

Yẽãẽòðĩĩíúé èĩããẽñ àèãẽãèòèçĩĩã (èèĩããèñòè-ãñèèã ìðèìã-àìèy è èãììàòèçàòìðó)

Áãóìð Ìããã Óãĩòãĩãã- Æèĩãĩãã
09.07.2010 ã.
Ìñèããĩãã ããĩãããĩãã 09.07.2010 ã.

In essence, only in the electronic form this or that historic source may be fully investigated by the researcher. To create an automatic lemmatizator is the foremost and the most difficult task of the electronic lexicography and source study. While elaborating it, not only the orthographical, phonetic and morphological variants should be taken into consideration, but also the dialectal formations.

Nĩçããĩãã

yẽãẽòðĩĩíúó àãç àãĩíúó ìèèðúãããò ìãúã ãĩðèçĩíúó ìãðãã èñòìðè-ãñèè ÿçúèĩçĩãĩãã.
Ìĩ ñóúãñòãó, èĩãĩĩ ã yẽãẽòðĩĩíúé àèãã òìð èèè èĩé èñòìðè-ãñèèè èñòì-ìèè ñòãĩãèòñy
ãĩñòóĩíúè èññèããĩããòãèp ã ìñèĩé ìãðã è ìñèãò áúòú ìñãĩãĩ èì ãĩ ãñãé ìñèĩòã ñãĩèò
ñãĩéñòã. Ìãèì èç ìãèãĩããã ÿðèèò ìããããĩãèò ìðèìãðĩã è ãĩèãçãòãèüñòã ìñèãããĩããĩ – ñèãì-èĩìèè
ìãããĩðĩãñèèò òãð, ààòèðóãìúò ìãðãúìè ããñyòèèãòèyìè XI ã.
(ñì. [Çãèèçĩyè 2003]). Áãçòñèĩãĩĩ, ãĩçĩãĩĩñòè èñòìðèèĩ-èèìããèñòè-ãñèĩãĩ ããèèçã
èñòì-ìèèĩã ñóúãñòããĩĩ ãĩçðãñòãò ìññèã ìðããããðèòãèüĩãĩ ìãðããĩãã èò ìã ãĩñòóĩíúé
èìñüòãòãĩúé yçúè – èĩĩããðòãèè ã yẽãẽòðĩĩíúé-ìããĩðĩúé àèã. Á yòì ñèó-ãã ñ ìñüüòp
èìñüòãòãĩúé ìãðããĩòèè ãĩñòãòì-ì ãúñòðì ìããò áúòú ñĩçãããú èñ-ãðĩúããpùèã òìðìíòèãçãòãèè
è èpãìíó òãèñòó òãèòè-ãñèè ããã çããèñèìñòè
ìò ãã ìãúããã. Áããñòã ñ òãì òãììñòú ìãó-ìé ìãèèèãòèè ðóèìñèñíúó èñòì-ìèèĩã ìĩãĩãèòãòì
ãĩçðãñòããò, ãñèè ìã ñĩðìãĩããããòñy ìã ìðìñòì òìðìíòèãçãòãèyìè, ã èò èèìããèñòè-ãñèè
ìãðããĩòãĩíúè ããðñèyìè, ò.ã. ñèĩãĩòèãçãòãèyìè è òãèñòãì. Nĩçããĩãã ããòìãòè-ãñèĩãĩ
èãììàòèçàòìðã ã yòì ñüññèã yãèyãòñy ìãðããĩ-ãðããĩé è ìãèãĩãã ñèĩãĩé çããã-ãé yẽãẽòðĩĩíúé
èãñèèèĩãðãòèè è yçúèĩçĩãĩãy ã òãèì. Ìðè ãã ìããĩòãĩãã ãĩèãèíú áúòú ó-òãĩú ìã òìèüèì
ñòãìããòãòìúã ñèĩãĩèçĩãĩèòãèüĩã ìãðãçòú, ì ì ìĩãĩ-èñèãĩíúã ìðòìãðãòè-ãñèèã, òìãòè-ãñèèã
è ìðòìèĩãè-ãñèèã ããðèãìòú, ã òãèãã àèããèòìúã ìãðãçãããĩãèy ðãçĩé ìðèðĩãú.

Nĩrúáñòáííéá

nĩrúáííáí íáðá-íy àèèáèòèçìíá, èõ nĩrúáííáðàçííá èñ-èñéáíéá íáíáðííàèìù, áí-íáðáúõ, äëy òíáí, -òíáú ìðèíáííáíéá èáììàòèçàòíðá áúèí ìàèñèìàèùíí yóòáèòèáíùì – òí-íùì è ìíèíùì – ìðè áú-èáííáíéè èáèñè-áñèèõ ááéíèõ, à áí-áòíðúõ, äëy òíáí, -òíáú ìðè ñíçááíéè ñèíáííóéàçàòáëy ìííéè áúòù òñòáííáééáíá àèèáèòíáy ìðèðíáà ðóèííèñè. Áúyáéáíéá è èçó-áíéá àèèáèòèçìíá, èàè èçááñòíí, ýäëyáòñy áàæíé -àñòùp èpáííí èñòíðèèí-èèíááèñòè-áñèíáí èññèááíááíéy. Èíáííí àèèáèòíúá íáðàçííáíéy áèèæá áñááí ñòíyò è àèíáíééá æèáíé ðá-è è ýð-á áñááí áá òàðáèòáðèçòpò. Ñ ìíííùùp ìñíáíáí èèp-à èáììàòèçàòíðá ìíáòò ñíçáááàòùñy óèàçàòáèè àèèáèòèçìíá è ýèáèòèðíííé ááðñèè òíé èèè èííé ðóèííèñè.

Äèàððíé-áñéày

èèàññèòèèáòèy ìñíáíííñòáé áðááíáðóññèèð àèèáèòíá áèèp-áàò ñèááòpùèá ìàðáìáòè-áñéèá áðóíù (ñì. [Èðúñüèí 1998]):

(1)
ìðáñéááyíñéèá àðòàèçìù[1];

(2)
ìðáñéááyíñéèá àèèáèòèçìù[2];

(3)
áíñòí-ííñéááyíñéèá èíííáàòèè áíñèñùìáííáí ìáðèíáà;

(4)
íáùèá ìñíáíííñòè äëy òíáí èèè èííáí áðáàèá áíáíðíá;

(5)
ðàçííáðáíáíùá èíííáàòèè òáò èèè èíùò ìòááèùíùò áíáíðíá.

È
ìðáñéááyíñéèì áðòàèçìáì ìòííèòèñy, á -àñòííñòè, áàðèàòèáííòù òíðì 3 è. áá. è ìí. -èñèá á ìðàçáííá íá -òù è -ø. Ñ-èòáàòñy, -òí ìíá ìòñòòñòáíáéá á áðááíáðóññèèð òáèñòáò ñáááðí-áíñòí-ííáí ìðíèñòíæááíéy, à á ìíááíðíáñéí-ìñèíáñèèò èñòí-íéèáò áúèá øèðíéí ìðááñòááéáíá á ðàçííáò òèíáò ñíðyæáíéy èàè á áá., òáè è áí ìí. -èñèá, á òí áðáíy èàè á áðááíáðóññèèò páí-çáíááíùò ìèñùìáííùò ìàìyòíéèáò – á ìñíáííí á 3 è. áá. ÷. I

ńĩđýæáíèý

[Çàèèçíýè 2004: 152–153]. Ááçõñéíáíí, òðááòáòñý òùàòáèúíáý ìðíááðéà ýòíáí áúáíáà, ìñéíèúéò ìàòáðéàè þáí-çàíááíúò èñòí-íèéíá á ýòí ìòííðáíèè èçó-áí áíðàçáí òóæá ìááíðíáñéí-ìñéíáñéèò ááííúò. Nĩññòááèáíéá ýéáèòðííúò ááðñéè ìááíðíáñéí-ìñéíáñéíé è áàèèòéí-áíèúíñéíé ðóéííèñáé, ñàýçáííúò ñ íáúèí ìðíòíáðáòí, ìðéáááííúé áúáíá íá ìáòááðæááò: áàèèòéí-áíèúíñéáý ðóéííèñú ñíááðæèò ñòíèú æá ðàçííáðàçíúá áàðéàòéáíúá òíðí, -òí è èñòí-íèè ìááíðíáñéí-ìñéíáñéèé. Áñòú ñíííááíèý ìíèááàòú, -òí è ñòááíá-çàíááíúá èñòí-íèèè ìðèè-àèèñú ðíáñòááííúè òáðáèòáðèñòéèáìè. Èñííèúçíááíéá ñíèíèíéé ýéáèòðíííé áúáíðéè òíðí ìçáíèèèí áú áíñííèíèòú ì-ááèáíúá ìðíááéú á ýçúèíáúò ááííúò. Èçó-áíéá áàðéàòéáííñòè ýòíáí òèíà òáí áíèáá áàæíí, -òí ìíá ýáèýáòñý èíáíááðííáéñéèì íáñéááèáì á ìðáñéááýíñéíì.

Á

èá-áñòáá ìðéíáðá ìðáñéááýíñéíáí àèáèáèòéçíá ìíèáçàòáèúíí, á -áñòííñòè, ðáñíðíñòðáíáíéá íéíí-áíèý -íáè/-ááè á Á Ā ñóúáñòáèòáèúííúò ìóæñéíáí ðíáà. Nĩññòááèáíéá ýéáèòðííúò ááðñéè ìáííáí è òíáí æá ñáíðíéèá ðàçííáí àèáèáèòíáí ìðíèñòíæááíèý ìðíòéáíðá-èò òñòíýáðáìñý ìíèíæáíèþ íá òñòíé-èáíí ðíñòá òàèèò ñéíáíòíðí á þáí-çàíááíúò èñòí-íèèèò. Òíèúéí ìíèíáñòááíáý ýéáèòðííáý áúáíðéá ñíñòááòñòáóþùèò ñéíáíòíðí ìçáíèèò òñòáííáèòú ìíáèèíóþ èàðòéíò ðáñíðáááèáíèý áàðéàòéáíúò ñéíáíòíðí íá -íó è -íáè, èíòíðúá, áíçííæíí, èìáþò ðíáñòááííúé ááíáçèñ (ñí. [Æíéíáíá 2009á: 224]).

Ááçõñéíáíí,

á èáííàòéçàòíðá áíèæíú áúòú ó-òáíú, íáíðéíáð, ìðííáðáííú ñ -æ-~, ñáíéñòááííúá áàèèòéí-áíèúíñéèì èñòí-íèèáì (ñí.: áíæ-ú, èæ-áíòò < èç+æáíòòú, ðæ-áòú < ðàç+æúæáòú è áð.). Ííè ìðáæáþò ìðáñéááýíñéèé ðáòéáèñ ìáèàòáèèçàòéè *zg, ìðíòéáííñòááèáííúé ñòáðíñéááýíñéííó ðáòéáèñò [-d’], á òàèæá ìááíðíáñéí-ìñéíáñéííó æá’ < æá’ < *-’d-’ (ñí.: íáúæáèòè, ðáæáýòè è ìíá.)[3]. Íðóíáðáííú ñ æá, òáèè ìáðàçíí, ýáèýþòñý ìðéíáðíí ñíáñòááííí ìááíðíáñéíáí àèáèáèòéçíá. Á òíí è áðóáíí ñéò-áá ñóúáñòáóþò æéíáðéíððáèòíúá íáíèñáíèý, èíòíðúá òáèæá áíèæíú ó-òáíú (ñí.: òóæ-èèè vs. ùþæáèèòú < *teudj-).

Íéíí-áíèá

1 è. íí. -èñéá -íí á ìðáçáííá íáú-íí ðáñííàòðéááþò èáè ìðáñéááýíñéèé àèáèáèòéçí, ó áíñòí-íúò ñéááýí ìíèò-èáèèè ðáñíðíñòðáíáíéá á Þæííé è Çàíááííé Ðóñè. Íáíáéí ñáíúá ðáííéá ñéò-áè ìíýáèáíèý òéáèòéáííáí -íí ìòíá-áþòñý á áàèèòéí-áíèúíñéíé ðóéííèñè è ìáðáíé-áíú èñéèþ-èòáèúíí ìáúíè òíðíáìè èííáðáòéáá ñ íóèááúí ñóóòéèñíí: àè-úíí, ìèá-úííñý, ìíèèòéèì è áð. [Æíéíáíá 2009á: 53]. Á ìáíáíúò ìðéíáðáò ìíæíí òñíàòðéááòú òííáòè-áñéíá ðàçáèòéá òéáèòéáííé èíííáàòéè, ñáíéñòááííé ñíáñòááííí áðááíáðóññéèì þáí-çàíááíúí áíáíðáí.

Ýéáèòðííáý

áúáíðéá èííáðáòéáíúò òíðí 2 è. íí. -èñéá íá ìáòááðæááò áúáíá íá óçèí àèáèáèòíí, ìááíðíáñéí-ìñéíáñéí, òáðáèòáðá èíííáàòéííúò òíðí òèíà ìðéáèòá, íáñéòá è íá èò ìòñòòñòáèè á þáí-çàíááíúò áíáíðáò, ááá, èáè óéáçúááàòñý, áúè ìðááñòááèáí áðóáíé òèí ááíáðáèèçàòéè – ìí ìáðáííò èáñéèííáñéííó èèáñíó.

Òí-íí

òàè æá íáńòíèò ááeí ñ äðóáíé ýðéíé íńíááííńòùþ, éíòíðàý ñáyçúáááòńý ñ ííááíðíáńeí-íńeíáńeèèè
áíáíðàìè – ñ íńíáúì ðàçáèèèáì ðàńeááyíńeèò áðóíí òèìà tort è
túrt (ńì.: ñðí-ùèú, ñúáðíáí
è áeííáíeý, íðáòáúìú è ò.á.). Òàèèá æá ðáðeáeńú íáíàðóæèáàþòńý á áàèèèèí-áíeúíńeèò
eńòí-íeèáò (ńì. [Æíeíáíá 2009á]. Ááçòńeíáíí, ýòí ááèèáò áíeáá ñeíæíé èàðòeíó ðáðeáeńàòèè
óèaçáííúò áðóíí, íí àíáńòá ñ òàì è íðááííeáááò áíáńáíeá áá ñòáìú á èáììàðeçàòèþ.

Áàæíá

íáńòí á ðóéííeńíúò eńòí-íeèáò çáíeìàþò ðàçíúá àeèeáeòíúá òèìú íáíáúáíeý òááðáúò
èèè ìýáèèè ááèèèèàòèííúò íáðáçòíá, éíòíðúá, íáńííáííí, èàè è äðóáèá àeèeáeòíúá
ýáèáíeý, áíeæíú ííeó-èòù íððáæáíeá íðè ííááèèðíááíeè íàðáèèáì èáììàðeçàòíðá.

Èèòáðàòóðà

Æíeíáíá

2009á – Æíeíáíá Í.Ô. É eńòíðèè ñeááyíńeíáí áàòèáà / Í.Ô.Æíeíáíá // Rozprawy Komisji J zykowej. – 2009. –
S. 221–229.

Æíeíáíá

2009á – Æíeíáíá Í.Ô. Éíðíóń äðááíáðóńńeèò ñíeńeíá Ìàðáíáńeńà Áððáìà Ñeðeíá. II: ÐíÁ, Ííáíá. 71à / Í.Ô.Æíeíáíá
// Russian
Linguistics.
– 2009. – 33. – C. 37–64.

Æíeíáíá

2009á – Æíeíáíá Í.Ô. Íðeńòáíeùá íááúéííeíá (è ðáðeáeńàòèè áðóíí òèìà túrt/túlt) / Í.Ô.Æíeíáíá // Ðóńńeèè ýçúè
á íàò-ííí íńááúáíeè. – 2009. – 1 1 (17). – Ñ. 127–133.

Çàèeçíýè

2003 – Çàèeçíýè Á.Á. Íðíáeáìú èçó-áíeý Ííááíðíáńeíáí éíááeńà XI ááèà,
íáeááííáí á 2000 á. / Á.Á.Çàèeçíýè // Ñeááyíńeíá ýçúeíçíáíeá. XIII Íáæáóíáðíáíúé
ńúáçá ñeááeńòíá. Èþáeýíá, 2003 á. Áíeèááú ðíńńeéńeíé ááèááàòèè. – Í., 2003. – Ñ.
190–212.

Çàèeçíýè

2004 – Çàèeçíýè Á.Á. Áðááíáíááíðíáńeèè àeèeáeò / Á.Á.Çàèeçíýè. – Í.: ßçúèè ñeááyíńeíé
éóèúóóðú, 2004. – 872 ñ.

Éðúñúèí

1998 – Éðúñúèí Á.Á. Äðááíéé ííááíðíáñéí-íñéíáñééé àèèèèò íà íáúáñéááyíñéíí ðííá / Á. Á. Éðúñúèí // Áííðíñú ýçúéíçíáíéy. – 1998. – 1 3. – N. 74–93.

[1] Èíááòñý á àèáo ñíððáíáíèá òíáí

èèè èííáí íáðàçíááíéy íáúáñéááyíñéíáí íðíèñíðíæááíéy á òíé èèè èííé àèèèèèèéé çííá íðè ááí òððáðá èèè ÷áñðè÷íí ñíððáíáíèè á äðóáèð àèèèèèèèð.

[2] Èíáðòñý á àèáo àèèèèèèííá ðàñðíæááíéy

è èçíáèíñú, áíçíèèèèè á íðáñéááyíñééé íáðèíá.

[3] Äðááíáðóññéáy íððííðáðèy òíðíúááòñý.