

# AUSTRIC: AN 'EXTINCT' PROTO-LANGUAGE

Paul K. Benedict

A leading Mon-Khmerist (Diffloth 1985) has recently raised the question: what happened to Austric? One can dilly-dally over this at great length but the answer is very brief: it became extinct.

The Austric hypothesis was developed shortly after the turn of the century by Wilhelm Schmidt (1906), a pioneer in the field of comparative Mon-Khmer linguistics. He had noted a general similarity in morphology, with even some resemblances in prefixes and infixes, between Austronesian (AN) and Mon-Khmer (MK), the latter grouped with the Munda languages of India to make up the Austroasiatic (AA) stock. True, he had found nothing in MK to compare with the elaborate suffixial apparatus of AN but, not surprisingly, he skipped lightly over this discrepancy. And what of the anticipated common stock of 'core' (basic) lexical elements? Linguists of that period, and well up into the middle of this century, laboured under the mistaken notion that lexical elements are pretty much the *last* place to look in setting up linguistic relationships; furthermore, in those pre-Swadesh days, that even when one does get around to inspecting these elements he need only apply an aphorism that Gertrude Stein might have expounded: a word is a word is a word. Henri Maspero, perhaps linguistically the most sophisticated of all the French sinologists of the first half of the century, predictably followed along these lines of thought in wrenching Vietnamese out of its native MK setting and misidentifying it as kindred to the Tai languages. It would have been equally predictable, given the circumstances, that a linguist such as Schmidt would seek to establish a genetic linkage between AN and AA. Thus it can be said that the birth of 'Austric' was expected; rather less expected has been a certain continuing enchantment with it on the part of some scholars, a reluctance to accept its demise. A review of the matter, as attempted in this paper, is thus in order.

The writer, when not busy with actual (Sino-Tibetan [ST], Austro-Tai [AT]) rather than fantasied language stocks, has from time to time given some consideration to the basic question here. He at first (Benedict 1942) *en passant* expressed a willingness to go along with Schmidt's hypothesis on the assumption that continued research in that still largely unexplored field would turn up a respectable body of AN/MK cognate sets for 'core' lexical elements, thereby 'fleshing out' Austric, so to speak. Much later, after a sizeable amount of data in the field had become available, he read a paper on the subject at the First International Conference on Austroasiatic Linguistics at Honolulu (Benedict 1973). Here he played the 'devil's advocate' role, first setting up as good a case as possible for

Austric, tying in both phonological and lexical evidence, before finally—and with some reluctance—reaching the conclusion:

AT and AA [the earlier levels of AN and MK] do not have a core vocabulary in common, despite the morphological similarity of the two language stocks, hence the idea of an ‘Austric’ superstock must be abandoned.

The homelands of both stocks must be assigned to adjacent areas in South-East Asia (SEA) and the similarities in overall patterning can be ascribed to areal factors. This corner of the Asiatic mainland has, in fact, over the past decades acquired a certain fame among linguists for the vast scope of its areal influences, involving even highly specific processes such as vocalic transfer (Benedict 1979), and there is no reason to suppose that similar influences were not also operative at an early period.

In the paper presented in 1973 the writer went rather further in suggesting that a handful of lexical similarities in basic vocabulary between AT and AA are indicative of an early relationship of substratum type between the two stocks. It now seems evident that he overplayed his ‘devil’s advocate’ role in this respect, paying insufficient attention to the alternative explanation in terms of ‘look-alikes’ or ‘comparabilia’ (Matisoff 1976), along with areal influences or ‘border phenomena’. Thus, the classical pair: PAN \*mata (as earlier reconstructed); PMK \*mat ‘eye’ are attention-getting, to be sure, but the syllable reduction is basically unmotivated (*di-* as well as mono-syllabic roots occur in MK); PAN \*maCa (with \*C a cover symbol for reflexes representing proto-level palatal or consonant cluster), as now reconstructed by Austronesianists, weakens the comparison (final \*-c is a prominent feature of PMK) while PAT \*mapra (for the earlier cited \*mapla; see Benedict 1990) makes matters even worse: finally, PAA \*mət, the likely reconstruction on the basis of the Munda cognates (see Pinnow 1959), takes us even further away from PAT \*mapra at the early time level at which these proto-comparisons must be made. We can, in fact, do much better by comparing the PST root for ‘eye’: \*myək (as now reconstructed; see Benedict 1976a), with final \*-k > \*-t shift after the medial \*-y- (lacking in AA), even obviating the problem of syllabic reduction!

The above is an excellent illustration of the useful linguistic rule of thumb: ‘look-alikes’ look less and less alike as we attain more precise reconstructions at earlier and earlier levels. Thus, the suggested ‘twin/two’ set of the Benedict (1973) paper, labelled as ‘doubtful’ even by Shorto (1976), must now be discarded in the light of recent MK evidence re the medial vocalism<sup>1</sup> while the discussion there of the suggested set for ‘dog’ overlooked Monic final \*-r, which vitiates the comparison. Both comparisons are also burdened with the problem of unmotivated syllabic

---

1. The recently reported Lai language of Southern China (Guǎngxī) has /bi/ for ‘two’, supporting a PMK reconstruction of \*ʔbiar or \*ʔbiaar type, in line with Shorto’s suggested \*biʔaar > ʔbia(a)r rather than the \*ʔbaar cited in Benedict (1973). See Benedict forthcoming, n.6.

reduction, found also in 'eye' (above) as well as in a fourth set: PAN \*busuk; PMK \*sook 'hair'.<sup>2</sup> A fifth, often cited set: PAN \*ikan; PMK \*ka (or \*ka') 'fish' has the same syllabic discrepancy along with 'problems' in both final and initial (the Munda languages point to \*q- as the PAA initial, as reported in Benedict 1973). These difficulties led the writer to reject the comparison in 1973 and the present reconstruction and analysis of the PAN form (Benedict 1990) supports that decision: \*šikan (for the cited \*išikan), from \*ši-ka-n, an archaic derivative (> Japanese *ika* 'squid') of the ubiquitous AT 'core' root: \*ka/ka' 'eat'.

It is entirely possible, of course, that the borrowing of one or more basic roots such as 'hair' (above) will eventually be assignable to the AT/AA 'border phenomena'. If so, the syllabic reduction factor requires that one set up the direction of the loan(s) as AT > AA. This direction is also required for the early AT > AA culture-word loans, notable those for 'copper' (> Munda group 'iron') and 'sugarcane', as pointed out in the 1973 paper. The Aslian languages maintain an archaic final \*-s (generally > -h ~ ø in non-Formosan AN languages) in the latter loan and an extensive corpus of Aslian material of this kind recently collected by Geoffrey Benjamin (pers. comm.) includes what appear to be parallel examples of early loans that preserve other archaic phonological features, including even a final \*-l (maintained as a lateral in Formosan languages only). The analysis of this material is still in an early phase and it is unclear how many, if any, of these loans can be established as roots at the PMK (possibly even PAA) level; in any event, they supply additional evidence for the priority of the AT-speaking peoples in the early emergence of 'high culture' in South-East Asia (Benedict 1975).

At the 1976 Toronto symposium on AT, Harry Shorto made a spirited defence of Austric, prompting a 'Comment' by the writer (1976b). The argumentation there need not be repeated in this paper; much of it concerns lack of agreement re lexical items. Shorto proposed a number of new AN/MK cognate sets, falling for the most part in what the writer termed the *ketketbongbong* class, involving lexical areas well outside 'core' vocabulary, with a striking attenuation or gap in this key aspect. As summed up by Benedict (1976b: 106):

...the point to be stressed is that a score of *ketketbongbong* do not one language stock make, since correspondences of this kind have significant value for comparative purposes *only when found in association with a corpus of core vocabulary items*.

---

2. The Formosan evidence, strongly supported by the Japanese (see Benedict 1990), points to an underlying trisyllabic root: \*busukas or \*bosokas (PA medial \*o and \*u were merged in PAN \*u), which yielded Malayo-Polynesian forms of \*busuk > buhuk type. The mid or mid-low height of the PMK vowel—the cited form reflects the Proto-Semai vocalism—thus can readily be explained as a retention rather than an innovation while the length can be interpreted through vocalic transfer (Benedict 1979). This root appears to be unrepresented in the mainland AT languages (Kadai, Miao-Yao); the indicated line of development, with syllabic reduction on-the-left along with vocalic transfer, is of the sort that characterises the Kadai family as a whole. PAT \*o is also maintained in this family, which may well have been the source of an early loan (as an alternative to a 'look-alike') for this lexical item.

The final clause was italicised lest the point be overlooked, as apparently has occurred in the more recent paper by Diffloth (1985). The nub of the writer's argument here lies in the fact that Shorto, a MK specialist with good access to a wealth of comparative AN data, failed to turn up an even halfway decent corpus of 'core' vocabulary correspondences. This is indeed the name of the game, at least in the South-East Asia region, on the basis of an apparent linguistic consensus at the present time, and on this basis Austric surely emerged as the loser. And the writer (1976b: 105), even 'at the risk of appearing whimsical', could quite honestly report his belief that Shorto's paper had served only to strengthen the case *against* Austric!

Almost a decade passed before the appearance of another defence of Austric, this also by a leading Mon-Khmerist, Gérard Diffloth (1985). Unlike Shorto, however, who proposed a sizeable number of new AN/MK cognate sets, Diffloth included only two 'possible Austric etymologies': for 'wood' (~'tree') and 'bone', apparently the prize members of his collection, preferable to certain 'runts' which he modestly kept to himself. The pair that he does offer are in fact 'look-alikes' or 'comparabilia', with Diffloth admitting to serious difficulties in reconstruction even at the PMK, let alone a PAA or 'Austric', level. It must also be noted that he has unparalleled MK comparative material at his disposal along with the most recent advances in the AN field, all of which he has made use of, hence his failure to do any better than Shorto simply further strengthens the null case here. Given the present circumstances, the notion that a presentable corpus of 'core' AN/MK (or AT/AA) cognate sets lies 'out there somewhere' waiting to be discovered strikes one as quixotic in the extreme.

Diffloth's citation of 'look-alikes' in an undertaking of this nature conforms to general linguistic practice, as does his 'sharpening the reconstructions of promising cases' (Schmidt had proposed the 'wood' etymology). Quite novel, however, is his attack upon what he labels 'Swadesh's "basic" notions', which is no less than a frontal assault upon the whole idea of utilising basic or 'core' vocabulary in comparative work. Granted a certain feeling of frustration on his part in being unable to uncover 'core' AN/MK cognate sets, one can hardly follow him in this rash line of thinking. After pointing out that given items of 'core' type can on occasion be replaced through factors such as 'taboo or euphemistic *passe-passe*', he makes the following summary statement:

Over the millennia, the recurrence of factors like the above [*cit. supra*] could easily wipe out 100 such 'basic words', or even 200 [i.e. both the standard Swadesh-type word lists], while protecting for us many hidden gems, such as 'to whittle bamboo strips' or 'scruff', not to mention 'smegma' and the like. Benedict may detest it, but the Austric hypothesis is still very much alive. (Diffloth 1985)

The writer does not 'detest' Austric; in fact, as an old supporter of the hypothesis he must admit to a certain sentimental attachment to it and a

sadness at its demise. He must confess, however, that he is dismayed at the thought of a proto-language with a corpus of reconstructed roots made up exclusively of 'scruff', 'smegma', and the like. One can hardly visualise comparativists recognising a Proto-Indo-European language put together in such a fashion, and there is no reason for them to accept an 'Austric' of this sort either. South-East Asia has its novelties as a research field, to be sure, but it is hardly extra-linguistic. The cardinal point, as emphasised above, is that roots of this kind, whether called *ketketbongbong* or not (they need not be reduplicated), can be used in establishing a genetic relationship only in conjunction with 'core' vocabulary. On theoretical grounds, of course, all 100 or even 200 'Swadesh list' roots could have been replaced in a given instance, but how would one prove it, especially in the case of proto-languages, such as PAT and PAA, that occupy contiguous territories? Perhaps, in the dim South-East Asian past, AT and AA not only influenced each other structurally but also exchanged a few scruffy or smegmatic words in the bargain, along with some 'high culture' items after AT had made the early advance towards 'civilisation'. None of this, however, justifies the setting up of an 'Austric' superstock. As reported at the outset of this paper, Austric is 'extinct'.

#### REFERENCES

- Benedict, P. K. 1942. Thai, Kadai, and Indonesian: a new alignment in Southeastern Asia. *Amer. Anthropol.* **44**, 576-601.
- ..... 1973. Austro-Thai and Austroasiatic. First Internat. Conf. on Austroasiatic Linguistics, Honolulu (revised version appears as chap. II in Benedict 1975).
- ..... 1975. *Austro-Thai: Language and culture*. New Haven: Human Relations Area Files Press.
- ..... 1976a, Sino-Tibetan: another look. *J. Amer. Orient. Soc.* **96**(2), 167-97.
- ..... 1976b. Shorto—In defense of Austric.—Comment. *Comp. Analyses Asian Afr. Langs.* **6**, 105-8.
- ..... 1979. Vocalic transfer: a Southeast Asia areal feature. *Acta Orientalia* **40**, 229-52.
- ..... 1990. *Japanese/Austro-Tai*. Ann Arbor: Karoma Press.
- ..... (forthcoming). How to tell Lai: an exercise in classification (to appear in *Ling. Tibeto-Burman Area*).
- Difflloth, G. 1985. What happened to Austric? Paper presented at the 18th Sino-Tibetan Conf., Bangkok.
- Matisoff, J. A. 1976. Austro-Thai and Sino-Tibetan: an examination of body-part contact relationships. Papers for the 1st Japan-US Joint Seminar on East & Southeast Asian Linguistics, Tokyo.
- Pinnow, H.-J. 1959. *Versuch einer historischen Lautlehre der Kharia-Sprache*. Wiesbaden: Harrassowitz.
- Schmidt, W. 1906. Die Mon-Khmer-Völker, ein Bindeglied zwischen Völkern Zentral-Asiens und Austronesiens. *Archiv für Anthropologie* **5**, 59-109.
- Shorto, H. L. 1976. In defense of Austric. *Comp. Analyses Asian Afr. Langs.* **6**, 96-104.

