C.C. UHLENBECK ON INDO-EUROPEAN, URALIC AND CAUCASIAN

FREDERIK KORTLANDT

In his early years, C.C. Uhlenbeck was particularly interested in the problem of the Indo-European homeland (1895, 1897). He rejected Herman Hirt's theory (1892) that the words for 'birch', 'willow', 'spruce', 'oak', 'beech' and 'eel' point to Lithuania and its immediate surroundings and returned to Otto Schrader's view (1883, 1890) that the original homeland must rather be sought in southern Russia and may have included some of the later Germanic and Iranian territories. It is clear that the Mediterranean region and the area around the North Sea can safely be excluded because the arrival of the Indo-Europeans was comparatively recent here, as it was in Iran and the Indian subcontinent. It is difficult to be more specific within the limits of central and eastern Europe and central Asia. Uhlenbeck was impressed by the lexical correspondences between Indo-European and Semitic which had been adduced in favor of an eastern homeland but pointed out that borrowings from Semitic may have reached the Indo-Europeans through an intermediary. He agrees that the Indo-European words for trees and animals point to a moderate climate but questions the possibility of a more specific localization as well as the concept of homeland itself.

Uhlenbeck identifies the Slavic word for 'dog' *pĭsŭ* with the Indo-European word for 'livestock' **peḱu* and its original meaning as 'domestic animal'. Unlike Hirt (1895), he recognizes that the Indo-Europeans were pastoralists before they became agriculturalists, as is clear from the absence of common words for 'plough', 'field', 'grain' and suchlike. While Armenian shares many agricultural terms with the languages of Europe, these are absent from Indo-Iranian. The common Indo-European vocabulary reflects a stage of development when weaponry was made of stone, wood, bones and hides (cf. Schrader 1890: 320–346). It includes words for 'cart' (ὄχος), 'wheel' (κύκλος), 'axle' (ἄξων), 'yoke' (ζυγόν), 'carpenter' (τέκτων), '(wooden) house' (δόμος), 'vessel' (ναῦς), 'to plait' (πλέκω), 'to weave' (ὑφαίνω), 'to spin' (νέω), 'garment' (εἷμα) and 'to clothe' (ἕννῡμι).

The population of Denmark and Scandinavia did not speak an Indo-European language before the advent of pastoralism, which puts these countries beyond the original homeland. The introduction of the cart was evidently more recent than the domestication of cattle, which were used as draught animals. All Indo-European languages have the same words for 'ox', 'sheep', 'goat', 'horse' and

'swine', but not for 'pig' and tame birds, which had not yet been domesticated, nor for 'donkey', which came from the south. The ancient character of cattle-breeding is corroborated by the words for 'livestock', 'to herd', 'herd' (πῶυ), 'herdsman' (ποιμήν), 'cow' (βοῦς), 'to milk' (ἀμέλγω), 'butter' (ἔλπος) and 'bull'. The word *steer* (ταῦρος, Latin *taurus*) is probably an early Indo-European bor-rowing from Semitic denoting 'wild bull', as is indicated by Lithuanian *taũras* 'aurochs' and Old Prussian *tauris* 'wisent', while *ox* is the old word for 'domesti-cated bull'. Cattle were the most important animals in prehistoric pastoralism, as is clear from such Vedic expressions as 'desire of cows' for 'struggle' and 'cattle-master of horses' for 'lord' and from the peculiar Greek word ἱπποβουκόλος 'horse-cowherd' for 'horse-herd'. Other domesticated animals included sheep, as is also clear from the words for 'wool' and 'lamb', and goats and horses, but not pigs, donkeys, chicken, ducks and geese, which were domesticated more recently (cf. Schrader 1890: 390).

While the Indo-European vocabulary contains an abundance of words re-flecting a pastoral society, there is no common agricultural terminology. The Indo-Iranians evidently belonged to a different cultural unity when the lan-guages of Europe, including Armenian, developed their agricultural terminol-ogy (cf. now Kuz'mina 2007). The Latin word *grānum* originally meant 'grain of corn', as is also clear from its Germanic, Baltic and Slavic cognates, while *hor-deum* 'barley' may have designated a wild variety. The Sanskrit cognate of the word for 'field' *ájras* means 'plain' while the word *vápati* means both 'throws' and 'sows'. The European words for 'seed' (Latin *sēmen*), 'to mow' (ἀμάω), 'to mill' (Latin *molere*), 'to plough' (ἀρόω) and 'plough' (ἄροτρον) are absent from Indo-Iranian (cf. Schrader 1890: 410). These etyma can now be identified with Hittite *šai–* 'to throw', *ānš–* 'to wipe', *malla–* 'to grind' and *harra–* 'to crush', re-spectively (cf. Kloekhorst 2007). The word for 'wine' has an Indo-European etymology (cf. Beekes 1987) but was limited to the Mediterranean countries (It-aly, Greece, Asia Minor), from where it spread to northern and eastern Europe and to the Middle East. The original Indo-European word probably denoted 'vine' rather than 'wine' because this is the meaning of Basque *ayen*, *aihen*.

Though metallurgy was unknown to the Indo-Europeans, they had words for *gold*, which is cognate with *yellow* and has been preserved in Germanic, Baltic, Slavic and Indo-Iranian, 'silver' (Latin *argentum*), cognate with 'bright' and pre-served in Celtic, Italic, Greek, Armenian and Iranian, 'copper' (Latin *raudus* 'piece of brass'), which is cognate with *red* and has been preserved in Italic, Ger-manic (where it may also be represented by *lead*), Slavic and Indo-Iranian, and *ore*, which became the word for 'brass, bronze' in Italic (Latin *aes*), Germanic and Indo-Iranian. There were no common Indo-European terms for 'iron', 'lead' and 'tin'. The original words for 'gold' and 'silver' were apparently replaced with the advent of metallurgy, the first in Italy (Latin *aurum*), from where the new term spread to Baltic, Celtic and eventually Basque, and the second in Spain

ond in Spain (Celtiberian *silabur, Basque zillar), from where the new term spread to Germanic, Baltic and Slavic. The latter word cannot be separated from Berber aẓref 'silver' and aẓarif 'alum', where a– is a nominal prefix. Here the word for 'alum' is evidently a borrowing from Punic and can be derived from the Semitic root ṣrp, Akkadian ṣarāpu 'to refine (metals by firing)', Hebrew ṣārap 'to smelt (metal)' (see Boutkan & Kossmann 2001 for references).

I conclude that Uhlenbeck was well ahead of his time in his discussion with Hirt and Schrader. He recognized that it is necessary to distinguish between two components of Indo-European language and culture, an older common inheritance which reflects a pastoral society and a later European complex with a common agricultural vocabulary, both of them dating from before the introduction of metallurgy. It is interesting that before the end of the 19th century he had already reached the position which has now become dominant among Indo-Europeanist scholars and is supported by the archaeological evidence (cf. Mallory 1989). The major point which he did not see is the crucial role of the domesticated horse in the Indo-European expansions (but see below).

The tentative localization of the Indo-European homeland was logically followed by the question if the proto-language could be related to other language families. Uhlenbeck remarked that the identity between the nominative and the accusative in the neuter, both singular and plural, points to an original absolutive case ("Passivus") which was identical with the bare stem (except in the o–stems), whereas the subject of transitive verbs was in an ergative case ("Aktivus"), marked by a suffixed *–s which he identified with the demonstrative pronoun *so (1901). His student Nicolaas van Wijk argued that the nominal genitive singular in *–s was identical with the original ergative (1902). Among other things, he adduced such constructions as Latin miseret mē 'I feel pity', pudet mē 'I am ashamed', where the logical subject is in the genitive case. Uhlenbeck claimed that the Indo-European proto-language was characterized by polysynthesis, suffixation and infixation and drew attention to its typological similarity to unrelated languages such as Basque, Dakota and Greenlandic. He also observed that the Indo-European mediopassive voice is reminiscent of the verbal construction with an incorporated dative and an object in the absolutive case which is found in Basque and North American languages. In a later study (1904), he adduced the strong resemblance between Basque and Indo-European nominal composition as an example of typological similarity between unrelated languages ("Sprachen zwischen welchen man selbst keine entfernte Verwandtschaft nachzuweisen vermag", cf. also Uhlenbeck 1913). While he considered a common origin of Eskimo and Aleut with the Uralic languages probable (1905a), he rejected the possibility that Basque is related to Uralic and Altaic and suggested that it might rather be of Afro-Asiatic provenance (1905b). After a detailed examination of the available evidence, Uhlenbeck concluded that a genetic relationship between Basque and Caucasian languages cannot be estab-

lished (1923), but later he changed his opinion and considered the latter to be highly probable ("de onmiskenbare verwantschap met het Kaukasisch", 1946: 17), regarding the Afro-Asiatic elements as borrowings.

Holger Pedersen has listed numerous examples of Russian impersonal sentences with an inanimate agent in the instrumental case, e.g. *tečeniem ego poneslo nazad* 'the current carried him back', *vetrom sneslo kryšu* 'the wind blew off the roof', and similar constructions in Iranian, Celtic and Germanic (1907: 134–140). He argues that this sentence type is older than the rise of grammatical gender in Indo-European and compares it with the ergative construction in North Caucasian languages, where the subject of a transitive verb is in the instrumental or in the genitive, as it is in Tibetan and Eskimo, and in the Armenian *l*–preterit, where it is in the genitive. For the Indo-European proto-language he proposes that the subject of an intransitive verb and the object of a transitive verb were in the absolutive (unmarked) case while the subject of a transitive verb was in the genitive when the agent was animate but in the instrumental when it was inanimate (1907: 152). After the differentiation between the ergative and the genitive, the former came to be used for the subject of intransitive verbs in the case of animate individuals, as distinct from collectives and inanimates. This three-way distinction was subsequently grammaticalized as masculine, feminine and neuter gender. In a later study, Pedersen argues that a comparison of the Indo-European and Uralic grammatical systems leaves no doubt about their genetic relationship ("Es liegt hier eine Summe von Übereinstimmungen vor, die den Zufall ausschliesst", 1933: 309) and proposes that athematic presents (τίθημι 'I put') were originally transitive while thematic presents (φέρω 'I carry') and perfects (οἶδα 'I know') were originally intransitive, which is reminiscent of Hungarian *várom* 'I wait for him' versus *várok* 'I wait'.

In the meantime, Uhlenbeck's doubts about the reality of an Indo-European proto-language had grown. He now defined the proto-language as the group of dialects spoken by the original community of Indo-European conquerors. Since the conquests of non-Indo-European territories took place at different times, the language of the later conquerors was no longer identical with the original proto-language. While the separate branches of Indo-European arose when the language of the invaders was adopted by local populations speaking quite different substratum languages, Uhlenbeck claims that Proto-Indo-European itself already consists of two unrelated groups of elements, which he calls *A* and *B* (1933, 1934a, 1937b). Here *A* contains pronouns, verbal roots and derivational suffixes whereas *B* contains isolated words which are not related to verbal roots, such as numerals, some kinship terms, and many names of body parts, animals and trees. Uhlenbeck compares *A* with Uralic and Altaic and attributes irregular features such as heteroclitic inflection and grammatical gender to *B*, for which one might think of Caucasian languages. The relation between Indo-European and Uralic can be extended to Eskimo (cf. Uhlenbeck 1905a, 1906, 1907, 1934b, 1937a,

1941, 1942). This is in accordance with the view that the Indo-Europeans arrived in southern Russia from the Asian steppes east of the Caspian Sea, where they allegedly led a nomadic life with horses, chariots and large herds of cattle.

Since such words as kinship terms and names of body parts are usually regarded as belonging to the basic vocabulary of a language, Uhlenbeck rejected the terms "genetic relationship" for *A* and "borrowing" for *B* and presented Proto-Indo-European as a "mixed language", an idea which had first been put forward by Sigmund Feist (1910). This is an unfortunate notion which is based on underanalysis of the data and can easily lead to muddled thinking about linguistic contact and language change (cf. Kortlandt 2000). The point is that the two components *A* and *B* have an entirely different status. The situation is reminiscent of Michif, which has been adduced as a prime example of a mixed language. Here we find numerous French nominal stems which were borrowed together with their determiners, e.g. *le loup* 'the wolf', *sa bouche* 'his mouth', *son bras* 'his arm', while the verbal stems and grammatical elements are purely Cree (cf. Kortlandt 2000: 123). Uhlenbeck himself adduces the Sanskrit influence on Indonesian, the French influence on English and the Romance influence on Basque as parallels (1941: 204f.). He appears to have realized his mistake because he later returned to an analysis in terms of genetic relationship and borrowing (1946).

The two major findings which Uhlenbeck has contributed to Indo-European linguistics are the reconstructed ergative (which was established independently by Pedersen (1907: 157), who provided the comparative evidence) and the twofold origin of the vocabulary. Both discoveries have been slow in their acceptance by the scholarly community. André Vaillant has identified the Indo-European ergative in *-s* as an original ablative, the animate accusative in *-m* as a lative (casus directivus), and the neuter pronominal ending *-t* with the instrumental ending in Hittite and the ablative ending of the *o*-stems in the other Indo-European languages (1936). The ending *-m* was originally limited to animate individuals, like the preposition *a* in Spanish, e.g. *veo a Pedro* 'I see Peter' (see Pottier 1968 for details). Robert Beekes has shown that the entire paradigm of the *o*-stems was built on an ergative case form in *-os* (1985). Blissfully ignorant of the data and unaware of the comparative evidence, Alan Rumsey has argued on typological grounds that there cannot have been a Proto-Indo-European ergative because this case is absent from the neuter paradigm (1987a, 1987b). Since his objection was effectively answered by Pedersen a hundred years ago (1907), there is no reason to return to the matter here. It illustrates how a tool which in itself is useful becomes harmful in the hands of the unskilled (cf. also Kortlandt 1995).

The idea of a genetic relationship between Indo-European and Uralic has become fairly well accepted among specialists (e.g. Collinder 1965, 1974). Gimbutas's theory that the Indo-Europeans moved from a primary homeland north of

the Caspian Sea to a secondary homeland north of the Black Sea (e.g. 1985) is fully in agreement with the view that their language developed from an Indo-Uralic proto-system which was modified under the influence of a North Caucasian substratum, perhaps in the sixth millennium BC (cf. Mallory 1989: 192f., Kortlandt 1990, 2002). Johannes Knobloch has suggested that the thematic vowel *–e/o– in the Indo-European verbal inflection represents an earlier object marker (1953). I have argued that the thematic present and the perfect originally had a dative subject, reflecting an earlier intransitive construction with an indirect object (1983). For Proto-Indo-Uralic we can reconstruct a genitive in *–n, which is reflected in the oblique stem form of the Indo-European heteroclitics, a lative-accusative in *–m, a dative-locative in *–i, an ablative-instrumental in *–t, which is reflected as both –t and –s in Indo-European, plural markers *–t and *-i, dual *–ki, personal pronouns *mi 'I', *ti 'thou', *me 'we', *te 'you' and corresponding verbal endings, reflexive *u, demonstratives, participles, derivational suffixes of nouns and verbs, negative *n– and interrogative *k– (cf. Kortlandt 2002). The rise of the ergative construction, grammatical gender and adjectival agreement can be attributed to North Caucasian influence and may have proceeded as indicated by Pedersen (1907). It is important to note that the accusative is of Indo-Uralic origin and therefore older than the ergative. This explains the peculiar construction of Russian *vetrom sneslo kryšu* 'the wind blew off the roof', where the inanimate agent is in the instrumental and the object is in the accusative. While the Indo-Uralic component of the lexicon (Uhlenbeck's *A*) has been a focus of research in the past, the identification of the non-Indo-Uralic component (Uhlenbeck's *B*) remains a task for the future. In view of the large number of consonants and the minimal vowel system of Proto-Indo-European, the northern Caucasus seems to be the obvious place to look (cf. Starostin 2007).

Leiden University

REFERENCES

Beekes, Robert S.P.
  1985   *The origins of the Indo-European nominal inflection.* Innsbruck: Institut für Sprachwissenschaft.
  1987   "On Indo-European 'wine'". *Münchener Studien zur Sprachwissenschaft* 48, 21–26.
Boutkan, Dirk, & Maarten Kossmann
  2001   "On the etymology of 'silver'". *North-Western European Language Evolution* 38, 3–15.
Collinder, Björn
  1965   *Hat das Uralische Verwandte?* Uppsala: Almqvist & Wiksell.

1974	"Indo-Uralisch – oder gar Nostratisch?". In: Manfred Mayrhofer et al. (eds.), *Antiquitates Indogermanicae*, 363–375. Innsbruck: Institut für Sprachwissenschaft.

Feist, S.
1910	*Europa im Lichte der Vorgeschichte*. Berlin: Weidmann.

Gimbutas, Marija
1985	"Primary and secondary homeland of the Indo-Europeans". *Journal of Indo-European Studies* 13, 185–202.

Hirt, Herman
1892	"Die Urheimat der Indogermanen". *Indogermanische Forschungen* 1, 464–485.
1895	"Der Ackerbau der Indogermanen". *Indogermanische Forschungen* 5, 395–402.

Kloekhorst, Alwin
2007	*The Hittite inherited lexicon*. Diss. Leiden.

Knobloch, Jean
1953	"La voyelle thématique –*e/o*– serait-elle un indice d'objet indo-européen?" *Lingua* 3, 407–420.

Kortlandt, Frederik
1983	"Proto-Indo-European verbal syntax". *Journal of Indo-European Studies* 11, 307–324.
1990	"The spread of the Indo-Europeans". *Journal of Indo-European Studies* 18, 131–140.
1995	"General linguistics and Indo-European reconstruction". *Rask* 2, 91–109.
2000	"On Russenorsk". *Amsterdamer Beiträge zur älteren Germanistik* 54, 123–127.
2002	"The Indo-Uralic verb". In: R. Blokland & C. Hasselblatt (eds.), *Finno-Ugrians and Indo-Europeans: Linguistic and literary contacts*, 217–227. Maastricht: Shaker. [electronic edition: www.kortlandt.nl/publications]

Kuz'mina, Elena E.
2007	*The origin of the Indo-Iranians*. Leiden: Brill.

Mallory, James P.
1989	*In search of the Indo-Europeans*. London: Thames & Hudson.

Pedersen, Holger
1907	"Neues und nachträgliches". *Zeitschrift für vergleichende Sprachforschung* 40, 129–217.
1933	"Zur Frage nach der Urverwandtschaft des Indoeuropäischen mit dem Ugrofinnischen". In: Walter Anderson (ed.), *Liber Semisaecularis Societatis Fenno-Ugricae*, 308–325. Helsinki: Suomalais-Ugrilainen Seura.

Pottier, Bernard
1968	"L'emploi de la préposition *a* devant l'objet en espagnol". *Bulletin de la Société de Linguistique de Paris* 63, 83–95.

Rumsey, Alan
1987a	"Was Proto-Indo-European an ergative language?" *Journal of Indo-European Studies* 15, 19–37.
1987b	"The chimera of Proto-Indo-European ergativity". *Lingua* 71, 297–318.

Schrader, Otto
1883	*Sprachvergleichung und Urgeschichte*. Jena: Costenoble.
1890	*Sprachvergleichung und Urgeschichte*. 2. Auflage. Jena: Costenoble.

Starostin, Sergej A.
2007	"Indoevropejsko-severnokavkazskie izoglossy". In: *Trudy po jazykoznaniju*, 312–358. Moskva: Jazyki slavjanskix kul'tur.

Uhlenbeck, C.C.
    1895   "Waar werd de Indogermaansche stamtaal gesproken?" *Tijdschrift voor Nederlandsche Taal- en Letterkunde* 14, 69–74.
    1897   "De voorgeschiedenis der Indogermaansche volken." *Taal en Letteren* 7, 1–25.
    1901   "Agens und Patiens im Kasussystem der indogermanischen Sprachen". *Indogermanische Forschungen* 12, 170–172.
    1904   "Eine baskische Parallele". *Indogermanische Forschungen* 17, 436–441.
    1905a   "Uralische Anklänge in den Eskimosprachen". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 59, 757–765.
    1905b   "Baskisch und Uralaltaisch". *Beiträge zur Kunde der indogermanischen Sprachen* 29, 305–307.
    1906   "Zur Eskimogrammatik". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 60, 112–114.
    1907   "Zur Eskimogrammatik". *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 61, 435–438.
    1913   "Baskisch und Indogermanisch". *Indogermanische Forschungen* 33, 171–172.
    1923   "Over een mogelijke verwantschap van het Baskisch met de Palaeo-Kaukasische talen". *Mededeelingen der Koninklijke Akademie van Wetenschappen, Afdeeling Letterkunde,* A 55/5, 105–137.
    1933   "Eine Bemerkung zur Frage nach der Urverwandtschaft der uralischen und indogermanischen Sprachen". In: Walter Anderson (ed.), *Liber Semisaecularis Societatis Fenno-Ugricae*, 396–397. Helsinki: Suomalais-Ugrilainen Seura.
    1934a   "Oer-Indogermaansch en Oer-Indogermanen". *Mededeelingen der Koninklijke Akademie van Wetenschappen, Afdeeling Letterkunde,* A 77/4, 125–148.
    1934b   "Eskimo en Oer-Indogermaansch". *Mededeelingen der Koninklijke Akademie van Wetenschappen, Afdeeling Letterkunde,* A 77/6, 179–196.
    1937a   "Über den Wert eskimoisch-indogermanischer Wortähnlichkeiten". In: Louis Hjelmslev et al. (eds.), *Mélanges linguistiques offerts à M. Holger Pedersen*, 110–114. Aarhus: Universitetsforlaget, København: Levin & Munksgaard.
    1937b   "The Indogermanic mother language and mother tribes complex". *American Anthropologist* 39, 385–393.
    1941   "Oude Aziatische contacten van het Eskimo". *Mededeelingen der Nederlandsche Akademie van Wetenschappen, Afdeeling Letterkunde,* NR 4/7, 201–227.
    1942   "Ur- und altindogermanische Anklänge im Wortschatz des Eskimo". *Anthropos* 37-40, 133–148.
    1946   "Gestaafde en vermeende affiniteiten van het Baskisch". *Mededeelingen der Koninklijke Nederlandsche Akademie van Wetenschappen, Afdeeling Letterkunde,* NR 9/2, 13–24.
Vaillant, André
    1936   "L'ergatif indo-européen". *Bulletin de la Société de Linguistique de Paris* 37, 93–108.
Van Wijk, Nicolaas
    1902   *Der nominale Genetiv Singular im Indogermanischen in seinem Verhältnis zum Nominativ.* Zwolle: Tijl.