A Note on Lexical Replacement in Khmer

Philip N. JENNER
University of Hawaii
[retired]

In a recent paper George Grace (1988) has reopened the matter of glottochronology. Of the several assumptions of that theory he isolates three as fundamental. These are: (a) the existence in every language of an unchanging set of meanings realized as a "basic" vocabulary; (b) the randomness with which this vocabulary is replaced over time; and (c) the constant rate at which it is replaced. As his title indicates, Grace is concerned with the third of these assumptions, and for his persuasive argument the reader is referred to his paper. The second assumption appears to be no less open to question than the third: to postulate randomness of replacement when, *a priori*, we know little or nothing of circumstances governing such replacement could be seen as taking too much for granted.

Here, however, I propose to return briefly to the first assumption, that of a universal basic vocabulary, which has generated so much debate during the past thirty years. It is generally recognized that every language has a body of lexical items which appear to be more resistant to replacement than others; it is recognized further that such items are likely to include the numerals, the pronouns, terms for body parts, terms for conspicuous features of the environment, and terms for certain cultural features. It was on the basis of this consensus that the original 100 –item and later 200–item diagnostic lists were framed by Swadesh and his followers.[1] Implicit in these test lists as well as in their subsequent application to problems of linguistic subgrouping was, and is, the supposition that what constitutes basic vocabulary in one language constitutes basic vocabulary in every other language. This is an assumption, or presumption, which is as daring as that of random replacement and that of a constant rate of replacement. It is one which should surprise us when we remember that its proponents have been professionally involved in problems concerning human behavior and culture. The claim that the most stable lexical domain in a group of American Indian languages matches the most stable domain in a given language group of South Africa, for example, is one which cannot be demonstrated until *after* relationships within every language group have been worked out by the comparative method. It is certainly too doctrinaire for many anthropologists and linguists today.

Khmer is an example of a languge in which that part of the lexicon which corresponds to the glottochronological test lists has been susceptible to confusion as well as to change. To show the risks involved in using Khmer in glottochronological comparisons I have gone through an expanded version of the 200–item list which has been a standard tool with fieldworkers of the Summer

[1] For background on glottochronology see Grace's bibliography.

Institute of Linguistics for many years. The modern Khmer equivalents of the items in this list fall into a number of slightly overlapping groups yielding a total of 299 items. Illustrating the overlap of these groups is the item 'head': the usual word in modern Khmer is *kapāla* /kɓaal/, a loan from Sanskrit; but most Khmer–speakers are likely to know that an older word is *tpūṅa* /tɓòoŋ/, a regular derivative of *tūṅa* /ɗòoŋ/ 'coconut', which is also in the test list. Again, Khmer has at least two different forms corresponding to the list's 'to fear', 'to drop', 'mother', 'road', 'to kill', 'thin', and 'dry'; it has at least three different forms for 'to eat' and 'river'. In the following description of the groups into which my Khmer items fall I cite a few Pear equivalents to show relationships and contrasts, drawing my data from Headley's extensive wordlist (Headley 1978).[2] Pear is a Mon–Khmer language commonly agreed to stand closer to Khmer than most others.

From my Khmer list we can first cull out a group of 24 loanwords. For all its modest size (8.0%) in relation to the total, this group could be deemed large in the light of the strong claims that have been voiced for the stability of the basic vocabulary. Because of the far greater elaboration of Khmer culture than those of all other Mon–Khmer–speaking peoples except the Mon, it is reasonable to suppose that many of these loans represent replacements of original Khmer forms now lost. This is clearly true of *kpāla* /kɓaal/ 'head', just cited; it is equally true of *go* /kóo/ 'cow', likewise from Sanskrit, which has replaced Old Khmer *t(h)mur* /tmur/, itself from a wordbase meaning 'to gnaw, nibble'. The needs to which this class of loans responded can only be guessed at. Loans taken into Khmer to fill lexical vacuums may include the forms for 'sea', 'gold', 'silver', 'buffalo', 'anatomical heart', 'husband', 'wife', 'ten', 'hundred', 'green', and 'yellow'. If this possibility could be demonstrated it would furnish rather strong evidence against the claim made for a universal set of meanings expressed as basic vocabulary. But other loans in this group surely replaced original Khmer forms now lost: 'sky', 'rainbow', 'dust', 'animal', 'heart', 'calf (of leg)', 'to see', 'man', 'woman', and 'road'. But these 24 loans are of interest chiefly to the extent that they distance Khmer statistically from its closest congeners. There are only two cases in which Pear shows a form resembling the Khmer form: Khm. *māsa* /míiəh/ 'gold': P. *mas*, and Khm. *krapī* /krɑɓəə̀j/ 'buffalo': P. *krəpaaw*. The first of these is likely to be a loan directly from Khmer into Pear, though both could be loans from an Austronesian source into Mon–Khmer generally. The other 22 Khmer forms in this group exhibit no resemblance to Pear—for example, *peḥ ṭūṅa* /ɓèh ɗòoŋ/ 'heart' (cf. P. *(kələ)paʔm*); *srī* /srəə̀j/ 'woman' (cf. P. *c(h)əmkhi(i)n*); *ṭa'pa* /ɗɑp/ 'ten' (cf. P. *raay*). It is worth noting that in addition to the form for 'ten' this group includes another numeral, *raya* /rɔɔj/ 'hundred'. The only word for the latter which is attested in Old Khmer is *śata* /sat/, from Sanskrit (Jenner 1974).

Next, we have a group of 115 irreducible monosyllables—that is to say, forms having the shape CV(C) and incorporating no derivational morphemes. For 19 of these items Headley furnishes no Pear form. Of the Pear forms that his wordlist does include only 33 show any resemblance to their Khmer equivalents, while 63 show no resemblance. Examples of resemblances are Khm. *ṭī* /ɗəəj/ 'earth': P. *thee(y)*; Khm. *sa'ka* /sɑk/ 'head hair': P. *suk*; Khm. *mwya* /múuəj/ 'one': P. *mooy*. Examples of nonresemblances are Khm. *neḥ* /néh/ 'this': P. *aan*; Khm. *sa*

---

[2] Wherever Headley's forms show the IPA symbol for vowel length I use geminated vowels. A few other liberties have been taken with his transcription. Throughout this paper the abbreviation Khm. refers to Khmer, P. to Pear.

/saɑ/ 'white': P. *bruuy*; Khm. *pāya* /ɓaaj/ 'cooked rice': P. *claaŋ*. Of particular interest is the circumstance that these 115 Khmer monosyllables comprise only 38.5% of the test list, which should surprise us when we remember that we are concerned with basic vocabulary in the glottochronological sense.

A third group of Khmer forms consists of 138 items having the shape of derivatives by affixation: either CCV(C) or a presyllable followed by (C)CV(C). Over eighty years ago Pater Schmidt (1906:71) issued a stern warning, not often heeded in the intervening years, that no Mon–Khmer forms should be compared at all until divested of any affixes they might carry. If this admonition is justified, as I think it is, it follows that the items in this group, accounting for 46.2% of the test list, are not best suited to comparison; this conjecture does not bode well for any glottochronological results. However, when we come to examine the items in this group more closely, we find that they fall into two subgroups entailing a curious possibility. The first of these subgroups comprises 101 items of which the derivation is known (see Jenner and Pou 1980–1981). The other consists of 37 items of which the wordbases are not known from Khmer data. Examples of the first are *bhnaṃ* /pnum/ 'hill, mountain' (prefix /p–/ + *naṃ* /num/ 'small cake', < Old Khmer *naṃ* /nɔm/ 'mound; small cake'; *saṃpaka* /sɑmɓaak/ 'bark' (infix /–vN–/ + **spaka* /sɓaak/, prefix /s–/ + *paka* /ɓaak/ 'to peel'); and *sravi̯a* /srawɤ̀ŋ/ 'to be drunk' (prefix /srv–/ + *vi̯a* /wɤ́ŋ/ 'to spin, to be giddy'). Examples of the second subgroup are *phtau* /pɗaw/ 'rattan' (prefix /p–/ + *–tau* /–taw ~ –ɗaw/); *chkē* /ckaaɛ/ 'dog' (prefix /c–/ + *–kē* /–kaaɛ/); and *ṭaṅkūva* /ɗaŋkòow/ 'worm' (infix /–vn–/ + **thkūva* /tkòow/, prefix /t–/ + *–kūva* /–kòow/). A large number of wordbases attested only in Khmer derivatives can be assigned meanings, but the three bases just cited have not been identified. These two subgroups together pose the greatest and certainly the most exciting challenge to the comparativist. It would appear reasonable to take it as a working hypothesis that items in the first subgroup, precisely because they can be analyzed, were formed in Khmer with or without parallel formation in sister languages, while items in the second subgroup were formed in an ancestral language and inherited intact by Khmer and some of its sisters. At the same time, in this third group of 138 items, as in others, we must expect a certain amount of borrowing from one Mon–Khmer language into another; it does not necessarily hold that the direction of borrowing in this case would usually be from Khmer or Mon into the languages of the preliterate peoples. It is perhaps too early to draw any firm conclusions from such comparisons as Khm. *svā* /swaa/ 'ape, monkey': P. *wa* and Khm. *jhāma* /cʰíiəm/ 'blood': P. *haam* (Jenner and Pou 1980–1981:342, 390), where Pear shows the loss of prefixes.

— In the larger subgroup of 101 Khmer derivatives 15 items are in Headley's Pear wordlist. Of the remaining 86 items only 10 (11.6%) exhibit more or less close resemblances to their Pear counterparts, while the 67 others are answered in Pear by forms showing no resemblance. Examples of similarities are Khm. *brai* /prɔ́j/ 'forest': P. *brii, prii*; Khm. *phlē* /plaaɛ/ 'fruit': P. *phlii*; Khm. *chṅāya* /cŋaaj/ 'far': P. *cəŋay, sŋay*. Examples of dissimilarities are Khm. *babaka* /ppɔɔk/ 'cloud': P. *guul*; Khm. *ṭaṃrī* /ɗamrəəj/ 'elephant': P. *khənaay*; Khm. *dhmeña* /tméeɲ/ 'tooth': P. *khooy*.

— In the smaller subgroup of 37 items formed on unidentified (and presumably non–Khmer) wordbases 4 are not in Headley's Pear wordlist. But of the remaining 33, 13 or 39.4% show more or less close resemblances to Pear forms; only 20 show no such resemblances. The proportion of resemblances between Khmer and

Pear is therefore twice as great here as it is in the larger subgroup. Examples of similarities are Khm. *khya' la* /kjɑl/ 'wind': P. *cyal, khyal*; Khm. *thma* /tmɑɑ/ 'stone': P. *thəmoo(k)*; Khm. *bhlau* /plɔ́w/ 'thigh': P. *pluu*. Examples of dissimilarities are Khm. *phkā* /pkaa/ 'flower': P. *paaŋ*; Khm. *khlā* /klaa/ 'tiger': P. *rəwaay*; Khm. *khmau* /kmaw/ 'black': P. *ca(a)ŋ, thaŋ*. It is worth noting, in addition, that while Khm. *prāṃ* /pram/ 'five' corresponds to P. *p(h)ram*, for 'twenty' modern Khmer employs *mbhai* /mphɔ́j/ 'one score' while Pear employs *paa khsay* 'two tens'.

Quite apart from problems of linguistic comparison, these 101 Khmer derivatives and 37 pre–Khmer derivatives (as we may term them for the moment) raise a fundamental question: are they or are they not to be regarded as items of a basic vocabulary? For the time being I see no possibility of answering this question. On the one hand, it is tempting to see all derivatives as having replaced original underived monosyllables. Thus pre–Angkorian *tpūṅa* /tɓuuŋ/ 'head' (better: 'cocount, noggin'), cited above, *trī* /trə̀əj/ 'fish' (better: 'wiggler'), and *'aṅguya* /ʔaŋkuj/ 'to sit' (better: 'to perch'), and many formations of like kind, could be mere tropes which have displaced monosyllables now lost. These three Khmer derivatives correspond respectively to Pear *kəət* (or *tooh*), *mɛɛl* (or *meel, meʔel*), and *kil*. On the other hand, corresponding to Khm. *bhnēka* /pnɛɛk/ 'eye', a transparent derivative, Pear has *mat*, an underived monosyllable presumbly borrowed from an Austronesian source. Are we to suppose that neither Khmer nor Pear had an original word for 'eye'? Certainly not. In all of these cases we may ponder in vain the reasons why so many items of the common lexicon are derivatives, but this one case warrants the belief that reasons do exist. In the absence of evidence I can only conjecture that these derivatives found in the common vocabulary of Khmer do represent replacements of original monosyllables. I should look for the rationale of these replacements primarily in the extra–linguistic environment of Khmer in late prehistoric times, when Khmer culture could only have been in great turmoil owing to the crushing weight of new influences emanating from China, India and Indonesia. The widespread use of Sanskrit in the courts of the various principalities that sprang up and the entry of thousands of (royal, administrative, military, religious, botanical, and other) loanwords into a hitherto unwritten language must have led to a linguistic awakening in the course of which many ancient monosyllables fell by the wayside in response to an urge to find new ways of expressing old concepts. The powerful presence of Sanskrit alone would account for a new partiality for more elegant, more vivid, more precise, and often more amusing ways of stating the commonplace. In all of this Khmer would have been in good company.

The remaining items in my Khmer wordlist fall into three small groups:

(1) One of these comprises 18 items in the test list for which Khmer has developed phrases (headword + attribute). Pear shows no equivalents suggesting a relationship. Examples are Khm. *satva slāpa* /sat slaap/ 'bird': P. *chem, chiim*; Khm. *bīṅa bāṅa* /piiŋ píiəŋ/ 'spider': P. *tuŋ maaŋ*; Khm. *boḥ viana* /póh wíən/ 'intestines': P. *wiik*. These Khmer phrases raise the same historical question as the derivatives just mentioned.

(2) Another group consists of the 4 additive compounds for 'six', 'seven', 'eight' and 'nine'. In Khmer these are formed with the numeral 'five' followed by the numerals 'one', 'two', 'three', and 'four' respectively: *prāṃ mwya* /pram mú

uəj/, *prāṃ bīra* /pram piir/ > /pram pýl/, *prāṃ pī* /pram ɓəəj/, *prāṃ pwna* /pram ɓùuən/. Corresponding to these Pear has a set of unitary numerals cognate with those of a good many of its sister languages: *kədɔɔŋ* (or *krɔɔŋ*), *kənuul* (or *kənuuj*), *k(r)ətii*, *kənsaa(r)* (or *kəsaal*). It is well known that Khmer stands apart from the other Mon–Khmer languages in this respect, and I know of no one who does not accept the Khmer set as an innovation.

(3) The third and last group is a residue of two items in the test list. For 'thou' and 'ye' Khmer has recourse, when absolutely necessary, to pronouns lacking any implication of number but qualified by singularizing or pluralizing attributes. In their place Pear has *boo* (or *por*) 'you (*sg.*)' and *nɔɔk* (or *sak*) 'you (*pl.*)'.

To recapitulate: Headley's wordlist furnishes no Pear forms for 56 items in the test list; this reduces the comparisons which can be made from 299 to 243. In only 58 cases (23.9% of this adjusted total) can a cognate or loan relationship be discerned between Khmer and Pear. Answering to my 24 Khmer loans are only 17 Pear forms, these yielding only two resemblances. Answering to my 115 Khmer monosyllables are 96 Pear forms including 33 resemblances. Answering to my 101 Khmer derivatives are 86 Pear forms with only 10 resemblances. The highest percentage of resemblances (39.4%) is found in the set of 37 pre–Khmer derivatives, where 33 Pear forms yield 13 resemblances. In my last three minor groups we find no resemblances at all.

It can be said with some confidence, however, that the number of *prima facie* relationships between Khmer and Pear would have been far higher had the items in my Khmer list been less liable to replacement. For no one can examine Pear without a feeling that it stands much closer to Khmer than 23.9% would suggest. If comparison based on a glottochronological list of basic vocabulary leads to this result, Headley's Pear wordlist contains several hundred items which have presumable cognates in Khmer but are not part of any test list. In working out the relationship between Khmer and Pear are we to ignore all this additional evidence?

Just as the selection of meaning making up the glottochronological lists is probably more intuitional than objective, so is it unlikely that the lexical items expressing those meanings could have the same stability or permanence in all languages. Moreover, the limitation of these lists to an arbitrary number of meanings is dictated by the statistical use to which they are put; but if the lexical items expressing those limited meanings are not replaced at a constant rate the lists are pointless: any language relationships determined by means of such lists cannot have the validity which has been claimed for them.

Certainly the glottochronological test lists do not impress me as suited to Mon–Khmer. Indeed, the Khmer evidence points to the conclusion that glottochronology puts the cart before the horse: "basic" vocabulary is identified by collecting possible cognates out of the general lexicon and developing the regular correspondences between them. In most culture areas there should be little difficulty for the experienced investigator in focusing his search on the more conservative and even archaic domains of the lexicon. In Mon–Khmer these have to do not so much with the numerals and pronouns and body parts as with the traditional occupations and crafts. The cultivation of irrigated rice in the lowlands and of dry rice in highlands employs a body of terms, such as names for tools and other equipment, not likely to give way to loans or neologisms and therefore apt to reveal cognates. The

hunting, raising and use of the elephant (Pou 1986), the keeping of water buffalo, weaving and loom construction, stonecutting, boatbuilding, house construction, fishing, hunting, the working of bamboo, and metalworking are all dependent on special terms likely to be of considerable age and permanence. To fail to tap all such resources is tantamount to setting aside much needed evidence.

BIBLIOGRAPHY

Grace, George W. 1988. Do languages change at a constant rate? An alternative view of basic vocabulary replacement. *East Meets West: Homage to Edgar C. Knowlton, Jr.* Edited by Roger L. Hadlich and J.D. Ellsworth. Honolulu: University of Hawaii, College of Languages, Linguistics and Literature, Department of European Languages and Literature, pp. 72–9.

Headley, Robert K. 1978. An English–Pearic vocabulary. *MKS* 7: 61–94.

Jenner, Philip N. 1974. Les noms de nombre en khmer. *JA* CCLXII.1–2:171–91.

Jenner, Philip N. and Saveros Pou. 1980–1981. A lexicon of Khmer morphology. *MKS* 9–10.

Pou, Saveros. 1986. Vocabulaire khmer relatif aux éléphants. *JA* CCLXXIV.3–4: 311–402.

Schmidt, Wilhelm. 1906. *Die Mon–Khmer Völker, ein Bindeglied zwischen Völkern Zentralasiens und Austronesiens*. Braunschweig: Vieweg.

received: April 1991
revised: 30 May 1992

356 Delezene Road
Elma, WA 98541
USA