

The Indo-Europeanization of Europe: An Introduction to the Issues¹

BRIAN D. JOSEPH

Distinguished University Professor of Linguistics
Kenneth E. Naylor Professor of South Slavic Linguistics
The Ohio State University

INTRODUCTION TO THE PLAYERS AND THE GAME

The panel represented in these pages explored events that are from the distant past but have had a broad impact, even into the present day, affecting everyone in the United States and in North America more generally, but even further, in all parts of the globe. The particular events in question are what can be called the Indo-Europeanization of Europe.

Indo-Europeanization refers to the arrival in Europe of a large family of languages, where the term *family* indicates that the languages stem from a common source, and thus have common parentage.² While there are many language families in the world, the particular family of interest here is the one that includes English, Greek, Hindi, Russian, and literally dozens of other modern—and ancient—languages ranging from India to Europe.³ This geographic spread is the basis for the usual name for the family, *Indo-European*. The family itself and the relations among the members of the family are often represented as a tree, mimicking the representation of a biological family; a slightly tongue-in-cheek, but nonetheless useful, rendition of the family with its 10 major branches is given in Figure 1, with a version given in Figure 2 that is more conventional in form, for linguists at least.⁴

1 Read 29 April 2017 as part of the *Indo-Europeanization of Europe* symposium.

2 Note that in the case of related languages, there is only one “parent.”

3 And now beyond, as well, due to colonization.

4 This type of representation for language relationships was lent to linguistics from stemmatology, the study of the relationships among different manuscripts. Charles Darwin may well have borrowed the idea of a biological family tree for relations among species from comparative linguistics of the mid-19th century; in his *On the Origin of Species*, he writes, “It may be worth while to illustrate this view of classification, by taking the case of languages” (Darwin 1860, 422). It should be noted that in biological phylogenetics, the family trees tend to be schematized with the “parent” node (representing the “root” of the tree) at the bottom, while linguists more conventionally have that node at the top (as in Figure 2, thus not a “root” in the same way as in biological “trees,” or real trees, for that matter, as in Figure 1).



FIGURE 1. A representation of the Indo-European language family as a real tree.

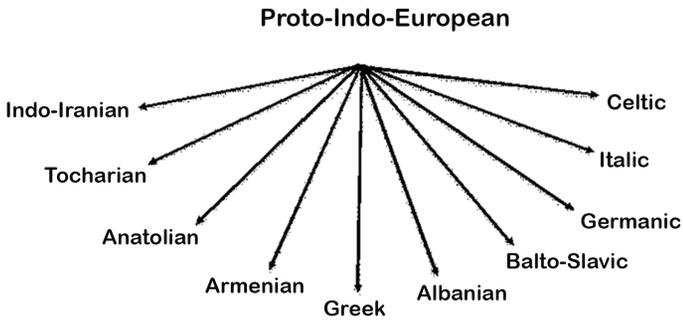


FIGURE 2. A schematic "tree" representation of the Indo-European language family.

The representation in Figure 1 gives the 10 major individual language branches of the Indo-European family from right to left in roughly the east-to-west position of each branch at its earliest attestation; no significance should be placed on the size, length, or thickness of the branches of the tree. The photo in Figure 1 is my own, taken on the island of Palawan in the Philippines in November 2017; Figure 2 is an original drawing.

Languages that stand in such a relation to one another are said by linguists to be “genetically related” where the operative sense of *genetic* is not the biological notion that panelist David Reich writes about, but rather the sense that derives from the original meaning of the word in Greek; the relevant Greek word is γενετικός (*genetikós*) and it means “pertaining to origins,”⁵ so a linguistic genetic relationship means that the languages have a common origin. Still, due to possible confusion with the biological notion, and to avoid any misleading and mistaken assumptions that biological DNA has anything to do with which language a given person speaks, many historical linguists these days use the term *genealogical* to refer to language relatedness.

Indo-European is an appropriate term, given where the languages of the family were located when serious study of them and of their interrelationships began in the late 18th and early 19th centuries, but the family has been called by various names over the years. Some of these names can still be encountered.

Indo-Hittite is one such name, based on the conventional label of the most fully documented language of the Anatolian branch of the family, *Hittite*,⁶ which is one of the earliest-attested Indo-European languages, written in a cuneiform script on clay tablets dating from the first half of the second millennium BCE. This appellation, however, has more to do with the theory behind the internal structure of relations of member languages within the family, with some linguists arguing for an early split of the Anatolian languages from the remaining branches. Whatever the merits of the Indo-Hittite hypothesis, we can say that talking about the Indo-Hittite-ization of Europe here, besides being a mouthful, would not quite work since the specific branches of Indo-Hittite that made it into Europe were not the languages of the Anatolian branch, that is, Hittite and the languages most closely related to Hittite, such as Luvian, Palaic, Lycian, and Lydian.⁷ So *Indo-Hittite* is not a suitable term for our purposes here.

Another such name is *Indo-Germanic*, based on German *indogermanische*, from which Indo-Germanicization of Europe would be a possible turn of phrase. However, this probably is infelicitous due to images it might invoke of a World War II-like German territorial expansion.

5 The adjective γενετικός is related to the noun γένεσις (*génesis*), meaning “origin.”

6 For what it is worth, the Hittites themselves referred to their language as *nešili*.

7 All of the languages of the Anatolian branch, despite being attested early in the second millennium BCE, died out early in the start of the Common Era.

Virtually never used now is the term *Japhetic*, though it does occur in pre-modern writings, for instance those of Karl Richard Lepsius (e.g., Lepsius 1981), a noted Egyptologist at the University of Berlin in the middle of the 19th century. This term was based on an idea popular in Europe in the 1600s that Japhet, one of Noah's three sons, was the ultimate source of the languages and peoples of Europe; his two brothers, S(h)em and Ham, were seen as the sources of the Semitic and Hamitic peoples, respectively. For the purposes of the panel represented here, talking about the *Japheticization* of Europe just does not have the same ring to it as *Indo-Europeanization* and in any case the term would be rather opaque to the modern reader.

Thus in what follows, the currently prevailing terminology of *Indo-European* and thus *Indo-Europeanization* is adopted, without further argument.

THE ISSUES

In the standard view of the Indo-Europeanization of Europe, starting from a small group of speakers many millennia ago somewhere in the region between Europe and Central Asia—the so-called Indo-European homeland—the Indo-Europeans spread westward into Western Europe, as well as eastward into the Indian subcontinent. One recent schematization of this movement is shown in Figure 3.

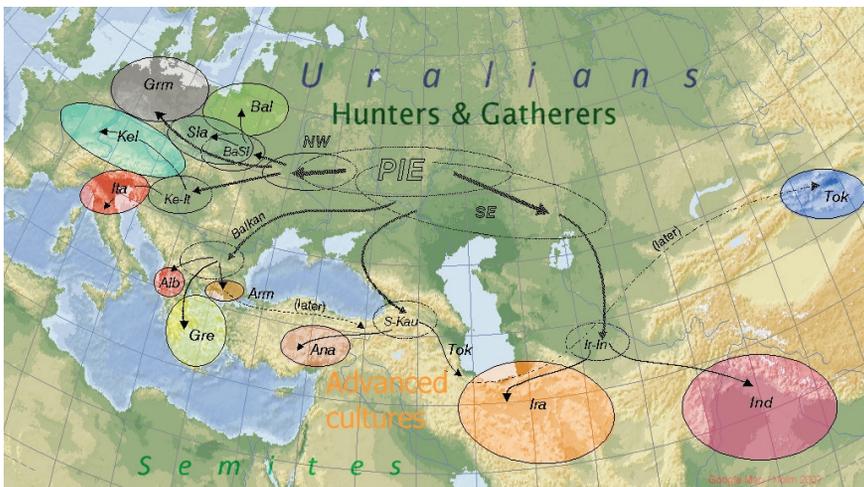


FIGURE 3. A map of possible dispersal patterns for the Indo-Europeans, showing possible migration routes between the 35th and 25th centuries BC. See Holm ([2007] 2008). Courtesy of Dr. Hans J. Holm.

I am being deliberately vague about various details here because many aspects of this movement are a matter of controversy. Still, some of the key issues can be identified as the following:

- Where was that homeland?
- When was that homeland?
- What led the Indo-Europeans to move as they did and where they did?

I offer brief answers to each of these key questions by way of introducing the issues dealt with in the panel and treated in the two papers that follow.

Regarding “where,” the key area of dispute is whether the Indo-European homeland is to be located somewhere in the steppes of Eastern Europe, in the steppes of eastern Ukraine and southern Russia (the home of the so-called “Kurgan” Culture; cf., e.g., Gimbutas 1956, 1970), or instead in Anatolia (Renfrew 1987).

As for the “when” question, the conventional date that linguists have generally worked with for the date of the Proto-Indo-European speech community is c. 6500 BP, a date that allows for the development of the diversity of structure and lexicon that one sees in the attested Indo-European languages of the second millennium BCE (specifically Hittite and some other Anatolian languages, Greek, and Sanskrit) while at the same time recognizing the essential commonality among them that allows them to be considered members of a single language family. Others, working with different methodologies, in particular borrowing the phylogenetic methods employed by evolutionary biologists, have come up with deeper dates, such as c. 9000 BP (cf. Gray and Atkinson 2003).

As for the “why” question, there is the view, associated with Renfrew’s Anatolian hypothesis, that it was the spread of agriculture that drove the Indo-Europeans’ mobility and their ultimate movement into Europe. Alternatively, in Gimbutas’s view, the Indo-Europeans were a nomadic pastoral society and moved about for the sake of their pastoralist economy but also in a way that was characterized by militaristic invasions aided by the use of the domesticated horse and chariots.

WHY ANYONE IS PURSUING THIS QUESTION AND WHY ONE SHOULD BE DOING SO

On April 11, 2017, at a meeting of the foreign ministers of the G-7 coalition, then-Secretary of State Rex Tillerson famously asked, “Why should U.S. taxpayers be interested in Ukraine?” We might ask the

same question here both in a slightly altered form and in its original form. The original form of Secretary Tillerson's question is especially appropriate here if the Indo-European homeland is in fact located in the general area of today's Ukraine, as just noted and discussed further in the articles to follow. The altered form would be: Why should what happened with the Indo-Europeans thousands of years ago have any importance to anyone now, especially people in the United States?

So why indeed should we consider the Indo-Europeanization of Europe to be an important and interesting issue? This question has several compelling answers.

First, it is a question that is intrinsically interesting because answering it involves an attempt to gain insight into events that occurred thousands of years ago. There is thus an aspect to this investigation that is rather like solving a mystery, trying to piece together bits of evidence and to reconstruct the distant past. Moreover, it is indeed a matter of piecing together different strands of evidence since it involves multidisciplinary perspectives and methodologies, drawing on linguistics, archaeology, biology, genetics, and even computer science, among other disciplines. So in a sense, there is something for everyone in exploring this question.

Second, the Indo-European languages are important in practical terms now for several reasons, having to do with geography and demographics, but also what we can call humanistic and scientific achievements. Linguists generally think of all languages as being important in the sense of being able to tell us something about the human faculty of language and about how language emerges and develops in individuals and in societies.⁸ Nonetheless, it is also a fact of life that through various accidents of history, various individual languages have come to take on a greater social, cultural, political, and economic importance for large numbers of people, and this can certainly be said about Indo-European.

For instance, Indo-European languages today are spoken by literally millions and millions of speakers. Seven of the top 10 languages as measured by the number of native speakers are Indo-European: Spanish, English, Hindi, Portuguese, Bengali, Russian, and Punjabi, totaling some 1,760,000,000 speakers as of 2010, nearly 26 percent of the world's population. Moreover, this number is greater yet if one adds in second-language speakers who, even if not native speakers, use

8 Another way of stating this assumption on the part of linguists—a way that is particularly appropriate in Philadelphia, the home of the American Philosophical Society—is that we hold it to be a “truth [that is] self-evident, that all languages are created equal.”

an Indo-European language on a regular basis. English is used as a second language in many parts of the world, for instance in India and Kenya, and French is used as a second language as well, for instance in West Africa. With approximately 900,000,000 such speakers the count swells to roughly 2,660,000,000—about 33 percent of the world's total population. This means that one in every three people in the world today natively speaks or routinely uses an Indo-European language. Thus the demographics of Indo-European are truly global.

Furthermore, from a political standpoint, Indo-European languages turn out to be important. That is, Indo-European languages are official state—that is to say, national—languages on all continents except Antarctica. This reckoning includes English in India, the Philippines, and Singapore, by way of taking in Asia; English and Afrikaans in South Africa; Portuguese in Angola and Mozambique; and French in Senegal and Benin, by way of taking in Africa.⁹

Even better perhaps, looking to nonpolitical reasons, it is easy to enumerate various great human accomplishments that happen to have involved Indo-European languages. For instance, several important religions have basic foundational works that were composed in Indo-European languages: Buddhism has the *Dhammapada*, composed in the Indo-Iranian (more specifically Indo-Aryan) language Pāli; Christianity has the New Testament, composed in Greek; Hinduism has the hymns of the Rigveda, composed in Vedic Sanskrit; and Zoroastrianism has the Gathas, composed in the ancient Iranian language Avestan. Moreover, many monumental works of literature have been composed in Indo-European languages; focusing just on ancient Indo-European languages, one can cite the ancient Homeric Greek epics the *Iliad* and the *Odyssey*, the Old English *Beowulf*, the Old Irish *Táin Bó Cúailnge*, and the Classical Sanskrit epics the *Mahābhārata* and the *Rāmāyaṇa*.

And there are significant intellectual achievements that focus on Indo-European languages. For instance, relevant here is the development some 2,500 years ago by Indian grammarians, of whom Pāṇini (c. fifth century BCE) is the leading figure, of as sophisticated and complete a grammatical account of a language, in this case Sanskrit, as any that has been done since. The sheer brilliance, complexity, and elegance of this grammar, known as the *Aṣṭādhyāyī*, a collection of nearly 4,000 brief sutras that strictly and exhaustively describe the totality of Sanskrit grammar, put it, according to some observers, on a par with

⁹ There is of course no basis for talking about an official language in Antarctica, but there are Russian and American bases of operation there. And we have it on good authority that English is spoken there on a regular basis, since two Society fellows, Ellen Mosley-Thompson and Lonnie Thompson, and their teams use English there on their frequent expeditions to the South Pole.

Einstein's theory of relativity.¹⁰ A further achievement centered on an Indo-European language that deserves mention here is the decipherment in 1952 by Michael Ventris (with help from others, especially Alice Kober) of the Linear B syllabic writing system found on Crete and the Greek mainland, and the demonstration that it reveals an archaic form of Greek.¹¹

Finally, one can say that investigating any aspect of Indo-European, whether for purely linguistic or linguistically allied archaeological purposes, makes one a part of a long and glorious scholarly tradition—over 200 years old (dating from the 1816 publication of Franz Bopp's *Über das Conjugationssystem der Sanscritsprache in Vergleichung mit jenem der griechischen, lateinischen, persischen und germanischen Sprache*)—aimed at discovering facts about the past that would not be determinable without attention to linguistic evidence.

The Panelists and Their Articles

It is this past that the two panelists discuss in their papers that follow. These two scholars present two different lines of investigation into early movements of Indo-European peoples into Europe.

The first author is Andrew Garrett, Professor of Linguistics and the Nadine M. Tang and Bruce L. Smith Professor of Cross-Cultural Social Sciences at the University of California, Berkeley. He is an Indo-Europeanist, specializing particularly in Hittite and the ancient Anatolian languages, and in the past decade has begun serious research on the native languages of California, especially Yurok. Most recently he has been working on refining the methodology of evolutionary biology (systematics) for working out linguistic phylogenies (see, e.g., Chang et al. 2015), which is the reason he has been invited to contribute here. In his paper, Dr. Garrett reviews the general question of Indo-European phylogeny and the deep chronology of the family from various perspectives, including both more traditional qualitative analysis and more recent quantitative analysis based on the methodology of biological systematics.

10 As an index of the scale of Pāṇini's grammar, one can note that there are popular means to understand relativity (e.g., the "relativity for dummies" seen at <http://www.thecrazyhistoryofhistory.com/2012/09/the-theory-of-relativity-for-dummies.html>), but no equivalent "Pāṇini for dummies." And while David Spergel, in his excellent presentation on April 27 at the American Philosophical Society's Spring 2017 Meeting entitled "Our Simple and Strange Universe," gave a compelling two-minute lesson explaining relativity, a comparable two-minute explanation of Sanskrit grammar seems an impossibility.

11 On Ventris and the decipherment, see Chadwick (1967); on Alice Kober, see Fox (2013).

The second author is David Reich, Professor of Genetics at Harvard Medical School and an Investigator with the Howard Hughes Medical Institute. His research focuses on finding complex genetic patterns that cause susceptibility to common diseases among populations. But his involvement in this panel is the result of his work on deep history of humankind through an examination of DNA evidence, including a 2009 investigation of the origins of the population of India that has proven to be a landmark study (Reich et al. 2009). In his article, Dr. Reich reports on the results of sophisticated analysis of the DNA of ancient peoples from various regions around the world, and on the basis of this analysis, he concludes that the ancient DNA evidence is counter to the claims of Renfrew's out-of-Anatolia hypothesis.

CONCLUSION

In sum, it seems fair to say that the Indo-Europeanization of Europe is indeed a fascinating question about the past that not only can be explored but also should be explored. Moreover, it should be taken seriously by one and all today, as it has the greatest relevance to our present.

REFERENCES

- Chadwick, John. 1967. *The Decipherment of Linear B*. Second edition. Cambridge: Cambridge University Press.
- Chang, Will, Chundra Cathcart, David Hall, and Andrew Garrett. 2015. "Ancestry-constrained Phylogenetic Analysis Supports the Indo-European Steppe Hypothesis." *Language* 91: 194–244.
- Darwin, Charles. 1860. *On the Origin of Species by Means of Natural Selection, or, The Preservation of Favoured Races in the Struggle for Life*. London: J. Murray.
- Fox, Margalit. 2013. *The Riddle of the Labyrinth. The Quest to Crack an Ancient Code*. New York: HarperCollins.
- Gimbutas, Marija. 1956. *The Prehistory of Eastern Europe. Part I: Mesolithic, Neolithic and Copper Age Cultures in Russia and the Baltic Area*. Cambridge, MA: Peabody Museum.
- . 1970. "Proto-Indo-European Culture: The Kurgan Culture during the Fifth, Fourth, and Third Millennia B.C." In *Indo-European and Indo-Europeans: Papers Presented at the Third Indo-European Conference at the University of Pennsylvania*, edited by George Cardona, Henry M. Hoenigswald, and Alfred Senn, 155–97. Philadelphia: University of Pennsylvania Press.
- Gray, Russell D., and Quentin D. Atkinson. 2003. "Language-tree Divergence Times Support the Anatolian Theory of Indo-European Origin." *Nature* 426: 435–39.
- Holm, Hans J. (2007) 2008. "The Distribution of Data in Word Lists and Its Impact on the Subgrouping of Languages. Data Analysis, Machine Learning, and Applications." In *Proc. of the 31th Annual Conference of the German Classification Society (Gfkl), Univ. Freiburg, 2007*, edited by C. Preisach, H. Burkhardt, L. Schmidt-Thieme, and R. Decker. Heidelberg-Berlin: Springer.

- Lepsius, Richard. 1981. *Standard Alphabet for Reducing Unwritten Languages and Foreign Graphic Systems to a Uniform Orthography in European Letters* (2nd rev. ed. London, 1863). With a new introduction by J. Alan Kemp. Amsterdam Classics in Linguistics, 1800–1925, 5. Amsterdam: John Benjamins. First published 1863.
- Reich, David, K. Thangaraj, N. Patterson, A. L. Price, and L. Singh. 2009. “Reconstructing Indian Population History.” *Nature* 461: 489–94.
- Renfrew, Colin. 1987. *Archaeology and Language: The Puzzle of Indo-European Origins*. London: Cape.