



Research on the Homology of Indo-European Languages from the Perspective of Comparative Linguistics

Jiaqi Liu(✉)

School of International Studies, Zhejiang University, Hangzhou, China
liujq557@zju.edu.cn

Abstract. This paper gives an overview of the development of comparative linguistics and Indo-European language family, and analyzes some cognates in Indo-European language family from the aspects of meaning, pronunciation and inflectional form based on comparative linguistics. It is expected to have a deeper understanding of the overall structure of Indo-European languages, and to better explore the development and basis of Indo-European languages.

Keywords: homology · Indo-European languages · Comparative linguistics · cognates

1 Introduction

Indo-European languages, which refer to a system of languages in parts of India and Europe. As Indo-European languages with a large number of speakers and covering a large area, their origin and development have attracted many scholars.

By modern times, these branches of Indo-European languages, though seemingly different or even quite different, had many things in common. Some of them have certain pronunciations, certain words in certain fields, and the same particular grammatical forms or syntactic features in certain sentence patterns, and so on. This linguistic similarities inspired some scholars began to think about where their languages came from and to extrapolate the development from these linguistic similarities.

Based on comparative linguistics, this paper analyzes some cognates in the Indo-European language family from the aspects of meaning, pronunciation and inflectional form, in order to have a deeper understanding of the overall structure of the Indo-European language family.

2 Comparative Linguistics and Indo-European Languages

The beginning of comparative linguistics began with the discovery of Indo-European languages. In the late 18th century, an English scholar and diplomat named William Jones (1746–1794), while working in India, noticed that some of the words and grammatical

features of the popular Sanskrit language in India at that time had something in common with some of the words and grammatical features of modern European languages, such as Latin and Greek. As a result, he is credited with the first discovery of Indo-European languages, and Indo-European linguistics was born.

Ruhlen (1987) described the distribution of the main language branches of Indo-European language family and Dravidian language family [4]. Although languages in different regions have been greatly different due to the influence of time, language contact and other factors, the homology of vocabulary, pronunciation and grammar can still be found in these different Indo-European languages.

Linguistics can be regarded as the historical study of comparative linguistics, which plays an important role in the study of the origin and evolution of regional languages. “The central task of comparative historical linguistics is to identify genetically related language groups, to reconstruct their ancestral languages, and to trace the development of each of them” (Thomason & Kaufman 1988:15) [8]. Once the spread and dispersal of languages has begun, they diverge in a wide variety of ways.

3 The Cognates and Their Meanings

Cognates are words derived from common ancient roots in related languages, from which phonetic correspondences can be found and phonetic variations can be derived (Fromkin & Rodman 2007:480) [3]. Even if they do not form strict correspondence in phonetics, based on the common Indo-European background, some languages have similarities in morphology, meaning and other aspects in cognate words. The similarities of languages in each language family or language family or the differences with other languages can provide a basis for constructing the language source.

The analysis of homology of words is not only helpful to infer the origin and migration evolution of language, but also to make certain inferences about national habits and living environment.

As can be seen from the Table 1, in each group of words, words with the same color have more obvious similarities. For example, the word “snow” is mostly derived from Indo-European nouns and verbs with the meaning of “snow”, which is cognate with “snih-” in Sanskrit, and “snih-” itself means “sticky”. So in Lithuanian, Irish and Russian we can find the evolution of “snih-” and its similar form. The Greek “χίονι” is the word with the meaning of “winter”, while the Sanskrit “hima-” mainly means “snow, ice, winter”. The Latin “nix” comes from the verb “ninguit”, which also comes from Italian “neve”, French “neige” and Spanish “nieve”, meaning “snow”. Others have to do with “winter, cold”, and some have to do with specialisation of words such as “fall”, such as “snowfall, snow” in English.

As for the word “river”, there is a rich vocabulary to describe different grades of rivers. To be specific, “river”, “brook” and “stream” also have different words to describe various forms of running water. Therefore, the words representing the above are likely to be interchangeable. Most words for rivers come from the roots for “flow” and “run”, such as the Greek “sreu-”, the Sanskrit “sru-”, which means “flow”, and the Sanskrit “ri-” means “flow, run”. Others, like “ápas” in Sanskrit, are cognate of words that represent “water”, which might have originally been described as “flowing water”. Others come

Table 1. Comparison of specific words in some Indo-European languages

	Snow	River	Bear	Bee	Sea	Rice	Grass
Greek	χιόνι	ποτάμι	αρκούδα	μέλισσα	θάλασσα	ρύζι	χορτάρι
Latin	nix	fluvius	ursa	apes	mare	oryza	herba
Italian	neve	fiume	orso	ape	mare	riso	erba
French	neige	rivière	ours	abeille	mer	riz	herbe
Lithuanian	sniegas	upė	turėti	bitė	jūra	ryšai	žolė
Irish	sneachta	abha	béar	beach	fairrge	ris	féar
English	snow	river	bear	bee	sea	rice	grass
Russian	sneg	reka	nesti	pchela	more	ris	trava
Sanskrit	hima-	nadí-	—	bhrama- ra-, ali-	sāgarag	vrihi-	trna

(made by the author)

Table 2. Contrast verbs related to “see”

English	see	look (vb.)	sight (subj.)	sight (obj.), look (obj.), appearance	show (vb.)
Dutch	zien	aanzien	geziht	aanzien	toonen
Greek	βλέπω	βλέμμα	ὄψη	ὄψη	φαίνω
Latin	vidēre	aspicere	visus	aspectus	mōnstrāre
Italian	vedere	guardare	vista	vista	mostrare
French	voir	regarder	vue	vue	montrer
Spanish	ver	mirar	vista	vista	mostrar
Lithuanian	matyti	veizdeti	matymas	išvaizda	(pa)rodyti
Russian	videt'	smotret'	zrenie	vid	pokazat'
Sanskrit	drc-	drc-	drsti-	drc-	dic-

(made by the author)

from verbs for “rush”, “crash”, or “roar”, which may have originally been “a roaring stream”, like the Latin “rivus” etc.

The analysis of the homology of these words can also make certain inferences about the national living environment. Like most in the Indo-European language of “snow”, “river” and “bear”, “bee”, “sea”, “rice”, “grass” and other words be homologous, the probability of large, their living environment, or migration survival environment is on the way to have snow, rivers, there are bears and haunted by bees, grassland, to see the sea, to be able to grow and eat rice. Not only nouns, but also verbs can show the homology between languages (Table 2).

The Indo-European word “weid-” stands for “see” and “know”, which probably represents the connection between “see” and “know”, and “know” is a result of “see”. The Indo-European “okw-” always appears in words related to the word “eye”; And “show” can be seen as the equivalent of “see” and “to be seen”. In Greek, “φάινω” means “bring to light”, “cause to appear”, “appear”. In Sanskrit, “bhá-” is similar, meaning “shine”. Therefore, from the perspective of different forms of verbs, language homology can also be analyzed and interpreted to a certain extent.

4 Phonetic Homology of Indo-European Languages

The centum and satem categories, for example, differ from regular soft palatal sounds. The palatalized /k/ in Proto-Indo-European sounds more like /ky/ and is pronounced “kyuh” rather than “kuh”. In satem languages, this palatalization is a single phoneme, which then evolves into the satem fricative “suh” or “shuh”. Therefore, the word “k̑mtóm”, which means “hundred” in Proto-Indo-European, becomes “śatam” in Sanskrit, “šimtas” in Lithuanian, and “sto” in some Slavic languages.

In centum languages, the /k/ sound in the palatalized soft palate is integrated with the /k/ sound in the pure soft palate. Taking the Germanic languages as an example, the two /k/ sounds have evolved into /h/, so the word “k̑mtóm”, meaning “hundred” in Proto-Indo-European, becomes “centum” in Latin, “(he)katon” in Greek, “cant” in Welsh and “hund” in Old English. Generally speaking, “hundred” is thought to come from the Germanic branch, and “cent” is thought to be a Latin loan. These two words look very different, but we can extrapolate from the /k/ sound:

- (1) drop the suffix “-red” for numbers: hundred → hund
- (2) convert /h/ sound to /k/ sound: hund → kund
- (3) since “c” was originally pronounced like /k/ in Latin → kuh
- (4) /d/ is the voiced version of the alveolar plosive /t/

When combine the steps above and pronounce the two words, kund and cent, they sound surprisingly similar. Therefore, the English words “hundred” and “cent” are cognates, but they can’t be seen at first sight. It is because they have been quite different in the long evolution.

5 Cognates and Their Inflected Forms

Unlike Chinese, Proto-Indo-European is a highly inflected language, and almost all languages of the Indo-European family of languages are inflected to some extent and are highly inflected. Inflection refers to the use of conjugation to express grammatical functions such as number, person, tense, mood, case, etc. The most obvious example in Indo-European is the suffix of a verb, take the verb “to bear”.

From Table 3 we can see the different inflections of the verb “to bear” in different languages. English has retained second-person variations even in modern English, with “bearest” being a second-person variant. However, such inflections have been diluted or even disappeared in many languages so far. In modern English, the second person

Table 3. The cognate root of “to bear”

Language	to bear	Meaning
Proto-Indo-European	bherō/bheremi	I bear
	bheresi	You bear
	bhereti	He/she bears
Sanskrit	bharā-mi	I bear
	bhara-si	You bear
	bhara-ti	He/she bears
old English	ber-u/ber-o	I bear
	biri-s	You bear
	biri-th	He/she bears

(made by the author)

Table 4. The cognate pairs of noun cases in Proto-Indo-European and Sanskrit

Proto-Indo-European (8)			Sanskrit (8)		
nominative	Eg. (horse)	eḱwos	nominative	Eg. (horse)	aśvaḥ
vocative		eḱwe	vocative		aśva
objective		eḱwom	objective		aśvam
genitive		eḱwosyo	genitive		aśvasya
dative		eḱwōy	dative		aśvāya
ablativus		eḱwōd	ablativus		aśvad
adessive		eḱwōy	adessive		aśve
possessive		eḱwō	instrumental		aśvena

(made by the author)

inflections have been eliminated and only the third person inflections, namely, “he/she bears”, have been retained.

From the Table 4, there are eight cases in Proto-Indo-European nouns: nominative, vocative, objective, genitive, dative, ablativus, adessive and possessive. Sanskrit carries on all eight cases perfectly, as does Old Iranian and Avesta, Latin has six, Old English has four, but modern English and even other Indo-European languages like most of the Romance languages have no noun cases. The absence of case change in this language leads to the need for the function of nouns to be determined by prepositions and word order. Of course, many Indo-European languages still have cases, for example, there are six to eight cases in the Polo-Slavic languages; Greek, German, and Icelandic also have four cases.

6 Conclusion

Comparative linguistics has several defects, one of which is that it fails to solve the key problem of dating (Cheng & Liu 2022) [1]. In the middle of the 20th century, Swadesh (1952:452–463;1955:121–137) successively put forward “lexicostatistics” and its improved version “glottochronology” [6, 7]. But since the beginning of the 21st century, some linguists have switched to phylogenetic methods used in evolutionary biology, especially Bayesian phylogenetic methods, like Sagart et al. (2019:10317) [5], Zhang et al. (2019:112–115) [9].

Based on the above, we can see that most languages in Indo-European language family not only have common root, affix or word meaning in lexical form, but also have certain homology and inheritance in case features. Although these features are not obvious in some languages, a closer look reveals that the Indo-European languages share a wide range of homologies. Just like Friedlaender (2009:464) said, linguistics is an effective tool in the study of the origin and distribution of peoples [2]. If we can compare and analyze their cognates and their characteristics of meaning change, morphological change, word order change and phonetic change, we can have a further understanding of their cognates.

References

1. Cheng G., Liu J.Q. “Asian Civilization Research from the Perspective of the Linguistics,” *Journal of Zhejiang University (Humanities and Social Sciences)*, Vol. 52, No. 2 (2022): 87–99.
2. Friedlaender J., Hunley K., Dunn M., Terrill A., Lindström E., Reesink G. & Friedlaender F., “Linguistics more robust than genetics,” *Science*, Vol. 324, No. 5926(2009), pp. 464–465.
3. Fromkin, V., Rodman, R. 2007. *An Introduction to Language* (8th edition). Beijing: Peking University Press.
4. Ruhlen, M. 1987. *A Guide to World’s Languages*. Publisher: Arnold.
5. Sagarta L., Jacquesa G., Lai Y.F., Ryder R.J., Thouzeau V., Greenhill S.J. & List J.M., “Dated language phylogenies shed light on the ancestry of Sino-Tibetan,” *PNAS*, Vol. 116, No.2(2019), pp. 10317–10322.
6. Swadesh M., “Lexico-statistical dating of prehistoric ethnic contacts,” *Proceedings of the American Philosophical Society*, Vol. 96, No. 4(1952), pp. 452–463.
7. Swadesh M., “Towards greater accuracy in lexicostatistic dating,” *International Journal of American Linguistics*, Vol. 21, No. 2(1955), pp. 121–137.
8. Thomason, S. G., Kaufman, T. 1988. *Language Contact, Creolization, and Genetic Linguistics*. University of California Press.
9. Zhang M.H., Yan S., Pan W.Y. & Jin L., “Phylogenetic evidence for Sino-Tibetan origin in northern China in the Late Neolithic,” *Nature*. Vol.569, No.7754(2019), pp. 112–115.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

