In the Root of the Indo-European Language Tree

Messele Zewdie Ejeta

Abstract

This work presents various collections of likely cognates from the Oromo language and the Indo-European language family and opens a new frontier of research in the study of the root of the Indo-European language tree and the origins of human languages. It presumes that those words that have been in use more frequently to express various concepts and physical matters may have been, of necessity, likely among the earliest words to have been formed, more commonly used since their formations, and hence less forgotten along the way of humanity's long journey. It points to a new and unique opportunity for a focused study on the origin of human languages.

Introduction

Our collective quest for a better understanding of the origin and evolution of human languages has attracted considerable efforts in the past and continues on today (Dominguez and Rakic, 2009). In Friedrich Nietzsche's work, the significance of language for the evolution of culture is suggested to lie in the understanding of language as a separate world beside the other world (Nietzsche, 1996). The innate desires and needs by early humans to express their observations of the environment in which they lived and thrived as intelligent beings may have been the drivers for the genesis of the use of words, phrases, and sentences. This thesis presumes that the availability of the environment in which early humans lived preceded the desires and needs to use words, phrases, and sentences to express their observations in that environment. The continuity of this use in the form of oral and written communications as well as cultures may be considered our inheritance of their innate desires and needs for the same purposes of living and thriving in the environment in which we find ourselves. Parallel to biological information provided by genes, human languages provide vital clues to human history (Pagel, 2000; Gray and Jordan, 2000). They are also conceived as cultural replicators with behavior and fidelity that can rival that of genes (Pagel, 2009). Some of the extensively researched and established language families are the Indo-European languages (Gray and Atkinson, 2003). The base of these languages is a presumed dead language known as the proto Indo-European.

Through the study of phonological changes, among others, linguists reconstructed the Indo-European language family tree from a presumed dead protolanguage Indo-European, which is believed to have been spoken some 10,000 years ago (Fitch, 2007). These reconstructions have inspired Charles Darwin's biological evolution (Fitch, 2007). It can be also argued that the evolution and history of human languages provide important clues to that part of early history of humanity and the world, which have not been understood well enough yet but are being approached through various scientific disciplines, including DNA analysis, archeology, anthropology, and paleontology. In addition to enriching our knowledge of the history of early humans, getting to the root of the Indo-European language tree, unraveling what languages and linguistic structures lie behind it, and making a scientific push along this research frontier may possibly lead to uncovering some clues towards the studies of the origin of human languages. In effect, this work takes a step towards getting new insights about the root of the Indo-European language tree and points to a new frontier in the study of the origin of human languages, a subject of an ongoing investigation by this author.

By presenting various collections of cognates, words with similar forms and related meanings, to their counterparts in some of the Indo-European languages including English, the findings in the current work bring in a new insight to bear on the presumption that proto Indo-European is a dead language. These collections of words are obtained from the Oromo language, which is used today by millions in East Africa, mainly in Ethiopia, and the Amharic language, Ethiopia's official language. In addition to bringing this new insight to the root of the Indo-European language tree, this pool of words has led to a focused ongoing research by this author on the origins of human languages.

Methodology

Some of the various means linguists and evolutionary theorists have used to bridge the gap between lexicons in different environmental settings are the studies of cognates and glossogeny. Lieberman, et al. (2007) used the frequency of word use to quantify the pattern of language change. Different criteria were used in the current work to determine whether a word in the Oromo language is a cognate of an English word or its cognates thereof in the Indo-European language family. The first criterion is that the Oromo word should have the same or a closely similar meaning to its English counterpart. The second criterion is that the two words, or its cognate

thereof, should have in common at least 50% of the consonants used in the cognates. Third, these cognates should have similar vowels that are used to pronounce the consonants. Fourth, the Oromo word is a closer cognate to an older version of Indo-European language than its counterpart in the English language. Fifth, due to the evolutionary history of words, certain consonants may be replaceable with others, such as *d* with *t*, *b* with *v*, *f* with *p*, or vice versa.

For an instance of consonant sound exchangeability, take the word *lubu*, an Oromo word that means soul. Its close cognates in the Latin, Gothic, Old English, Middle English, and English are *lubēre*, *lubō*, *lufu*, *lufu*, and *love*, respectively. While most of these cognates use *lu-* sound in them, according to these words, the older languages have seemingly closer cognates to the Oromo language word *lubu* than the English word *love*. Such an evolutionary departure shows divergences both in meaning and word spelling.

I start the comparison of various words commonly used in the Oromo language from the word *eye*, which is then used as the bridge and window from the Oromo and Amharic languages to the Indo-European and other international languages (see Table 1).

Table 1. Possible cognates of the word *eye* as a potential bridge between various international languages

No.	Term	Language	No.	Term	Language
1	Acs	Latvian	17	Occhio	Italian
2	Auga	Icelandic	18	Ochi	Romanian
3	Auge	German	19	Oci	Slovenian
4	Ayin	Amharic	20	Oeil	French
5	Begi	Basque	21	Oga	Swedish
6	Ege (Eage)	Old English	22	Ojo	Spanish
7	Eie (Ie)	Medieval English	23	Olho	Portuguese
8	Eye	Danish	24	Oko	Serbian
9	Eye	English	25	Oko	Slovak
10	Ghajn	Maltese	26	Oko	Polish
11	Glaz	Russian	27	Ollo	Galician
12	Goz	Turkish	28	Oog	Dutch
13	Ija	Afan Oromo	29	‘Oog	Afrikaans
14	I’y	Hebrew	30	Oye	Norwegian
15	Je	Haitian	31	Oyen	Ukrainian
16	Jicho	Swahili	32	Yan	Chinese
			33	‘Yn	Arabic

I then present comparisons of over sixty other words collected from the Oromo language with their likely cognate counterparts in the various Indo-European languages. These cognate words that are compared are categorized into various groups of use including for and related to 1) human anatomy, 2) human emotion including faith and conflict, 3) living necessity, 4) living environment, and 5) time concept. These categories are selected with the presumption that those words that have been in use more frequently to express these categories of uses may have been likely, of necessity, 1) among the earliest words to have been formed, 2) more commonly used since their formations, and hence 3) less forgotten along the way of humanity's long journey.

The basic reason to start the comparison with the word *eye* is its exemplary nature to the above three factors used for selecting the words I compared. As a sensory organ, the physical human organ eye provides a virtual bridge between the physical world and the complex thought processes that take place in the brain. In other words, this word is considered to be particularly significant for this study because the sensory organ that this word describes is considered vital in forming and maintaining the virtual bridge between the external physical world and the internalized information processing system. In its functional state, there is a one way steady flow of information over this virtual bridge, which is then processed by the brain. The processed information is used to take spontaneous actions, stored as memory, or permanently discharged by way of forgetfulness. The functionality of this sensory organ doesn't seem to vary drastically among humans in different natural environmental settings or the physical world. In essence, under normal conditions, the mobility of this sensory organ is comparable to that of its carrier in a conscious functional state. Therefore, this study hypothesizes that this sensory organ is implicitly the most frequently used in its class for the functioning of the memory apparatus, thus lending itself to the corollary that the word that describes it is likely to be very resistant to evolutionary changes. Furthermore, this study hypothesizes that a word that stands for a matter that is observed by the eye or engages the mind more frequently has less propensity for evolutionary changes.

On the other hand, there are words that are associated with various physical and perceptive matters that are stored and passed down from one generation to the next generation. These words may not necessarily have the same level of mobility as the word *eye* or the sensory organ described using this word. For instance, the word *rock* refers to the same physical matter irrespective of the natural environmental setting in

which it is found. However, the utility of this term for the purpose of describing or conveying information about this particular physical matter in different environmental settings, or languages, goes only as far as the ability and propensity of the user to adapt an equivalent term for the same matter in the lexicons used in these different environmental settings. Thus the qualitative difference between what the two words, *eye* and *rock*, describe can be characterized as pseudo permanently mobile (PPM) and willing permanently mobile (WPM), respectively. The differences in these attributes may be due to two possible reasons: 1) the differences in the spatial and temporal origins of these attributes and 2) the memory losses in the spatial and temporal mobility of the information processing system that is used to store and transmit these attributes. Irrespective of the sources of these differences, the state of human languages in our world doesn't seem to be readily flexible to adaptive utility simply because of our propensity to retain, perhaps subconsciously, what we inherited from our ancestors. It is the understanding of such salient features of human languages that will likely lead to their systematic further mapping and using that as a springboard to the study of the origins of human languages.

Collected Likely Cognates

After focusing specifically on the cognate structure of the word *eye* among the Oromo, Amharic, Indo-European, and other international languages, this study sought the various collections of cognates that are observed by the eye or engage the mind more frequently for each of the five categories of use described above. Summaries of the collected words are presented in Table 2 to Table 6 for these five different categories of word uses.

Table 2. A collection of cognates for or related to human anatomy (partial sources: Gamta, 2004 and www.dictionary.com)

Oromo	Amharic (Greek)	Latin	Old Norse (Indo Euro.)	German (Goth)	Old English	Middle English	English
ija	**ayin**			**auge**	**ēge, ēage**	**eie, ie**	**eye**
hadda					hēafod, hafod	he(v)ed	head
hari					hǣr	haire, heer	hair
luka			leggr			leggr	leg
gura		auris	eyra		ēare	ere	ear
fana	fana				fōt	fōt	foot
huddu					buttuc	buttok	buttock
mata		mēta					meta
kintiri		clītoris					clitoris
fuchi		vāgīna					vagina

Table 3. A collection of cognates for or related to human emotion including faith traditions and conflict (partial sources: Gamta, 2004 and www.dictionary.com)

Oromo	Amharic (Greek)	Latin	Old Norse (IE)	German (Goth)	Old English	Middle English	English
lubu		lubēre		(lubō)	lufu	lufu	love
du'e			deyja			dien, deien	die
kafana	kafan (kóphinos)	cophinus				cofin	coffin
dhara		errōr				errour	error
haramu	haram						harem
gadhe					gāhi	gai	gay
milki				(glūck)		luc, lucke	luck
kolf(e)an			hlǣa	(hlahjan)	hlæh(h)an	laughen	laugh
bada					bæddel	badde	bad
gad(e)an	(hádēn)	satis		saths	sæd	sæd	sad
gali					gǣan	gol	goal
dula		duell					duel
raya		reg(ere)	(reg)				regiment
werara	warara				werre	werre	war
tarre	tarta				ārǣdan	arrayen	array
eye					gēse	yes, yis	yes
	kineliq		knouleikr		knoulāc	knoulec he	knowledg e
holi					hālig	holi	holy
gofta	gheta			got, guth			god
alaba			flap			flagge	flag
kora(saba)		congressus				congres se	congress

Table 4. A collection of cognates for or related to living necessity (partial sources: Gamta, 2004 and www.dictionary.com)

Oromo	Amharic (Greek)	Latin	Old Norse (Indo Euro.)	German (Gothic)	Old English	Middle English	English
buddena				(fōdjan)	fōda, fēdan	fode	food
kita	kita						pizza, pitta
bedde				badi	bedd	bedd	bed
farso					bēor, bjōrr	bere	beer
dabala		duplus				duplus	double
dughe(an)			drekka	drinkan	drincan	drinken	drink
kuti	kurat		kuti		cytten	cutten	cut

Table 5. A collection of cognates for or related to human environment (partial sources: Gamta, 2004 and www.dictionary.com)

Oromo	Amharic (Greek)	Latin	Old Norse	German (Goth)	Old English	Middle English	English
ishi	isua				hēo, hīo	sēo, sīo, sīe	she
nama	(ōnoma)	nōmen	nafn	name (namō)	nama	nama	name
damme							madame
laga	(lākkos)	lacus			lagu	lak(e)	lake
lencha					lēo	leon	lion
walda				wild	wilde	wilde	wild
fora					feorr	far, fer	far
Sorsa						Sīrius	Sirius
kamadi					hwǣte, kweit	whete	wheat
nughi	nug						negro
kello					geolo	yelou	yellow
hadi			hvītr	(hweits)	hwīt	whit(e)	white
farda	faras			pferd			horse
komme			koma		cuman	comen	come
boro			borg		burg	borogh	borough
ummata		commūnitās				comunete	community
soke			sœkja	suchen	sēcan	seken	seek
sacha'e					serchier	serchen	search
kilensa	(klīmat)		(klei)			climat	climate
dhaka					rocc	rokk(e)	rock
	ayer	āēr				eir	air

Table 6. A collection of cognates for or related to time concept (partial source: Gamta, 2004 and www.dictionary.com)

Oromo	Amharic (Greek)	Latin	Old Norse (Indo.Euro.)	German (Gothic)	Old English	Middle English	English
yero			ār	jahr	gēar	yeer	year
bari(te)	barra		bjartr	(bairht)	breht, beorht	breht, beorht	bright
dukana					deorc	derk	dark
bara		æra			ār		era
wache			vaka	(wakan)	wacan (wōc)	waken	wake
fuldur(a)		fūtūus				futur	future
dur		dūrāre				durer	during

Discussion

These various collections of cognates from the Oromo, Indo-European, and other international languages appear to present a new insight into the early history of languages, cultures, and humanity. This data is likely to reopen the view that Indo-European is a dead protolanguage. It is presumed that to the extent that cognates for the same physical matter or conceptual expression exist among different languages, they show the possibility of these cognates being of the same origin and divergences of word usage over time or the resistance thereof. These findings are in agreement with the results of DNA analyses of samples of people around the world, which indicate that humanity migrated out of East Africa tens of thousands of years ago (Zhong, et al., 2010; Li, et al., 2008; Jakobsson, et al., 2008; Liu, et al., 2006). Thus, the pool of cognates provided in this work are likely to present vital clues about the presence of unbroken virtual bridges in chaotic sound bites in various languages of the world, including those outside the Indo-European language family. This significance of this work can be seen on various levels. It gives clues about how much social consciousness and adaptation to the natural environment, including plant and animal domestications, had been established before early humans started to migrate out of Africa. These clues present us some leads to the unanswered question of whether indigenous crops developed in Ethiopia or those presumed to have arrived from the Fertile Crescent launched Ethiopia's early food production (Diamong, 2005). Two possible hints for this lead are the likely changes from *nughi* to *negro* and *kamadi* to *wheat* (see Table 5). The possibility of social organization consciousness prior to the

migration out of Africa cannot be ruled out. If the word *flag* proves to be a cognate of *alaba* (Table 3), both of which have similar meanings, it is likely to explain the striking similarities of the tricolor flags of many nations around the world. If we consider a possible change over time and space of blue for black, we would readily observe that many nations around the world use the tricolors of black (blue), red, and white for their flags. For the Oromo, the tricolors of black, red, and white symbolize deep meanings that encompass the past, the present, and the future laced with faith traditions. Granted that these collections of cognates support our established knowledge through the study of our DNAs, they will not only narrow down the domain space for the study of the origin of human languages, but also highlight the ongoing progress of humanity's enlightenment.

References

Diamond, J. (2005). Guns, Germs, and Steel: The Fates of Human Societies. W. W. Norton and Company. New York, London.

Online Dictionary, www.dictionary.com (last accessed July 14, 2010).

Dominguez, M. H. and Rakic, P. (2009). Language evolution: The importance of being human, *Nature* 462, 169-170 (12 November 2009), doi:10.1038/462169a.

Fitch, W. T. (2007). Linguistics: An invisible hand. *Nature*, 665-667.

Gamta, Tilahun (2004), Comprehensive Oromo-English Dictionary, Karrayyuu Publishing, New York.

Gray, R. D. and Atkinson, Q. D. (2003). Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* , 435-439.

Gray, R. D. and Jordan, F. M. (2000). Language trees support the express-train sequence of Austronesian. *Nature* , 1052-1055.

Jakobsson, M., Scholz, S. W., Scheet, P., Gibbs, J. R., VanLiere, J. M., Fung, H., Szpiech, Z. A., Degnan, J. H., Wang, K., Guerreiro, R., Bras, J. M., Schymick, J. C., Hernandez, D. G., Traynor, B. J., Simon-Sanchez, J., Matarin, M., Britton, A., van de Leemput, J., Rafferty, I., Bucan, M., Cann, H. M., Hardy, J. A., Rosenberg, N. A., and Singleton, A. B. (2008), Genotype, haplotype and copy-number variation in worldwide human populations. *Nature*, 998-1003

Li, J. Z., Absher, D. M., Tang, H., Southwick, A. M., Casto, A. M., Ramachandran, S., Cann, H. M., Barsh, G. S., Feldman, M., Cavalli-Sforza, L. L., Myers, R. M. (2008). Worldwide Human Relationships Inferred from Genome-Wide Patterns of Variation. *Science*, 1100-1104.

Lieberman, E. Michel, J.-B., Jackson, J., Tang, T., & Nowak, M. A. (2007). Quantifying the evolutionary dynamics of language. *Nature* , 713-716.

Liu, H., Prugnolle, F., Manica, A. and Balloux, B. (2006), A Geographically Explicit Genetic Model of Worldwide Human-Settlement History, The American Journal of Human Genetics, 230-237.

Pagel, M. (2009). Human language as a culturally transmitted replicator. *Nature Reviews Genetics*. 405-415.

Pagel, M. (2000). In C. Renfrew, A. McMahon, & L. Trask, *Time Depth in Historical Linguistics* (pp. 189-2007). Cambridge, UK: The McDonald Institute for Archaeological Research.

Zhong, Hua, Hong Shi, Xue-Bin Qi, Chun-Jie Xiao, Li Jin, Runlin Z Ma and Bing Su (2010), Global distribution of Y-chromosome haplogroup C reveals the prehistoric migration routes of African exodus and early settlement in East Asia, *Journal of Human Genetics*.