

# Ôèëíñáíåòèéà â èëíñáèñòèéå: ýçûéè, äåðåâüy, ýâïëþöèý

Àâòíð Äèáíà lâäääàñáíà íóðáàéñáà

19.08.2012 á.

Íñéáñáíà lâñéáííá 27.08.2012 á.

## Summary. Languages

evolve through time. But the past of languages could be untangled like the species are put in the tree of life. Just as DNA is used in biology to uncover the relationships between species, the most stable linguistic features (e.g. the basic vocabulary, etc.) seem to identify the phylogeny of languages. Here, we evaluate the phylogenetic relationships between Slavic languages using a group of methods. Our results correlate with the traditional classification and prove the stability of the chosen methods.

It might be that some very ancient language had altered little,

and had given rise to few new languages,

whilst others [&hellip;] had altered much, and had given rise

to many new languages and dialects.

Ch. Darwin &lsquo;The Origin of Species&rsquo;, 1859

Бçûé íæíí èç áàæíáéøé òåñíñáííá

ðàçâéòèý ðåéñáéà, êòòíðûé íáiyáñöý áìåñòá ñ íèí íà íðòíýæåíèé áñåé èñòíðèé ðåéñáéà. Òí, íáñéíéüéí áæéåéíí á íðòíýæåíèé ýçûéñá íú ííæåí çàäéýíóðü, áí íííáíííí çàäéñèò íò ñéíðíñòé èçíáíáéý ýçûéñá íú ãäééíèò. Íáéíðûá ýéáíáíóðü ýçûéñáíé ñòðóéòòðû ýâéýþöñý áíñòàòí-íí óñóíé-èéñùíè è ííáóð ñéóæèòü ñâíáíí ðíäà áåíáíè, ííñéáñáíàòåéüñíñòýé ÁÍÉ, íà íñííáá ñðåááíáéý èòòíðûé ííæííííñòðíèòü áåðåááí ýâïëþöèé ðééíñáíåòé-áñéáíá áäðåááí.

Ôèëíñáíåòè-áñéíá

äðåááí áðåáó, íòðåæàþùéé ýâïëþöèþ ðàçéè-íûð áéäáíá èééé áðóäéø ñóúñíñòáé, èíåþùéò íáùåáí íðåäéà. Ðàçéè-áþò òðé èéäññá áåððééí ðéëíñáíåòè-áñéíáíí áåðåááá: èéñòüý, óçéú è éíðåáí. Èéñòüý ýòí éííá-íûð (éííóðåáñá) áåððééíû, òí áñòü òá, éíòíðûí èíòéááíòíí òíéüéí íáíí ðåáðí. Èáæäáüé èéñò íòíáðåæåáá ìáéíðûé áéäá èééé ñòúñíñòðü, íáéíðéáð, ñíâðåáííñüé ýçûé. Èáæäáüé óçáéé íðåñòàáéýáð ýâïëþöèéííñá ñíáúòéå: ðàçääéáíéá íðåäéíáíí áéäáá íà áááá èééé áíéåá, éíòíðûá á áæéüíáéøáíí áíéïëþöèéííñòðíàéé íáçäáéñéí. Êíðåáíû áûäééáííáý áåððééíá áåðåááá, íòíáðåæàþùáý íáùåáí íðåäéà áñåð ðåññíáòðéááíñòðü ñóúñíñòáé. Ð, áðå òéëíñáíåòè-áñéíáí áåðåááá íðéíýóí íáçüááðû «áåðåáýé». Áçäéííá ðàññíéíáíéá áåðåááé íáçüáááñòðü ñóàééé íðéíáýóñý íáðåááòé-áñééáá ááðåááá.

Âûýâéáíéá

ôèëíñáíåòè-áñééø íòíðåáíéé ýçûéñá íáíá èç ñòíðíí íðåäéáòá èññéåäááíéý ñðåááíéòåéüí-éñòíðè-áñéíáíí ýçûéñíáíéý. Íáíáéí á ííñéááíá áðåáíý, íííéíí ðåðåæéòííñüð èéíñáéñòé-áñééø ñòàééé íðéíáýóñý íáðåááòé-áñééáá ááðåááá.

Ðàçðàáíòàííûå èçíà÷àëüíí äëý ðåðáíèý çàääà÷ ýáîëþöèíííí è ííëåéóëýðííé áèïéíäéè. Íäíàéíí ñëåäöåóå ìòíåòèòú, -òí íðëëíæåííà åàííûò íàòíäíâ èëéíååèñòèéå íåëäåäåò ñåíèòé íñíååííñòýìé, ñâýçáííûè ñ íðåäíåòííé íáëæñòþ. Ñòòèò òàéæåå ãííéòú 1 òíí, -òí éþáíà ååðååíí íðåäñòååéýåò ñíáíé èëøú íáíó èç äèíòåç áçàéííòíøåíéé íåæäöå òàéñííàé, íñíéíëüéó åñå íäåééè èëøú óíðíû, íííá íðåäñòååéåíéå ðåàæüíûò íðòðåññíå, èíäþùèò ååñüíà ñëíæíóþ ñòðóéòóðó è íðèðíäó.

Nóùåñòåóþò ðàçëè÷íûå íåòíäû ðèéíååíåòè÷åñéé ðåáéíñòðóéöèé, èíòíðûå ååéýñòý íá ååå åððííû: äëñòåíøéíííí-íàòðè÷íûå íåòíäû è ñòòðåòèñòé÷åñééå íåòíäû. Íåðååý åððííà íåòíäû íñòðåííà íà ðåñ-åòåå íàòðèòû ðåññòýíéé, íðè ýòíí ðåññòýíéå íííéíàòñý èåéé íåðå ðàçëè÷éý. Íðèíåðåàíè åèñòåíøéíííí-íåòðè÷íûò íåòíäû íñòðåéòü: íåòíä íåçåååðåíííí íííà ñòðåäíååíí (UPGMA)

[Michener  
and

Sokal, 1958], íåòíä ñâýçûååíèý áéèæåéøèõ ñíñåååé (NJ) [Nei and Saitou, 1987], íåòíä íñòðåííèý ðèéíååíåòè÷åñéé ñåòè Neigh 2004]. Ñòåòèñòé÷åñééå (éee èèñéðåòíûå) íåòíäû ðåáíòåþò íåíñòðåäñòååíí ñ íñéååíååòåéüíñòýìé åàííûô, à íå ñ êíýôòèøéåíòàíè èò ñòíñòååà, è ðåøåþò çàääà÷ó íðèíèçåöèé. Ñòååéè ñòåòèñòé÷åñééò íåòíäû ííæíí íàçååòü íåòíä íåéñèíæüííé ýéíííéé (MP)

[Tassy and Darlu, 1993],

íåòíä íåéñèíæüíííí íðååäííííåéý, íåòíä Áåéåñà [Holden et al., 2005].

Çàääà÷à ååíííí åèññéååíååíèý çàéëþ÷åéåñü  
å ñòååíåíèé ðèéíååíåòè÷åñééò ååðååüååå ñëååýíñééò ýçûéíâ, íñòðåííûò ñ èñííéüçíååíèåí ñëååóþùèò íåòíäû ðèéíååíåòè÷åñéé:

éåðåððè÷åñéåý

ééåñòåðèçàöèý íåòíäíííí åçååøåíííí ñòðåäíååíí (WGMA) íà íñííåå åêíðååéýøéíííí ðåññòýíéý, ðååéèçíååííàý íàíè á íåéåòå Wolfram Mathematica 7.0;

íñòðåííèå

ðèéíååíåòè÷åñéé ñåòè íåòíäíííí NeighbourNet, ðååéèçíååííûò á íðåðåííàíííí SplitsTree [Huson and Bryant, 2006].

éåðåððè÷åñéåý

ééåñòåðèçàöèý íà íñííåå ðåññòýíéý Ëåååíøòååéíà [Ëåååíøòååéí, 1965] è  
Áàíåðåò&ndash;Ëåååíøòååéíà [Chakrabarti, 2003], ðååéèçíååííàý á íåéåòå Wolfram Mathematica 7.0.

íñòðåííèå

ðèéíååíåòè÷åñééíííí ååðåååå íåòíäíííí Áåéåñà ñ èñííéüçíååíèåí íðåðåíííí íðåóéòå MrBayes [MrBayes].

Å ååíííí èññéååíååíèé ðåññíàòðèååþòñý

åååíååååòðú ñíåðåíåííûò ñëååýíñééò ýçûéíâ: ñéíååíñééé, íéæíâ- è ååððíåéóæéöééé,  
÷åññééé, ñéíååöééé, óéðåéíñééé, ååéíðóññééé, ííéüñééé, ðóññééé, íåéååíñééé,  
åíéååååñééé è ñåðååñòååòñééé. Ñíåéåñíí òðåäéøéíííí èéåññèéåæöè ñëååýíñééò ýçûéíâ, åûåååéýþòñý ñëååóþùèå åððííû:

1. ãíñòí-ñëååýíñéåý:

ðóññééé, ååéíðóññééé, óéðåéíñééé, ðóñèíñééé;

2. þæíí-ñëååýíñéåý,

êîòîðàÿ â ñâîþ î÷åðåäü äåéèòñÿ íà:

çàïàäíóþ

Íñääðóííó: ñåðáîõíðâàòñèé è ñëîâåíñèé;

âñòî÷íóþ

ííäääðóïïö: ìàéåäïíñêèé, áïëääðñêèé, öåðêïâíî-ñëàâýíñêèé;

3. çäïàäíî-ñëàâýíñêàý,

êîòîðàÿ îäðàçääåëÿåöny íà:

ëåõèòñêèå ÿçûêè: îïüñêèé,

êàøóáñêèé;

cautious,

–åñîñëîâàöêå ÿcûêè·

÷åøñêèé, ñëîâàöêèé,

Iàòåðéèàëìí äëÿ èññéëåâîâàíèý iññéóæèëè  
 èäéñè-åññéèå åâèéïèöü, ñíäåðæàùèåñý à ñíèñéàò Ñâîâåðøà ñíâðåìåííûò ñëàâýíññéè  
 ýçûñéâ, è íåéïòìðûå åðàììàòè-åññéèå iðecíàéè (ñíðåìåííèå äâíéñòåâíííâ ð-èññéà,  
 iàèè-÷éå ðòðåò ðíáíâ, íåèè-÷éå iñðåâåéäíííâ ðòðòéèéý, ñóðôðéññàòèý èâà íåèàééà ÷-àñòòðíàý  
 òîðìà ñëàâííåðåçíâàíèý, èññéüçíâàíèå ñíòíèåðòèåííûò òîðì äëÿ åûðåàæåíèý åèàâíðåìåííûò  
 cíà-åíèé äëàäàíèà). Èñòí-íèéëí ääéñè-åññéèò åäííûò äëÿ 12 ñëàâýíññéèò ýçûñéâ iññéóæèëà  
 áàçá ääííûò, ñíäåìåíàÿ Ëðàññéàëìí è äð. [Dyen et al., 1992],  
 ååå ñëàâíòòðüù éèàññéòðóþòñý íà èíäíàòû (íäííèòðåííûà ñëíâà, èíäþùèå íàùåå  
 iðèññòíæåâåíèå è iñòíæåå çâðò-÷-àíèå à åâóò è áíèåå ñàíññòíþòåëüíûò ýçûñéà),  
 ñíííèòðåëüíûà èíäíàòû è “; íà èíäíàòû”; Á ñéó-÷-å ïðèíàíåíèý íàòíà NeighbourNet è  
 Áåéåññà ñëàâíòòðüù èåææåíí èç àíàéèçèðóåííûò ýçûñéâ áúëè çàéíàëðíâàíû è  
 iðåâíåðåçíâåíû à áéíàòðíûå öäíí-÷-éé, èåææåûé yéàâåíò èíòòðûò ñíòðåâòñòåâàæ  
 iðecíàéò íàèè-÷éý/iññóòñòåòèý ñíòðåâòñòåòþùåé èíäíàòû (yéàâåíò èíàéèðíâàéñý “; 1”;  
 èèè-“; 0”; ñíòðåâòñòåâíí) à íííèòðåòíí ýçûñéâ. Òàéèò íàðåçíí, äëÿ ñíèñéà Ñâîâåðøà  
 èç 200 cíà-åíèé ÷-èññéí èíäíàò ñíñòåâèëè 476. Aíàéíæ-÷íû íàðåçíí áúëè çàéíæèðíâàíû  
 åðåììàòè-åññéèå iðecíàéè.

Ííëó÷åííûå Õåçóëüòàòû áî íííäíì ñõîæè.

èñòðè÷áñéíáí ðàçâèòèý íàðíáíá, áíâíðyùèõ íà ýòèõ ýçûéàõ, òàé è áéèýíéý äðóäèõ ýçûéíá. xòí éàñàåòñý áíñòí÷íñéèõ ýçûéíá, iíæíí çàìåòèòü, ÷òí èõ ðàçääéáíéå íðíèçíøëí áíñòàòí÷í ðàíí, ÷òí ìòðàæàåòñý á ìòääéáíèé ðóññéíáí ýçûéá, iíäååðååíñý áéèýíéþ ðeíí-óäíðñéèõ ýçûéíá è ííááíðíáñéíáí áéàéáéòå. Iòíàòèí, ÷òí áíááéáíéå áðàííàòè÷áñéèõ íðíèçíàéíá èá÷áñòåíí íá áéèýáò íá òíííéíáéþ óeëíáíáòè÷áñéíáí ååðååá, iíñòðíáíííá íáòíáíí Áàéáñà, è óeëíáíáòè÷áñéíé ñåòè, iíñòðíáíííé íáòíáíí NeighbourNet.

Â õíäå èññéááíáíéý áúéà èñííéüçíáíá  
öåéáý áðóííáíá ìåòíáíá òèëíáíåòè÷áñéíé ðåéííñòðóéòè, íáíàéí áúéè ííéó÷åíú  
êà÷áñòåííí íáíðíòéåíòå÷éáúá ðåçóëüòàòú. Ááííúé òàéò ñåéäåòåëüñòåóå Íá óñòíé÷éåíñòé  
è áääéåàòííñòé íðèíáíýáíúò íáòíáíá áéý ðåøåíéý ííñòååéáííé çäääà÷é.

Â õíäå áäéüíáéøèõ èññéááíáíéé  
íéàíéðóåòñý ñòåáíéòåéüííá èçó÷áíéå ñëåáýíñéèõ ýçûéíá ñ áðóäèíé áðóííàéë ýçûéíá,  
ðàñííñòðåííúò Íá ðåðèòííðèé Ðíññéè, á ÷àñòííñòé òeíí-óäíðñéíé è òþðéñéíé  
áðóííàéë. Ðàñøéðåíéå íáðå÷íý áíáéëçèòåíúò ýçûéíá ñåýçàíí ñ òåí, ÷òí áàííúá  
áðóííü ýçûéíá ðàñííñòðåííúò íá òåððèòíðèé íáííé ñòðàíú è ðàçâèåàéëñü á  
óñéáéýö áçàéíííá áéèýíéý.

## Niñéñé èèòåðåòóðû

Chakrabarti, 2003 Chakrabarti S. Mining the Web: Discovering Knowledge from Hypertext Data. Morgan Kaufmann Publishers, 2003. Pp. 71–72.

Dyen et al., 1992 Dyen I., Kruskal J.B., Black P. An Indo-European classification: a lexicostatistical experiment // Transactions of the American Philosophical Society. 1992. Vol. 82. Pp. 1–132.

Holden et al., 2005 Holden C.J., Meade A., Pagel M. Comparison of maximum parsimony and Bayesian Bantu language trees // The Evolution of Cultural Diversity: a phylogenetic approach. London, 2005. Pp. 53–66.

Huson and Bryant, 2006 Huson D.H., Bryant D. Application of Phylogenetic Networks in Evolutionary Studies // Molecular Biology and Evolution. 2006. Vol. 23. 1–2. Pp. 254–267.

Michener and Sokal, 1958 Michener C., Sokal R. A statistical method for evaluating systematic relationships // University of Kansas Science Bulletin. 1958. 1–38. Pp. 1409–1438.

Moulton and Bryant, 2004 Moulton V., Bryant D. An Agglomerative Method for the Reconstruction of Phylogenetic Network // Molecular Biology and Evolution. 2004. Vol. 21. 1–2. Pp. 225–265.

MrBayes MrBayes: Bayesian Inference of Phylogeny. Ååá-ñàéò: <http://mrbayes.csit.fsu.edu/index.php>.

Nei and Saitou, 1987 Nei M., Saitou N. The neighbor-joining method: a new method for reconstructing phylogenetic trees // Molecular Biology and Evolution. 1987. Vol. 4. ¹ 4.  
Pp. 406&ndash;425.

Tassy and Darlu, 1993 Tassy P., Darlu P. La Reconstruction Phylogénétique. Concepts et Méthodes. Paris, 1993.

Éåâåíøòåéí, 1965 Éåâåíøòåéí  
Â.È. Äâîè-íûå êñäû ñ èñlðàâëåéàì âúïäääíèé, âñòàâîé è çàìåùåíèé ñèìâîéâ //  
Äíéëäû ÁÍ ÑÑÑÐ. 1965. Ô. 163. ¹ 4. Ñ. 845&ndash;848.