

THE SPREAD OF THE INDO-EUROPEANS

Frederik Kortlandt

The publication of Mallory's book (1989) has rendered much of what I had to say in the present contribution superfluous. The author presents a carefully argued and very well written account of a balanced view on almost every aspect of the problem. Against this background, I shall limit myself to a few points which have not received sufficient attention in the discussion.

First of all, the relation between archaeology and linguistics is a precarious and asymmetrical one (cf. already Schmitt 1974). Mallory's lucid discussion of the problem (1989: 164-168) should be required reading for anybody who ventures into this realm of shadows. It is a methodologically legitimate activity to look for archaeological traces of a linguistic group, but the converse does not hold. Speculations about the linguistic affinity of a prehistoric culture are futile because it is reasonable to assume that the vast majority of prehistoric linguistic groups have vanished without leaving a trace. Thus, it is certainly attractive to assign the ancestors of the speakers of Proto-Tocharian to the Afanasievo culture (cf. Mallory 1989: 62 and 225), but we must never forget that the very existence of the Tocharian texts which have survived is a purely accidental fact of history, due to a number of factors which happened to concur thousands of years after the eastward migrations of the Indo-Europeans. It is not merely possible, but very probable that many groups of Indo-Europeans migrated eastward before the ancestors of the Indo-Iranians, and that the distinguishing feature of the Tocharians is merely the preservation of their historical records. If the differences between East and West Tocharian lead us to date Proto-Tocharian to the second half of the first millennium BC, this still leaves a gap of two or three millennia after the purported arrival the Indo-Europeans in the area. Many things may have happened in the meantime.

The real argument for an early eastward migration of the ancestors of the Tocharians is the remarkably archaic character of the attested languages (see now Penney 1989 for a point of particular importance). It has often been argued that Tocharian has special connections with the western Indo-European languages. In my view, this is the result of a methodological bias in our way of reconstructing Proto-Indo-European. As Mayrhofer has noted (1983), the history of reconstruction can be described as a gradual shift away from the languages on which the reconstruction is primarily based. The similarities which link Tocharian to the western Indo-European languages reflect precious archaisms which were obscured by more recent developments affecting the dialectal area from which Greek and

Indo-Iranian were to evolve. The bias is strengthened by the presence of later parallel innovations in the latter two branches, e.g. in the development of the middle voice (cf. Kortlandt 1981: 130).

Similarly, Mallory's inconclusiveness about the westward Indo-European migrations (1989: 257) appears to result from a search for archaeological evidence beyond what can be motivated from a linguistic point of view. If we follow the traditional opinion and assign the ancestors of the speakers of Celtic and Germanic to the La Tène and Jastorf cultures, respectively, this again leaves us with a gap of two millennia after the Corded Ware horizon to which the ancestors of the western Indo-Europeans may have belonged. Here again, we can be sure that a lot of things happened in the meantime, and it is most probable that many linguistic groups were irretrievably lost.

This leads me to the second point I want to make. There seems to be a general tendency to date proto-languages farther back in time than is warranted by the linguistic evidence. When we reconstruct Proto-Romance, we arrive at a linguistic stage which is approximately two centuries later than the language of Caesar and Cicero (cf. Agard 1984: 47-60 for the phonological differences). When we start from the extralinguistic evidence and identify the origins of Romance with the beginnings of Rome, we arrive at the eighth century BC, which is almost a millennium too early. The point is that we must identify the formation of Romance with the imperfect learning of Latin by a large number of people during the expansion of the Roman empire. Similarly, we may identify the formative period of Proto-Indo-European with the earliest expansions of the Indo-Europeans.

The issue involved here is partly terminological. Elsewhere I have presented a relative chronology of 22 stages for the phonological developments which characterize the formation of Old Irish (1979). All of these developments are posterior to the Ogam inscriptions, which lack the characteristic features of the Old Irish language. If we use the term "Primitive Irish" for the period before the apocope (my stage 15) and the term "Archaic Irish" for the period between the apocope and the syncope (my stage 19), we may wonder about the applicability of the term "Irish" to the Ogam inscriptions; it may be more appropriate to speak of the variety of Insular Celtic spoken by the ancestors of the Irish. In any case, no reconstruction of Proto-Irish on the basis of Old Irish and later materials comes close to anything resembling the language of the Ogam inscriptions. Since the latter can hardly be older than the beginning of the Christian era and the syncope may be dated to the sixth century, it will be clear that I have little confidence in a theory which relegates Proto-Indo-European to the fifth or sixth millennium BC. The radical changes which embody the formation of Irish in the first half of the first millennium AD are probably due to imperfect learning by speakers of an unknown substrate language which was lost forever.

Perhaps the best example of a disintegrating proto-language is furnished by the Slavic material. Apart from the rise of *x* all the major developments which differentiate Slavic from its Baltic prototype are usually dated to the first millennium AD (e.g., Shevelov 1964, Kortlandt 1982). The earliest dialectal divergences within Slavic which have survived into historical times can hardly be older than the fourth century, and the last shared innovations of the entire group, such as the rise of the neo-acute tone, may be dated to the ninth century. The modern dialectal situation is essentially the same as it was in the twelfth century. When we reconstruct Proto-Slavic, the result can largely be identified with the language of the ninth century, apart from the dialectal differentiation which started half a millennium earlier, apparently in connection with the earliest expansion of the Slavic territory. It is reasonable to assume that many dialects arose and disappeared at earlier stages, but it is not obvious that the term "Slavic" is appropriate before the expansions of the first millennium AD.

This brings me to the third point I want to make here. If a proto-language can be dated to the period of its expansion, the mechanism of this process must be examined in detail. It comprises two phases, each of which has its own dynamics. First, a number of people have to move from their original homeland to a new territory. Second, a larger number of people must find it expedient to adopt the language of the intruders. Both developments are determined by specific social and economic circumstances.

Population movements are determined by three factors. Firstly, there must be a reason to leave one's homeland. This factor has rightly been stressed by Anthony, who observes that people living along the boundary between the poorer lowland steppe and the richer upland forest "risked periodic exposure to severe stress, for small variations in precipitation, temperature, population density, or deforestation rate would dramatically alter the local distribution of critical resources in these fragile borderland communities" (1986: 292). This periodic exposure to severe stress prompted expansion when the opportunity presented itself. Secondly, there must be a place where life seems to be better in order to make the journey worth while. This is a reason to expect migrations toward rather than away from more developed areas such as Assyria in the third and second millennia BC. Thirdly, the cost of the journey must not be prohibitive. It is now generally recognized that the domestication of the horse played a crucial part in reducing the cost of physical mobility.

The expansion of Indo-European presupposes not only the migrations of Indo-Europeans, but also the adoption of Indo-European languages by local populations. With respect to this issue Mallory refers to Barth's work (1981) in a discussion which is really too short. The complexity of the problem is illustrated by the following passage, which I cannot refrain from quoting at some length (Mallory 1989: 260f.):

“Barth examined the linguistic relations between the Pathans and Baluchi on the Afghan-Pakistan border. The Pathans were the more numerous, the wealthier, better armed, and even possessed a better military reputation. Nevertheless, it is the Baluchi who have been making the sustained linguistic assimilation of the Pathans. The Baluchi social structure is hierarchic and encourages vertical relationships between local leaders and clients. The various bands offer opportunities for social advancement within these hierarchies, and displaced Pathans in a frontier situation are attracted individually and in groups to join Baluchi communities. On the other hand, the more egalitarian society of the Pathans was ill-suited to absorb foreigners who could only enter it either in roles despised by the Pathans or by undertaking a more complicated process to being admitted as an equal in Pathan society. The nub of the issue here is not weapons, wealth or population size but the social permeability of the competing social organizations. As numerous historical instances testify, pastoral societies throughout the Eurasian steppe are typified by remarkable abilities to absorb disparate ethno-linguistic groups. Indo-European military institutions may have encouraged membership from local groups in the form of clientship which offered local populations greater advantages and social mobility.”

This must have been the decisive force in the spread of the Indo-European languages.

Starting from the linguistic evidence and trying to fit the pieces into a coherent whole, we arrive at the following picture. The best candidate for the original Indo-European homeland is the territory of the Sredny Stog culture in the eastern Ukraine. The attested languages reflect a number of waves of migration to the east, north of the Caspian Sea (Tocharian, Indo-Iranian), to the south, west of the Black Sea (Anatolian, Greek, Armenian, Albanian), and to the west, south of the Baltic Sea (Italo-Celtic, Germanic). As Mallory notes, there may have been a fourth, abortive wave of migration to the southeast, west of the Caspian Sea, which is not reflected in the linguistic records, perhaps because the Indo-Europeans were assimilated to the local population at an early stage. The earlier migrations yielded the peripheral languages (Tocharian, Anatolian, Italo-Celtic), which did not take part in the late Indo-European innovations of the central dialects (Indo-Iranian, Greek, Germanic, Balto-Slavic, etc.). Some innovations affected only a part of the central dialects, such as the assibilation of the palatovelars (which did not reach Greek and Germanic) or the loss of aspiration in the voiced stops (which did not reach Greek and Indic). Other developments had a more local character. An interconsonantal laryngeal voiced the following stop in North Iranian (Avestan, Sogdian) *dugdar-* ‘daughter’, but not in its Persian and Indic cognates. This must have been a very early development. It appears that Phrygian was rather closely related to Greek (cf. now Lubotsky 1988), Thracian to Armenian (cf. Kortlandt 1988), and Venetic to Italic. The position of Illyrian remains unclear.

The Indo-Europeans who remained after the migrations became speakers of Balto-Slavic. If the speakers of the other satem languages can be assigned to the Yamnaya horizon and the western Indo-Europeans to the Corded Ware horizon, it

is attractive to assign the ancestors of the Balts and the Slavs to the Middle Dnieper culture. If the origin of this culture “is to be sought in the Sredny Stog, Yamnaya and Late Tripolye cultures” and this phase is “followed by a middle period where the classic Corded Ware amphorae and beakers appear” (Mallory 1989: 248), the course of events corresponds nicely with the development of a satem language which was drawn into the western Indo-European sphere of influence. The disintegration of Balto-Slavic is closely parallel to that of Indo-European: the Slavs migrated to the west, the south, and the east, the Latvians to the north, and the Prussians were assimilated to the Germans. The deceptively archaic character of the Lithuanian language may be compared to the calm eye of a cyclone.

The resulting picture can be summarized as follows.

Eastward migrations:

1. Tocharian.
- 2a. Indic.
- 2b. South Iranian.
- 2c. North Iranian.
- (3. East Slavic.)

Southward migrations:

1. Anatolian.
- 2a. Greek.
- 2b. Phrygian.
- 2c. Armenian.
- 2d. Thracian.
- 2e. Daco-Albanian.
- (3. South Slavic.)

Westward migrations:

- 1a. Italic.
- 1b. Venetic
- 1c. Celtic.
2. Germanic.
- (3. West Slavic.)

Once again it must be emphasized that many linguistic groups may have vanished without leaving any historical record.

We must now examine how the view developed here can be related to Gimbutas' theory of two homelands and three waves of migration into the Balkans. The main objection which can be raised against Gimbutas' scheme (e.g., 1985: 198) is that it starts from the archaeological evidence and looks for a linguistic interpretation. As a consequence, the scheme does not fit the linguistic evidence very well.

It seems to me that we arrive at a much better representation if we start from the linguistic side and try to find an archaeological corroboration. The natural solution then is to link Gimbutas' first wave (4400-4200 BC) to the ancestors of the Anatolians, her second wave (3400-3200 BC) to the ancestors of the Greeks and the Phrygians, and her third wave (3000-2800 BC) to the ancestors of the Armenians and the Thracians. If this identification is correct, the satemization process can be dated to the last centuries of the fourth millennium. It is possible that the speakers of Italo-Celtic must be assigned to the Globular Amphora culture, and that Germanic grew out of a later component of the Corded Ware horizon. Since the beginnings of the Yamnaya, Globular Amphora, Corded Ware, and Afanasievo cultures can all be dated between 3600 and 3000 BC, I am inclined to date Proto-Indo-European to the middle of the fourth millennium, and to recognize Proto-Indo-Hittite as a language which may have been spoken a millennium earlier.

If we can identify Indo-Hittite and Indo-European with the beginning and the end of the Sredny Stog culture, respectively, it will be clear that the linguistic evidence from our family does not lead us beyond Gimbutas' secondary homeland and that the Khvalynsk culture on the middle Volga and the Maykop culture in the northern Caucasus cannot be identified with the Indo-Europeans. Any proposal which goes beyond the Sredny Stog culture must start from the possible affinities of Indo-European with other language families. It is usually recognized that the best candidate in this respect is the Uralic language family, while further connections with the Altaic languages and perhaps even Dravidian are possible. The hypothesis that Indo-European is genetically related to a Caucasian language family or to Afro-Asiatic seems much less probable to me. What we do have to take into account is the typological similarity of Proto-Indo-European to the North-West Caucasian languages. If this similarity can be attributed to areal factors, we may think of Indo-European as a branch of Uralo-Altaic which was transformed under the influence of a Caucasian substratum. It now appears that this view is actually supported by the archaeological evidence. If it is correct, we may locate the earliest ancestors of the speakers of Proto-Indo-European north of the Caspian Sea in the seventh millennium (cf. Mallory 1989: 192f.). This is essentially in agreement with Gimbutas' theory. [Cf. now Kortlandt 1989.]

REFERENCES

- Agard, F.B.
1984 *A course in Romance linguistics 2: A diachronic view*. Washington DC: Georgetown UP.
- Anthony, D.W.
1986 The 'Kurgan culture', Indo-European origins, and the domestication of the horse: A reconsideration, *Current Anthropology* 27, 291-304.

- Barth, F.
1981 *Selected essays*, 2 vols., London: Routledge & Kegan Paul.
- Gimbutas, M.
1985 Primary and secondary homeland of the Indo-Europeans, *Journal of Indo-European Studies* 13, 185-202.
- Kortlandt, F.
1979 The Old Irish absolute and conjunct endings and questions of relative chronology, *Ériu* 30, 35-53.
1981 1st sg. middle *-H₂, *Indogermanische Forschungen* 86, 123-136.
1982 Early dialectal diversity in South Slavic I, *Studies in Slavic and General Linguistics 2: South Slavic and Balkan Linguistics*, 177-192.
1988 The Thraco-Armenian consonant shift, *Linguistique Balkanique* 31, 71-74.
1989 Eight Indo-Uralic verbs?, *Münchener Studien zur Sprachwissenschaft* 50, 79-85.
- Lubotsky, A.
1988 The Old Phrygian Areyastis-inscription, *Kadmos* 27, 9-26.
- Mallory, J.P.
1989 *In search of the Indo-Europeans: Language, archaeology and myth*. London: Thames and Hudson.
- Mayrhofer, M.
1983 *Sanskrit und die Sprachen Alteuropas: Zwei Jahrhunderte des Widerspiels von Entdeckungen und Irrtümern*, Göttingen: Vandenhoeck & Ruprecht.
- Penney, J.H.W.
1989 Preverbs and postpositions in Tocharian, *Transactions of the Philological Society* 87, 54-74.
- Schmitt, R.
1974 Proto-Indo-European culture and archaeology: Some critical remarks, *Journal of Indo-European Studies* 2, 279-287.
- Shevelov, G.Y.
1964 *A prehistory of Slavic: The historical phonology of Common Slavic*. Heidelberg: Carl Winter Universitätsverlag.