

Philology and archaeology

SIR GERARD CLAUSON

Sir Gerard Clauson was trained as an orientalist at Oxford and developed his wide interests in archaeology and oriental languages during a long career in the Colonial Office, from which he retired as Assistant Under-Secretary of State in 1951. He is a past President of the Royal Asiatic Society and has already written in these pages on the decipherment of the Indus script (1969, 200-2). He was stimulated to write us this article by reading Professor McNeal's 'The Greeks in history and prehistory' which we published last year.

Reprinted from
Antiquity

Volume

Year

Professor McNeal's paper 'The Greeks in history and prehistory' (*Antiquity*, 1972, 19 ff.) is a refreshing essay in scepticism, all the more welcome at a time when, as the Editor has recently pointed out in another context, credulity is dangerously in the ascendant. But scepticism can be carried too far, as the 18th-century divine pointed out to the Free Thinker, when he said (I may not have got his wording quite right) 'I suppose, Sir, that you do not believe in the existence of your backside because you have never seen it.' In what follows I hope to show that Professor McNeal is unduly sceptical regarding the archaeological significance of philology.

I make no apology for speaking of 'philology' where Professor McNeal speaks of 'linguistics'. In the epilogue, entitled 'An old-fashioned look at the linguists', to *Turkish and Mongolian Studies** which, to be frank, was irrelevant to my main themes, and irrelevant to linguists, I took exception to the word 'linguistics' on the ground that it was the base-born offspring of a Latin stem and a Greek suffix, and suggested that philology and linguistics were different branches of the same discipline, philology ('love of words') being the study of written languages, and linguistics (something to do with 'the tongue') the study of the spoken word.

First on the subject of nomenclature, as one whose ethnical designation is derived from the

name of the tribe which is generally reckoned to have been the junior partner in the Anglo-Saxon invasion of England, I do not see any grave objection to using the name of an obscure Thessalian tribe as an ethnical designation of much wider import, provided that it is made quite clear in what sense it is used. After all many ethnical and geographical names have equally haphazard origins. We still speak about China because the first Indo-Europeans who had a name for that country, a group of adventurous Sogdian merchants, became acquainted with it when it was ruled for a brief period in the 3rd century BC by the Ch'in dynasty. The early Turks for some centuries either side of about AD 800 called it 'Tavghach' because they first made contact with it when northern China was ruled by the Tavghach tribe (in Chinese transcription T'o-pa) as the Yüan Wei dynasty in the fourth to sixth centuries AD. The Russians still call it Kitay because they first became aware of its existence when it was ruled by the Kitay/Kitay tribe as the Liao dynasty in the 10th to 12th centuries AD. Humpty Dumpty put the matter in a nutshell when he said: 'When I use a word, it means what I intend it to mean, neither more nor less.'†

I quite agree with Professor McNeal's objection to the ambiguous use of the word 'Greeks' to designate a 'well defined entity'

† Lewis Carroll, *Alice through the looking glass* (People's Edition, 1889), 114.

* Royal Asiatic Society Prize Publications Fund, 1962, Vol. xx.

without its being made clear whether a racial, cultural or linguistic entity is meant. If by racial entity we mean a pure bred stock with no admixture of alien elements, we are speaking of something which is extremely rare. It is possible that there were recently, or are still, one or two pure-bred stocks living in particularly inhospitable or inaccessible parts of the world. The Maoris, perhaps, were before the coming of the white man, and the Eskimos may still be, but not for much longer. But obviously neither the 'Greeks' nor any other peoples in the Mediterranean area in, or later than, the 2nd millennium BC could be confidently described as a pure-bred stock, and so a racial entity. Culture is man-made and never remains static for very long. If a 'culture' is qualified by an ethnic, it must, to be meaningful, be further qualified in time and space. Professor McNeal's definition of 'Greek culture' (p. 21) is a good example. Even so such a culture is not necessarily homogeneous. Early Ottoman culture is usually regarded as homogeneous, but the mosque of Murad I at Bursa looks remarkably like a Venetian palazzo. The explanation is that it was built by an Italian architect whom the Sultan picked up in a miscellaneous bunch of Christian prisoners and put to work at his trade. If racial and unqualified cultural entities are ruled out, all that seems to be left is linguistic entity, that is, that by 'Greeks' we should mean simply the people who spoke Classical Greek, or the earlier cognate language of the Linear B tablets, or the modern dialects. Speaking as a philologist, but without specialized knowledge of Greek as thus defined, I suggest that any language which exists over a long period evolves in much the same way; compare for example the evolution from Vedic Sanskrit to the modern Aryan languages of India or from the language of the Avestan Gāthās to modern Persian. But this does not imply, and must not be allowed to imply, that over the period of nearly three and a half millennia during which the language is known to have existed, its speakers have been racially homogeneous, or that there has been a parallel steady and uninterrupted evolution of their culture.

As one unprejudiced by any previous

knowledge of the subject I fully accept Professor McNeal's diagnosis of the present sorry state of prehistoric Aegean studies. But I do challenge his negative attitude to the utility of correlating comparative philology with archaeology in order to reconstruct cultural history.

The origin of human speech and its relationship to the methods of intercommunication employed by our cousins the apes are subjects which the philologists, and even the linguists, can cheerfully leave to the anthropologists, within whose sphere they clearly lie. What does concern the philologists is that there is no reasonable doubt that at various remote dates in the Stone Age, when hunting/food-gathering bands were scattered over those parts of the world which were then habitable as sparsely as the plums in a grossly under-endowed plum pudding, various bands evolved languages, in the full sense of the term, in complete isolation from one another. There is ample archaeological evidence to show how widely these bands were distributed at this remote period, and the fundamental differences between, say, the languages of the Eskimos and the Bushmen show that they must have been evolved in complete isolation from one another. The only question is how many such independent languages were invented, and that we shall never know.

Some of these primitive languages have no doubt disappeared without leaving any trace. A few have survived as 'loners' without any known relatives, for example Burushaski spoken in Dardistan, and one or two Palaeoasiatic languages like Yukagir and Giliak (Nivkhi). Sumerian, which has long been extinct, has every appearance of being a 'loner', in spite of numerous attempts to foist relatives upon it, some grotesquely improbable.

In addition to these 'loners' there are a good many groups of languages which are clearly related to one another and can be described as families. Some, like the Yeniseyan group in Siberia (one living language, Kettish, and two which were partially recorded before they became extinct) are very small. Others, like the Finno-Ugrian, Semitic, Turkish, Mongolian and Tungus, are of moderate size. By far

and away the biggest is the Indo-European group.

Objection has been taken to the use of the word 'families' for languages, with the implication that they are descended from a common ancestor, on the ground that it is improper to use a biological model for languages, which are not organisms. I have some sympathy with this attitude because in a recent paper* I pointed out that the whole basis of glottochronology was false because it assumed that all languages behaved in exactly the same way as an organic substance, so that the proportion of every original vocabulary which remained in use declined at the same steady rate, defined as a 19 per cent reduction per millennium, in the same way that the proportion of C14 in an organic substance declines at a steady rate. But I suggest that these two cases are quite different. A language is an abstract thing created in the first instance, and then continuously modified, by human minds. The exclusive link between a language and the pure-bred hunting/food-gathering band which created it ceased to be exclusive when the descendants of the original band were subjected to racial admixture and the language was adopted, voluntarily or under pressure, by people who had originally spoken another language, but it is broadly true that a proportion, and probably quite a large proportion, of the people who now speak some modified form of the original language are lineal descendants of members of the band which originally evolved it. From generation to generation descendants of the original band have modified the language in various ways. It is surely rather perverse to object to the use by analogy of such terms as a family of languages, descended from a common ancestral language, for languages of which so many speakers are distant blood-relations by virtue of their common descent from the band which evolved the ancestral language.

Professor McNeal has a good deal to say against Schleicher's *Stammbaumtheorie*, but apparently for rather different reasons. Some

* A lexicostatistical appraisal of the Altaic theory, *Central Asiatic Journal*, 1969, XIII, 1-23; Russian translation in *Voprosy Yazykoznaniiya*, 1969, v, 22-41.

Indo-Europeanists, particularly in the Soviet Union, have become so discouraged by their endless struggles through the matted jungle of the Indo-European languages and their grammars, that they are inclined to say that they are descended not from a single ancestor but from some kind of a syndicate of ancestors. But those who like myself have devoted themselves to the study of smaller families, in my case Turkish, have no doubt that Schleicher, even if he started life as a botanist, was right regarding the families in which we specialize. It is hard to see why the Indo-European family should be the odd man out. But it is easy to see that, in view of the large number of Indo-European languages and the wide differences between them, their common ancestor, and probably its immediate descendants, must have become extinct so long ago that it is now practically impossible to deduce what exactly they were like.

This is the looking-glass image of the old British custom of setting up a financial trust and declaring that the capital shall be divided between the remaining beneficiaries not more than 21 years after either the death of the last beneficiary, or the death of the last descendant of Queen Victoria, who were alive when the trust was first set up, whichever be the later. After two world wars and other upheavals it became clear that it was practically impossible to ascertain who, of the descendants of Queen Victoria, were alive at any recent date and it is now the practice to substitute King George the Fifth for the Old Queen.

It is one of the tragedies of 19th-century comparative philology that because nearly all the great scholars were Indo-European by stock, and Classicists by education, they did not realize the importance of working out their techniques in a simpler field, but advanced bravely on the problem of comparative Indo-European philology, which is by far the most difficult problem of its kind. It is as if a man started out to climb Mount Everest without first making it his business to master the techniques of mountain climbing on easier gradients and at lower altitudes.

The reasons for the special difficulties of

Indo-European are obvious. Even if the African, American and Australian continents are excluded as newcomers, Indo-European languages are, or have been, spoken as native languages as far east as the frontiers of China, as far west as Portugal, nearly as far north as the arctic circle and nearly as far south as the equator. The earliest texts in one at least of them goes back beyond the middle of the 2nd millennium BC, and there has been an almost uninterrupted and steadily increasing volume of texts ever since.

While the Indo-Europeanists were wrestling with this intractable problem, the students of the Finno-Ugrian languages, mainly Finns and Hungarians, were quietly working away at their own languages and in fact proved that they constituted a family of languages with a common ancestor before the existence and extent (as it was then known) of the Indo-European family was finally proved.

The study of the Semitic languages, especially Hebrew, has a very long history. I do not know when exactly the existence of a Semitic family of languages was proved, but I do not think that any Semitist would now deny that they all go back to a common ancestor.

The most fascinating exercise in comparative philology is vocabulary analysis, that is dividing the subject matter of a language, preferably the oldest surviving member of the family, into three parts: (1) concrete things and abstract ideas (including adjectives and verbs) for which the language has its own native words; (2) things and ideas for which it uses loan-words, because its speakers were not aware of their existence until they established contact with some other people that was familiar with them and had its own words for them; (3) things and ideas for which it has no words at all.

An analysis of this kind can project the history of the language and its speakers much further back than the earliest written texts in it. For example the earliest surviving Mongolian text dates from the 13th century. The earliest Turkish loan-words in the language were almost certainly borrowed eight hundred, perhaps even a thousand, years earlier. It throws a flood of light on the cultural history of

its speakers and, in favourable circumstances, on the geographical surroundings in which the language took shape. Examples to prove these points follow shortly, but it is easy to find examples much nearer home. For example, until comparatively recently no European wore a two-piece sleeping suit, and there were no European words for it. Then Europeans who had lived in India brought back such garments with them and also a Hindustani loan-word, 'pyjamas'.

So far as geographical surroundings are concerned I may perhaps be permitted to take examples from the language families with which I am most familiar. The treatment is necessarily a very summary one: I hope to set out the evidence in much greater detail elsewhere. Early Turkish had native words for steppe animals like the steppe fox (*Canis corsac*) and hot-country insects like the scorpion and tarantula. It had a very elaborate terminology for a number of domesticated animals, but not pigs, and native words for some cereals; but it had Tokharian loan words for sesame and probably barley, a particular kind of plough, and a particular kind of bovine, *öküz*, cognate to English *ox*, which probably meant a castrated bull, the only kind suitable for drawing a plough. It had no words for northern animals or birds, large river fish, sea fish, sea creatures, sea birds or seaweed. This suggests that the Turkish homeland was in what are now the Mongolian steppes, a long way inland and south of the thick Siberian forests and the great Siberian rivers. It also gives some indications of the Turkish cultural debt to the Tokharians.

Perhaps I might be permitted to make a short digression at this point. It is one of the sorrows of those interested in Turkish and Chinese philology that so little of the Tokharian vocabulary has survived. For example, applying the normal rules of Tokharian phonetics to the Indo-European words for 'barley', the Tokharian word should have been something like *alpi* which is very like Turkish *arpa*. Again the Chinese word for 'camel', which is disyllabic and so a loan-word, might well be Tokharian, but the Tokharian word has not survived. Thus, although some Tokharian loan-words can

be firmly identified in both languages, there may be others of which the Tokharian original has disappeared.

Tokharian also throws some light on the nature and chronology of the dispersal of the Indo-European tribes. Indo-Europeanists have long been puzzled by the fact that Tokharian is a *centum* language with a phonetic structure similar in many ways to Greek, Latin and some other Western languages, but is separated from them by a solid block of (Indo-)Iranian *satəm* languages. Edwin Pulleyblank, a Professor of Chinese, has recently put forward what looks like a convincing explanation. He has found evidence which indicates that the ancestors of the Tokharians made contact with China at a very early date, possibly in the 3rd millennium BC. He suggests that they were the first Indo-European tribe to leave the main body in an eastward direction, that when they did so all the tribes spoke *centum* languages, and that the Indo-Iranian tribes moved east later and evolved the phonetic changes typified by the equation *centum/satəm* after they had made this move.

Unlike the Turks, the Mongols had a substantial number of native words for large river fish and for northern wild animals such as the reindeer, but they used Turkish loan-words for steppe animals such as the steppe fox and the wild goat, and for some cereals, and had no words at all for scorpion or tarantula. Of the domesticated animals they had native words for dog and fairly elaborate vocabularies for horses and for pigs (and a different series for wild pigs) but they borrowed the whole elaborate Turkish vocabulary for bovines, sheep, camels, donkeys and in the earliest period mules (later they used a Chinese loan-word). Like the Turks they had no sea vocabulary at all.

This suggests that the Mongolian homeland lay north of the steppes in the southern Siberian forests, where they could learn to domesticate the pig, and in the valleys of some Siberian rivers with their large fish, but well away from the sea. It also shows that they learnt from the Turks how to domesticate some animals and to grow cereals.

It is difficult to speak with the same precision

of the Tungus languages. Most of them are dialects spoken by small or very small tribes scattered thinly over Siberia east of the Yenisei as far north as the Arctic Ocean and as far east as the Sea of Okhotsk. Only the Manchus had a history of 300 years of literacy and, although there are extensive dictionaries of Manchu, they all date from a period at which they had been established for some time in China and acquired many Chinese habits. Subject to these reservations the position appears to be as follows. Unlike the Turks and Mongols they had an extensive sea vocabulary. They had native words for northern animals like the reindeer, which they domesticated, for large river fish and for the Manchurian tiger. Of the domesticated animals they had native words for dogs and an extensive vocabulary for pigs (and a different series of words for wild pigs). But for all other domesticated animals they used Mongolian loan-words, most of them Turkish loan-words in Mongolian. They had only loan-words for steppe animals like the steppe fox and no words at all for scorpion or tarantula.

This suggests that their homeland was in Siberia north of the Mongols and extended at any rate to the eastern seaboard.

Admittedly in all these cases vocabulary analysis by itself cannot date the cultural changes proved by the loan-words. Some Turkish loan-words in Mongolian were obviously borrowed, by reason of their phonetic structure, much earlier than others, but tentative dates can be put on them only by correlating them with the history of ancestors of the Mongols like the Kitān as recorded in the Chinese historical texts. But it should be possible to correlate this evidence with the evidence of physical archaeology which, at any rate in part of the area concerned, seems to be able to produce fairly firm dates for archaeological cultures, and so to produce a more complete prehistory of these peoples than would be possible if they were kept rigidly apart.

Admittedly, too, I have stumbled more or less by chance on about the most favourable examples of the use of vocabulary analysis which it would be possible to find. In these cases, for example, the confusion caused by

words changing their meanings in time and space, that bugbear of the Indo-Europeanists, is reduced to a minimum, and the difficulties caused by exchanges within the group, which are common, for example, in the Semitic languages, do not occur at all.

The Indo-Europeanists have devoted a considerable amount of time and energy to comparative grammar, which is of little or no significance in the Turkish, Mongolian and Tungus families, and their concern with vocabulary has been principally to trace the history of words which are unquestionably Indo-European, including borrowings and lendings within the family. There are of course many easily identifiable loan-words in medieval and modern languages, for example Turkish and Finno-Ugrian loan-words in Russian, but these add very little to the information about cultural history which can be obtained from other sources.

But probably from the 3rd millennium BC, or even earlier, Indo-European languages began to spread over wide areas in which other languages had previously been spoken. It is hard to believe that they did not acquire some loan-words from those languages; the difficulty is to identify them. It is fairly easy, for example, to identify the Babylonian loan-words in Hittite, but some of the other lending languages may have completely disappeared. Who knows, for example, what language the Tripolye people of the Ukraine spoke?

On the other hand, if a particular word is peculiar to a single language, or small group of

languages, there is a good chance that it is a loan-word. For example there is no acceptable Indo-European etymology for the Tokharian word *tmām* (A) / *tumane* (B) 'ten thousand', which was subsequently borrowed by the Turks as *tümen*. Professor Pulleyblank has suggested that it is a very old Chinese loan-word. 'Ten thousand' is now *wan*; in Middle Chinese it was something like *myan*. But there is reason to believe that at a much earlier date there were initial consonantal clusters, which have now completely disappeared in Chinese, and that *dm-*, which existed in Classical Tibetan, a language cognate to Chinese, though it has now disappeared in the modern language, was one of them. If so, 'ten thousand' in Proto-Chinese was probably something like *dman*. The conclusion seems to be that although the Tokharians lent some words to Chinese, including probably the word for 'camel', their herds had never increased to sizes at which they felt the need for a numeral as high as ten thousand until they had established contact with China, and that when they did they borrowed a Chinese word for it. It is possible that the Turks borrowed it from the Tokharians for the same reason, but perhaps more probable that it was borrowed when Buddhist scriptures, which have a proclivity for high numbers, were translated from Tokharian into Turkish.

And so I suggest that if this subject is looked at in a wider context than the Aegean, Professor McNeal's paper, brilliant though it is, goes too far in what he calls the negative direction, and that his prohibition is too categorical.

Is not your voice broken? . . . and every part about you blasted with Antiquity?

2 Hen. IV

Rejuvenate yourself with a book from Heffers:
the bookshop at 20 Trinity Street, Cambridge

heffers: