (b) the largely subordinating and synthetic Munda languages. The Nicobar languages here share the typical characteristics of the eastern group; subordinating participial constructions, which frequently occur in Munda, are here completely absent. According to recent investigation² the extensive subordination and the extremely synthetic structure of the Munda languages are of a secondary nature, just as the analytic structure and general co-ordination in the east are not original but are probably the result of secondary development and, in part, of the influence of the Thai and other languages. The complicated Munda conjugation, if pursued back to its origins in older Proto-Munda, shrinks to a fraction of its present size; thus in Proto-Munda no pronouns as indicators of either subject or object were incorporated into the verb. The decisive question here is whether the analytic simplification developed independently in the individual eastern groups or whether it began in a period before the formation of these groups, as was the case with the increased structuralization in the later period of Proto-Munda. If the latter supposition is correct, the Khmer-Nicobar languages form just such an historical unified group as the Munda languages do. It is unfortunately not yet possible to give a reliable answer to this question.

5.6. Morphological systems with declension and conjugation are present only

² See p. 141, n. 3.
in Munda. In the eastern languages with the exception of Nicobarese morphological systems are practically wholly absent. Special syntactical relationships and loosely attached particles there assume the function of the Munda declension and conjugation. Nicobarese again has an intermediate position: it possesses obvious remnants of an older morphological system. In Car Nicobarese there are three cases of the pronouns are distinguishable, 'absolutive' or 'nominative', 'interrogative', and 'oblique'. The first person singular is thus cin or cuw-ö in the nominative, oic, cu-ö oic in the interrogative, and cu in the oblique. As in Munda, the singular, dual, and plural of the pronoun are differentiated, as well as inclusive and exclusive forms of the first-person pronouns in dual and plural. Some of the comparative forms in Nancowry differ greatly from the corresponding morphemes in the positive, e.g. wiilih 'young', olyaalə 'younger', keoo 'hot', kenooy-wa-yan 'hotter'. Continuative forms of the verbs are created by various alternations of the root with the infix -en-, the suffix -a, and the suffix -yan-de; e.g. ikaafə 'to sing', kennoifs-wan-de 'to continue singing'.

Passive formations in -ö or in -a have already been mentioned, as the continuative forms have been; they agree exactly with the Munda forms. Very different from the Munda, on the other hand, is the use of directional affixes; e.g. Nancowry - which indicates motion downward, in ten-fat-ka-fe 'to break something throwing it to the ground'. It is evident from Munda and Nicobar that Proto-Austroasiatic possessed a fairly well developed morphological system that was further expanded in Munda, reduced in Nicobar, and lost in the other languages. This loss of the morphological system suggests not only a common development but even that this development began in a period in which the individual groups had not become independent. In general the following rule holds true for the development of languages: if only one of a group of languages introduces some important innovation, the fact that the other groups do not possess this innovation may not be regarded as a sign that they still have any common ties; however, if all but one of several groups of languages introduce an innovation, this fact may be regarded as an indication that the innovating languages possess common bonds. An early separation of Nicobarese from the rest of the groups seems very likely.

5.7. Another morphological feature is the frequent use of certain compositional forms: i.e., the words, when they are used in a compound, possess different form, which is usually one affix shorter than the normal form. This occurs mostly in Sora but to a lesser degree also in the other languages, especially Kharia; e.g. Sora tan-liy-on 'cow', compositional form taan-on, in guptar tan-mar-on 'cowherd' (literally 'graze-cow-man-the'), Kharia soreen 'stone', compositional form -sor in thom-sor 'to stone'. Here there is an agreement with Khasi, which likewise has compositional forms; e.g. ksa 'ring', kti 'hand', sah-ti 'ring for the finger'. In the other languages there do not seem to be any

---

1 See 5.2 above.
special compositional forms, though an old remnant may perhaps be present in
Khmer mo- (m-) to muov 'one', e.g. in m-'dan' 'once'. In Nancowry Nicobarese
fiŋkaa and daam (raam) are respectively used in combination with numeral
classifiers for hey 'day' and hōtəm 'night'. Whether old compositional forms are
present here is not entirely certain. However, such pairs as ha-kook 'to fire a gun',
hen-kook 'cannon', ha-teeₚₚ 'to sail', hen-teevo 'sail' reveal good comparable
parallels, as do the frequent phonetic changes in Nicobarese when affixes are
employed.

Thus from a structural standpoint the classification is: (a) the eastern
languages except Nicobar and Khasi; (b) Khasi, Nicobar; (c) Munda. Special
compositional forms probably existed at an earlier period; the fact that they
have been in varying degree preserved in Munda, Khasi, and Nicobar constitutes
a bond between these languages but is no evidence for a common development.

5.8. In the use of numerous numerical coefficients the Nicobar languages
agree completely with the other eastern languages and thus differ from Munda,
in which such words are little used. The word-order in Munda, Nicobar, and
Khasi is numeral, numeral classifier, noun; in Mon and Khmer noun, numeral,
numeral classifier; in Bahnar usage varies: e.g. Kharia moloi boko'b ore'j-ki
'five bullocks' (literally 'five head[s of] cattle'), Car sōm ta-ka nitō 'ten children'
(literally 'ten persons child'), Khmer kun prus pi 'nok 'three boys' (literally
'child masculine three persons'), Bahnar pluŋ 'baar tooŋ 'two barks' (literally
'bark two piece') or 'baar tooŋ pluŋ. It is remarkable that with regard to the
position of the numeral coefficients Munda, Nicobar, and Khasi again stand in
opposition to Mon and Khmer. That the use of numeral classifiers was extended
in the east and restricted in the west is probably the result of the influence of
neighbouring non-Austroasiatic languages.

5.9. With regard to phonology the following points are important:
(a) Central, unrounded back, or rounded front vowels (i, o, u, y, y, ø) :
prevalent in Khmer-Nicobar, rare in Munda. For Proto-Munda *i and *o must,
however, be reconstructed.

(b) Quantitative opposition of vowels: prominent in Khmer-Nicobar
excepting Khasi; less prominent in Munda. Quantitative opposition is to be
assumed for Proto-Munda.

(c) Contrast voiceless versus voiced stops: lacking only in Nicobarese. In
Nancowry d appears as a kind of liquid. The loss of voiced stops in Mon, Khmer,
Riäng, etc. is demonstrably secondary.

(d) Aspiration: lacking only in Nicobarese in the east and in Sora and
Gutob in the west. For the older period aspiration is to be assumed. It was,
however, clearly diphonemic, just as it is to-day in Khmer, etc. In Munda second-
arily monophonemic as a result of the loss of the other initial consonant clusters.

(e) Retroflex consonants: only in Munda excepting Sora. Retroflex con-
sonants are completely lacking in the east; t in Khmer and d in Mon are graphemic
equivalents of phonetically preglottalized [d'], which may be alveolar or dental. In Car Nicobar \( t \) (i) is no genuine retroflex. Proto-Austroasiatic did not have retroflex consonants; \( t \) was probably dental, \( d \)-alveolar. This is still the state of affairs in Sora and Palaung.\(^1\) Retroflexes were probably developed in Proto-Munda under the influence of Dravidian and/or Aryan languages.

(f) Non-plosive, glottalized consonants: in Munda and Nicobarese not initially but only finally. Initial implosive consonants are present in Mon, Khmer, Bahnar, etc., but not in Palaung and some other languages. All these sounds which have in part become phonemes, are of a secondary nature and have developed independently. Only the tendency to introduce such sounds is a common feature. The close correspondence of Munda and Nicobarese is again noticeable.

(g) Tonemes were originally completely absent. In the languages in which they now occur, e.g. Mon, Sre, Riang, they are clearly secondary developments.

(h) Consonant clusters are numerous in the eastern languages except Nicobar, rare in Munda, absent in Nicobarese. Clusters of two, and not more than two consonants undoubtedly occurred in Proto-Austroasiatic.

(i) Number of syllables in words: Munda inclines towards polysyllabism, the eastern languages to monosyllables. The polysyllabic structure of Nicobarese is the result of the fact that many affixes have there been indissolubly combined with roots. It is certain that the older languages had both one- and two-syllable words and that the latter have mostly been secondarily shortened in the east.

This phonological evidence again leads to a threefold classification: Munda, Nicobar, and the other Austroasiatic languages. Nicobar and Munda have three points of correspondence; Nicobar and Mon-Khmer, four points. The development of the phonemic system of Proto-Austroasiatic to Munda, Nicobarm and the remaining languages indicates that the former two groups were each separated from the family.

5.10. The numerals present a different situation. In this respect the Munda languages together with Mon-Khmer and one part of the Malaccan languages form one group in contrast to Khasi, Nicobar in another group and to the remaining Malaccan languages in yet a third group: e.g. Mundari bar 'two', Bahnar 'ba', South Sakai 'mbaar'; but Khasi ar, Palaung aa, Riang kaar, Nancowry a, North Sakai nar. Similarly Mundari pe, Mon pi, South Sakai 'mpe?' three, in contrast to Khasi lai, Palaung uve, Riang kwai, Nancowry huee; North Sakai ne?\. An exception is the number for 'ten'; here Munda agrees with Palaung-Wa and not with Mon-Khmer: Mundari gel, Kharia ghol, Palaung go, köör, Riang kol, but Nancowry jom, Car sam, Khasi jipeu, and Mon d. Bahnar jit. An historical consideration of the numerals shows that the phonological similarity of the numerals for 'two', etc. in Munda and Mon-Khmer is not at all the necessary result of a close connection between the two groups.

it consists rather in a common preservation of old forms, which have been changed in Khasi and Nicobar. The initial *- of the Munda and Mon-Khmer numeral for ‘two’ is absent, for example, in Khasi, Palaung, and Nicobar. Riang (belonging to the Palaung-Wa group) has initial *k-, Wa initial *l- or *r-, North and Central Sakai initial *n-. But it is false to try to explain this state of affairs, as Sten Konow has, by positing an old prefix *b-. The original initial sound, which was part of the root, was itself *b-. Ple-Temer, a Malacca language, gives us a clue: it has three variants for the numeral ‘two’: *nar, *panar, and *børnar. Now, *n- is an infix, which also occurs in Santali: *ba-na-r ‘both’ from *bar ‘two’. *banar easily changes to *panar and thence to *pnar and finally to *nar. *k-, *r-, and *l(a)- are undoubtedly prefixes that have replaced the old initial consonant: *k-bar became *kaar, etc., *l-pe and *l-pai became *l-we, *lue, and *lai. The absence of initial *b- in *aar is probably to be explained by positing a prefix *a-, so that *a-bar became *a-war and thence *aar. In the Munda languages we have a similar development in the numeral for ‘four’: Santali *pon, Mundari *u-pon, Gutob *vun, but Sora *un-ji. *Un- (ji is a plural suffix) surely developed out of *u-phun (cf. Kharia *p phon) from *u-pon; this is a purely Munda affair, cf. Mon *pon, Palaung, Khmer *pon, Riang *kh pwon, Nanceowry *poon, Car *f n. Thus if both Mon-Khmer and Munda have an old initial *b- in this case, this constitutes no evidence for any particularly close degree of relationship; the two groups have not added a *b-prefix at any common stage of development; they have simply retained an old feature of the language, and such a retention of course presupposes no interdependency. The loss of *b- in Palaung-Wa, Khasi, and Nicobarese may indeed point to an earlier common stage in these languages, but this is not a necessary conclusion, especially inasmuch as Riang in the Palaung-Wa group has initial *k-.

5.11. In regard to vocabulary the Munda languages have, according to W. Schmidt and Sten Konow, far greater similarities with the Mon-Khmer languages than for instance with Khasi. The important words for ‘eye’, ‘leg’, ‘blood’, ‘fly’, ‘hair’, ‘water’, ‘nose’, etc. agree with Mon-Khmer, but not with Khasi; and Schmidt’s classification, which was based upon the numerals, thus receives support. There are, however, some instances in which the Munda languages lie closer to Palaung-Wa than to Mon-Khmer, e.g. in the words for ‘sun’, Mundari *si ji, Palaung *sh o jii, but Khmer *th i jai. Khasi agrees here with Palaung and with Mundari with *si j, as does Nanceowry with *h en < *si j. If one or two groups employ variant words, this need not reflect an old state of affairs. For instance, the Car Nicobarese word for ‘sun’ is *ta-vuuõi. In the words for ‘water’—Santali, Mundari *da?, Kharia *da?, Mon *da?, Bahnar *da ak, Nanceowry *da ak, Car *mak (< *um-dak), but Palaung *oom, Khasi *um—it is certain

that the old usual word was *daak. The Khasi and Palaung words belong with the Kherwari um, Bnahar, Besisi hum ‘to bathe’. Since Lawa in the Palaung-Wa group also has haum in the meaning ‘to bathe’, it is clear that a change of meaning with a secondary suppression of the old word for water (perhaps for reasons of tabu) has taken place in Khasi and Palaung. The close relationship of Munda and Mon-Khmer is thus by no means so apparent as Schmidt and Konow assume. It is nevertheless evident that in regard to vocabulary there are fewer correspondences between Munda and Nicobar than between Munda and Mon-Khmer and Munda and Palaung-Wa. Their relationships can be cleared up only through more investigation, which unfortunately still remains to be done.

6. The foregoing investigation has led to the following tentative conclusions:

6.1. Schmidt’s thesis that the Munda and Mon-Khmer languages are to be grouped together and apart from Nicobar, Khasi, and Palaung-Wa cannot be maintained, since it did not take into consideration the most striking and important common features of the Khmer-Nicobar languages, and since it did not make allowances for the similarity of Munda to Nicobarese and, to a lesser degree, to Khasi.

6.2. The differences of Nicobarese from the other eastern Austroasiatic languages are so great that it should be ranked as a separate group within the Khmer-Nicobar languages, a group that on the one hand has close relationships with Munda and on the other with Khasi. Though most of the similarities shared by Munda and Nicobarese result less from close relationship than from a conservative tendency, an old contact relationship seems likely.

6.3. Within the Austroasiatic family the Munda languages form a clearly defined group that stands in opposition to all the eastern languages, i.e. the Khmer-Nicobar group; and the Munda languages undoubtedly are more similar to Proto-Austroasiatic than the other members of the family. From a morphological viewpoint they are far more conservative than Nicobarese and Khasi, from the standpoint of vocabulary they surpass the Mon-Khmer languages in their preservation of ancient word stems and word forms.

6.4. Whether the Khmer-Nicobar languages form a unified group, as the Munda languages do, remains uncertain. Such a unity is in any case not as evident as it is in the Munda languages, and it may depend only upon secondary, structural resemblances. There is as yet a complete dearth of any investigations which might guide us in tracing the individual members of the Khmer-Nicobar family back to a Proto-Khmer-Nicobar which might then be compared with Proto-Munda. It may be that there never was a Proto-Khmer-Nicobar and that the individual members of the Austroasiatic family—Munda, Nicobar, Khasi, Palaung-Wa, Mon-Khmer, Sakai, Jakud, Semang—constitute, historically speaking, sub-families independent of one another and traceable solely to Proto-Austroasiatic. In this case the Khmer-Nicobar grouping would have only structural and geographical justification.
6.5. The former geographical distribution of the various sub-groups with regard to one another may be inferred from the above-mentioned linguistic similarities. The lines of relationship Mon-Khmer/Palaung-Wa/Khasi, Munda/Nicobarese, Munda/Mon-Khmer, Nicobarese/Khasi are particularly important here. These areas of contact suggest the following older spatial distribution:

```
Munda
<table>
<thead>
<tr>
<th>Palaung-Wa</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>
Mon-Khmer       |
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Nicobar</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Khasi</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Malacca</td>
</tr>
</tbody>
</table>
```

(three groups)

It is interesting to note that one must assign a northern position to Nicobarese if one wants to avoid excessive spatial dislocations. The spread of the various groups out of this compact original homeland, the location of which cannot be discussed here,\(^1\) can be imagined to have taken place as follows. The Munda peoples were the first to leave; they migrated westwards. The Proto-Nicobar tribes went to the south, then possibly settled west of the Mon-Khmer on the coast, and later migrated to the islands; they were the second group to leave. The Khasis, who perhaps went west like the Mundas, constituted the third group. At a later period Austroasiatic-speaking peoples occupied increasingly widespread territories and imparted their languages to other peoples, e.g. to the originally non-Austroasiatic Semang on Malacca.\(^2\) Concerning the Malacca languages, which cannot be dealt with here in more detail, cf. C. O. Blagden in Skeat and Blagden, *Pagan races of the Malay Peninsula*, London, 1906, Vol. II, 432 ff.

7. Nahali, a language spoken in Central India and bordering on the Kurku area, has not yet been considered. Its classification is particularly difficult, for the language is not yet sufficiently well known or investigated. It is possible that Nahali is completely separate, as R. Shafer\(^3\) assumes, but it may also constitute a separate branch of Austroasiatic. It is at any rate not Munda. Nahali exhibits a number of words that cannot be explained as Austroasiatic, as Dravidian, or as Indo-Aryan. Its morphological system, on the other hand, is obviously connected with that of the Munda languages: thus, for example, all the Nahali tense suffixes may be compared with corresponding suffixes in Munda. The reliability of these comparisons is, however, reduced through frequent divergences in meaning. The present state of investigation does not permit any definite judgment. We may

---


perhaps come closest to the truth if we assume that Nahali possesses an isolated non-Austroasiatic substratum that has been partially replaced by an Austroasiatic stratum which has also provided Nahali with its inflection. But the resolution of this problem awaits further research.\(^1\)

8. A provisional classification corresponding to the present state of knowledge takes the following shape:

Western group (Nahali-Munda)

(A) West : Nahali (?)
(B) East : Munda

(a) North
Kherwari (Santali, Mundari, Korwa, etc.)
Kurku

(b) South
1. Central : Kharia, Juang
2. South-East : Sora, Pareng, Gutob, Remo

Eastern group (Khmer-Nicobar)

(A) West : Nicobarese (Nancowry, Car, etc.)
(B) East : Palaung-Khmer

(a) West : Khasi
(b) North : Palaung-Wa (Palaung, Wa, Riang, Lawa, etc.)
(c) East : Mon-Khmer (Mon, Khmer, Bahnar, Sre, etc.)
(d) South : Malacca

1. Sakai
2. Jakud
3. Semang

The position of Yumbri has not been considered here. According to Salzner, it constitutes a special Austroasiatic group; it may, however, belong to the Palaung-Wa languages. And, finally, the picture is not complete without the note of Austroasiatic languages with an Austroasiatic substratum, namely (1) the Champa languages and (2) Vietnamese and Müöng.

\(^2\) *Sprachenatlas des Indopazifischen Raumes,* Wiesbaden, 1960, pp. 4, 6.