

ЇЇÄÅËÈ È ØÅÖÍÍËÍÄÈÈ ĐÅÀËÈÇÀÖÈÈ ØþÐÊÑÊÈØ ËííííÁÍØ Â ĐÓÑÑÉÍ-ØÀØÀÐÑÉÍÉ ËÅ

Àâòîð íââçîðâà ïëüâà Àâáíèðâà  
16.08.2012 à.  
Íñëéâíå ïáíâéâíèå 27.08.2012 à

**Summary.** The

article is presented the base models of new Russian-Tatar lexicographical database. Lexicographical database consists of interrelated components (Russian and Tatar) with an independent structure. The components are merged by semantic codes at the level of lexical equivalents. Each component contains a grammatical, semantic and derivational information. Component of the new Turkic language will have a structure similar to the structure of the Tatar component. The article discusses the basic design problems of new Turkic components and also technologies of extension of lexicographic database by adding new components.

Äëÿ òåîðåòè-åñêèò èññëåâíàíèé ïï  
 êïïüþþoåðñíé èëíåâèñòèåå è ðåçðåâîòèè èëíåâèñòè-åñêèò ïðèëíæåíèé (ñèñòåì  
 ìàðøéèííå îðåâîñå, èíöîðìåöèííå îñèñå è äð.) åíüþòíå çíà÷åíèå èíåðå  
 ñíçäåíèå ñíåöèåéèçèòíåíûò èäñèèòíåðåòè-åñêèò ðåñòðñíå ñ ååðåèèçèòíåíûò  
 èëíåâèñòè-åñêèè àííþòåöèÿ. Õóïéòèííåëüíûò îñíåâííñòè òåèèò èäñèèòíåðåòè-åñêèò  
 ðåñòðñíå îñðåâäëþþoñý èðóåíî îñòåíöèåëüíûò çåäà-, äëÿ ðåðåíèé ëîòòðûò ïï è íðíåðèòðþþoñý.

Â ìàñòïýùåå âðåïì ñïçääìí  
 çìà÷èòåëüííà ëîëè÷åñòåí ííñäñôóíéöèííàëüíûõ ëäåññèéíäðåòè÷åññèéõ ðåññóðñíâ. Äëÿ  
 òþðéññèéõ ýçûêíà ííæíí ìòåðòèòü íàðàëéëüíûé ííëåáé-ñëíâàðü òþðéññèéõ ýçûêíà  
 Õòðåöéíâí ëëíâññèéõ ÷åññèíâí íáúññôåà (Turk Dil Kurumu) [Türk Dil Kurumu, 2012], íðíâéò ìàøéíííâí  
 ñííàá áåøééõññèíâí ýçûêà [Ìàøéííûé, 2012], ýéâéòðíííûå ñëíâàðè éàçàðññèíâí ýçûêà  
 [Ðóñññé-éàçàðññèéé, 2011]. Äëÿ òàøàðññèíâí ýçûêà åúà å 90 ð ãíäàò ïðíøéíâí áâéà  
 áúëà ðàçðåáíòàíà èíñöåíöèÿ ìàøéíííâí ñííàá [Áóñàðååâ, 1995] è ðàçðåáíòàíû  
 ðàçéè÷íûå ýéâéòðíííûå ñëíâàðè (ýéâéòðíííûå ñëíâàðè ABBYY Lingvo, èíøåðñåò-ñëíâàðè ýçûêà  
 òàøàðññèéõ ìèñàðåëéé [Èàçàíññèé, 2012] è  
 äð.). Îäíàéí àéooàëüííé ýâëÿòöñý çàäà÷à ðàçðåáíòèè ñïåöèàëëçèòðíâàíûõ  
 èåññèéíäðåòè÷åññèéõ ðåññóðñíâ äéÿ ðàçéè÷íûõ òàðåðòè÷åññèéõ è íðèéëàäíûõ öåëåé,  
 òàéèõ, êàé ñðàâíéòåëüííà èçó÷åíéà òþðéññèéõ ýçûêíà, èéíââèñòè÷åññèíâ íåðñíâ÷åíéà  
 ññèñòåí ìàøéíííâí íåðåâíâà è äðooäèõ íðèéëàäíûõ èéíââèñòè÷åññèéõ òåðñííëíâé.

Â ÍÈÈ «Їðèëëàäíàý ïâìèòèà»  
 Âëàääìè è àoéê Ôàòàðñòáíà áåëëåðöñý ðàçðàáîòêà ðóññêî-òàòàðñêîé ëåêñèëtâðàòè-åñêîé  
 áåçú áäííûò [låâçïðñâà, 2012] íà ïñííâå ðóññêî-òàòàðñêîáí  
 ñëåäàðöý íåúâíî ìêëí 50 000 ñëåâ ííä ðåëëåðöèåé Ô.À. Âäíèåâåà [Ðóññêî-òàòàðñêèé, 1997]. Íâàÿ  
 ëåêñèëtâðàòè-åñêàÿ áàçà áäííûò (ËÁÄ) ïðëåíòèðñâáíà, ïðåæäå áñâåñ, íà ïðèëíæåíéÿ  
 á íåëàñòè áåòíàòè-åñêîé íåðåáíòè è òåêñòíà (àííîòèðñâíèå êíòíñíà òåêñòíà,  
 ìàòèííûé íåðåáíà è êíòíðàöèííûé ííèñ). Ðàçðàáíòàíàý ííäéü ËÁÄ íçâíèÿò ðåñòèðþòü  
 áäííûé ëåêñèëtâðàòè-åñêèé ðåñòð íà äðóàëå òþðëññêèå ýçûèé. Íäéëþ-áíèå íâûò  
 òþðëññêèé ëííííáíò áûíííýåðñý íà ïñííâå ííäéé è òåðííëíäé, ðàçðàáíòàííûò  
 äëÿ òàòàðñêîé ëííííáíò áàçû áäííûò.

Êåéñéêîåðàòè÷åñéâý ááçà äáííûô  
ñïñòòèò èç áçåèìñýçáíûô ðóññéîé è òàòàðñéîé êííííáíò, èíåþbuø íåçàâèñèíóþ  
ñòðóëòòðò è íáúåëíýåíûô íðë ííííûê ñåìàòè÷åñéèò êíäíâ íà óðííâ íå éåéñè÷åñéèò  
ýéâéâåëäíóíâ. Êííííáíòå áéý ãðóäíâí ðþðéñéíâí ýçüñéâ áóäåò êíäòü ñòðóëòòðò,  
ííííáíóþ ñòðóëòòðå òàòàðñéîé êííííáíòû, è ñåýü ñ ðóññéîé è òàòàðñéîé  
êííííáíòå ñòðóëòòðå ñòðóëòòðå ñòðóëòòðå ñòðóëòòðå ñòðóëòòðå ñòðóëòòðå.

Êæäääý èç êïííííáíó ñíááåðæèò  
âðåàííäåðè÷åñéóþ, ñíáíáíóðè÷åñéóþ è ñëüâíáðåçíâðåðåëüíóþ èíðíðíàðöþ. Åðåàííäåðè÷åñéäý  
-åñíðííü ðåðåðåðåðííé åñíííííåðíí (T-åñíííííåðíí) îðåííåðåðåðííåðíí åñíííííåðíí: ñëüâíáðåðåðíí

ìàðàìàòðû, èâè ñàðàíòè-÷åñêèé êïä, ñèïâàðíàÿ ôîðìà ïñïíàû, îïðôîëäè-÷åñêàÿ ôîðìà ïñïíàû, îïðôîëäè-÷åñêèé è îïðôîëíà-÷åñêèé òðëü ïñïíàû. îïðôîëäè-÷åñêàÿ ôîðìà èñïíëüçóàòñÿ îðë ïðïæääåíè ìïååðöíñòíûô ôîðì ñèïâàðíàÿ ïðëñïäæíåéý àôôèëñàëüíûô îïðôå è èñæåò (íà ñïäïàäàòðû ñì ñèïâàðíé ôîðìé ïñïíàû çà ñ-àò èâèéò-ëèåí âïòðåäííéò èçïåäíåíè (íàïðëàð, ÷åðäåäàíàíèé èëè ìøñòòñòàéý ñèïâàðíà).

Íäíí èç íáçíá-÷áíéè ñáíáàíò-÷åñéïáíí áíàá Íááñíá-÷åíéà ñáýçè ðàçéë-÷íúó  
éíííííáíò, þòí íñíááííí áâæíí íðè ðàñøèðåííè èáéñéèéíáðåöé-÷åñéíé ááçû íá ãððåéà  
òþðéñéèá ýçûéé. Ííðóíéíæ-÷åñéèé è ííðóíííéíæ-÷åñéèé óèíü íñííáú íáíáõíæíü äéý  
ñáýçè ní ñéíáàðáí íéí-÷áíéè è éñííéúçóþðñý íðíáðáííæ ííðóíéíæ-÷åñéïáíí áíàééçá  
é ááíáðåöéé.

Íðîöâññ ãíáàâéäíéý á  
ëåêñèéñðåòàôè÷åññóþ áàçó äàííûõ ííâíé êííííáíòû äéý äððåâññ òþðññéñññ ýçûéà  
áóäåð ñññòíýòû èç äåóó ÿòàíññ: ðàçðàáíòèà ñëíâàðÿ ïñññà è ðàçðàáíòèà ñëíâàðÿ  
íéí÷-àíéé.

Íðè ñíñooðíâíéè ñíèíâàðöý íñííâ  
ãëàâíâíàÿ íðíáæâíà çâéëþ÷åâðöý â íðâéòè÷åññî íòññóðñòâèè åâóýçû÷íúô ñíèíâàðåé  
íâæäö ýçûñâàè òþðêññéíâí ñâíâéñòâà (çà íâíïëøèì èñéëþ÷åíèâí â áèää òàðàðññéí-òóðåöéíâí  
è óððåöéí-òàðàðññéíâí ñíèíâàðåé). Â óí ãâ áðâíÿ ìíâàðöý ííæâñòâí òþðêññéí-ðóññéèö  
åâóýçû÷íúô ñíèíâàðåé, òâéëô, êâé ðóññéí-êâçâðññéèé, ðóññéí-óçââðññéèé, ðóññéí-êèðâèçññéèé  
è äð., êîòðñâ è ííâóò áûòü ííæâíû ã íñííâó íðíâéòèðíâàíéÿ ííâû òþðêññéèö  
êííííâíò ÈÁÄ.

Ñîçääìéà ñïèïààðý ïñïíà äëÿ ïïàíé  
éñïíïíàíóù ñïñòñèò èç äåooý ýòàïíà: ååòïíàòè-÷åñëïíä í è ðó-ííä. Åí åðåïíý àåòïíàòè-÷åñëïíä  
ýòàïí ïðèçäåäèöñý ñïíñòàåéäíèå ñïíäåðæèííä ðóññëïé èñïíïíàíóù ËÄÄ ñ åôïäàïè  
ñéïäåäðûñ ñòåðåé äåóýçü-÷íäí ðóññë-òþðëññëïä ñïèïààðý. ïðè ëíáåððæåíèè  
ñïäïäåäþþùñ äåéñïåí å òþðëññëïé èñïíïíàíóù ñïçäåäðñý ïíàäý çäïèñü ñ òåí æå  
ñïäïäåíòè-÷åññë-÷íäí ëíäí, -÷òí è ó ñïäïäåðåé äåéñïåíù å ðóññëïé èñïíïíàíóù ËÄÄ. Åñëè  
äëÿ ðóññëïé äåéñïåíù ðóññë-òþðëññëïä ñïèïààðý àíäéïä-÷íàÿ äåéñïåíà å ðóññëïé  
éñïíïíàíóù äåéññë-÷åñðåòè-÷åññëïé áåçü íå íåðïäèöñý, òí äëÿ ïïàíé çäïèñè òþðëññëïé  
éñïíïíàíóù ååäíåððåòðåòñý ïíàûé óíéèäüíûé ñåïàíòè-÷åññë-÷íä.

Íñiêà çàâåðøáíéý àâòòíàòè÷-åñêîáí  
yóàïá ñíñiñòàáééáíéý íñiñòðíáííáý ÁÁ òðåááòå ãíííéíèòåäëüííé ôèëíëíäè÷-åñêíé  
yéñiáðöèçû. Íðiñöåññ ðàçðàáìòèé ñëíñàðöý íéñí÷-áíéé ííâíé òþðêñéíé êííííáíòú  
íà÷-éíàåðöý ñ íñiñòðíáíéý òàåéèöû ñíñòååðòñòåéé íåæöí áòðèéññåéüûíè íñðåáàìàìé  
áíàååéýåííáí è òàòåðñéíáí ýçûéíâ. Çàòåí ðàçðàáåòùåàþþöý íñðòíòåéè÷-åñêéà íðååééà  
äéý åííáðåöèé áññåô åíçííæíû ñëíñåíòðí ííâíáí ýçûéâ ñ ó-åðíí íðèíýòû ïðåðàíè÷-åñíé  
íà éíèé÷-åññåô áòðèéññåéüû ñíðòåí (á íâñòíýùåé áâðñéè íà áíéåá 5). Íðè  
ðàçðàáòéå ñíðòíòåéè÷-åññééò íðååéè íæðåééëåëüíí íñðéññòíäèò áùåååéåíéå áññåô  
íñðòåíèé-åññééò è íñðòíííéíâ-åññééò öèííá ñíñíá ííâíáí ýçûéâ. Â çàâåðøáíéè áññå  
óéàçáííûå áèäû íéñí÷-áíéé áàòòíàòè÷-åññéé åáíåðéðþþöý è çàíííýþþöý à òàåéèöû ÁÁ  
ñí ñòðóéòòðíé, áíâíéé-ííé ñòðóéòòðíå òàòåðñéíé êííííáíòú ÉÁÁ.

Ðåàæëèçàöèý ãïïíïéèòåðëüíûô  
òþðéñéèô ëïïíïíáíô à ËÁÁ íïçâïéèò íïëó÷èòü ýóðååêòèåíûé èíñòðóïåíô àåòïïàòèçàöè  
ñðååíéòåðëüíûô èññéäåíâáíé è ðàçðåáíòè ïíïäïýçû÷íûô ëèíåâèñòè÷åññéèô  
òåðíïéïåé, îðååéåå áññåáí, óåðíïéïåé ïíïäïýçû÷íûô ìïèñéà è ìàøéíïäí íäðååíäà  
æëü áéècëöïåñòåííûô ýçûéïà áíóðé òþðéñéàí ñåjåéñòåà.

Äëÿ  
àåòìàòèçàöèè ñðàâíèòåëüíî-òèïïëìàè-÷åñèèõ èññëäåìàíèé òþðêñèèõ ýçûêîâ è  
ðåàäèçàöèè ëíòåäðèõìàíúô òþðêñèèõ êíïííåíò à êåñèèõðåðåò-÷åñèíè ààçå àäííûõ  
íåíáõíàèìà íåùäÿ ìíäåëü ìíðòíëíæè. ìíäåíáÿ ìíäåëü ìíæåò áûòü ñíçààíà íóðàí  
ðåñòèðåñíèv ìíäåëé òàðàðñîéíè ìíðòíëíæè. ðåàäèèçíåñííè ã FÅÄ ìíðòíëíæ-÷åñèè

Íàðàìåòðû ñëèïåîòðî ã óþðéññéëõ ýçûêåõ îíæíï ïðåðàçäåëëöü íá ïðïñòùå è ñëïæíûå [Óàèëëïâ, 2011]. Íðïñòùå íàðàìåòðû ïðåäñòàâëå (âíçíïæíï íàéë÷ëå, èëéå íòñóðòñòåðåéå íàðàìåòðà). Ñëïæíûå íàðàìåòðû ïðåäñòàâëåíû ãåóïý è áíéåå àëüòåðíàðéåíùìè çíà÷åíéýìè è ííäóö áúòü íáÿçàðåëüíûìè èëëè ôåéññéëüðàðéåíûìè. Ôåéññéëüðàðéåíûå íàðàìåòðû íòéë÷àþðñý ìò íáÿçàðåëüíûð òåì, ÷òî ííäóö íòñóðòñòåðåíûå â ñëèïåîòðî. Áñå ïðïñòùå íàðàìåòðû ýâëýþòñý ôåéññéëüðàðéåíûìè. Ëíèñàíéå óþðéññéëõ ãðàììàòè÷åññéëò êàðåäñíðéé ïí ãàííûì êðèòåðèýï ííçâíëýÿ ðòðàçèòù ñíåðòëëéõ êàæäíå ýçûêå à ðàìéåò ñíéååðñàëüíé êéàññéòðéëåðéé. Á ñåïþ í÷åðåäü, óñéëéàíéå ãíðèçííòåëüíûð è íåæôðíàíåñûð ñâÿçâé íåæäó ðàçíïýçûñíûìè èííïíàíòåì äàåð åíçíïæíïñòü óñòàííåëåíéý áíéåå ñëïæíûð ñíòåðåðñòåðéé, íàðèåð, à òåð ñëó÷àýõ, èíäåà íàíà è òà æå óóíëöéý à ðíäñòåðíûð ýçûêåõ ðåàæçóðñý ååéíèòåìè ðàçíûð óðíåíáé (ííðôåìû, ñëíàà, ñëíàññí÷åðåíèý è äð.).

Áëàãîääàðíñòè

Èññëåâíààíé  
âùíííéáíí íöð ôéíàíñíàíé íïäåâðæêå ÐÃÍÔ â ðàìèàö íàó÷íí-èññëåâíààòåëüññéíå  
íöðâéòå ÐÃÍÔ<sup>1</sup> 11&ndash;14&ndash;16024à/Â.

Nïèñîê ëèòåðàòóðû

Türk Dil Kurumu, 2012. Türk Dil Kurumu [Yéáéòðííúé ðåññóðñ]. URL: <http://tdk.org.tr> (ääòà íáðåùåíëv: 10.03.2012).

Ìàøèííûé, 2012 Ìàøèííûé ôîíä ààøêèôñêïâ ýçûèà [Ýëåâôòõííûé ðåñóôñ]. URL: <http://mfbl.ru> (ààðà ìàðåìóùåíÿ: 15.02.2012)

Дóññéî-éàçàõñéèé, 2011 Дóññéî-éàçàõñéèé è éàçàõñéî-ðóññéèé  
ñéïâàðü [Ýéåéòðííûé ðåñóðñ]. URL: <http://sozdik.kz> (äàòà  
íàðàùàíèé: 21.10.2011)

Áóõàðàåâ, 1995 Áóõàðàåâ Ð.Ã., Ñàôèóëëíà Ô.Ñ.,  
Ñóëáéíàíâ Ä.Ø. è äð. È êíöâïöè è àzøéííâ Ôíàä Ðâñíóáëèéè Òàòàðñòàí // Òàòàðñèéè  
ýçûé è ííâûå eíöîðìàòèííûå òåðííèíàèè. Ñâðèÿ: Éíòåëéåò. ßçûê. Èííüþòåð. Åûï.2.  
Èàcàíü, 1995. N. 20&ndash;35.

Êàçàíñêèé,  
2012 Êàçàíñêèé èèíääñäðàôè÷åñêèé ôíïä [Ýëåêòðííûé ðåñóðñ]. URL: <http://klf.ksu.ru> (ääòà íàðàùåíÿü: 5.04.2012)

Íââçîõîâà,  
2012 Íââçîõîâà Í.À., Ñàëèìâ Ô.È., Õàèèìâ Á.Ý., Æàòèàòóëëè  
À.Ð. Íðääìéçàöèÿ èíóïõàöèíííâ Ìèñêà à ðóññéí-òàòàðñéíé èâéñèéíâðàôè-÷åñéíé  
ääçá ääíûò. // Íðèéëääíay èéíââèñòèà à íàóéå è íàðäçâàíèè: Náðíèé  
òðóäâ VI íæäöíàðííé íàó-ííé êíñôåðåíöè.   
Ñíá 2012

Đóññêî-òàòàðñêèé, 1997 Đóññêî-òàòàðñêèé  
ñêîñàðñ / ùüä ðåä. Ô À Äáíèåñà ì· ËÍÑÁÍ 1997

Õàêèìâ, 2011 Õàêèìâ Á.Ý., Äèëüíóëëéí Đ.À. Ê ðàçðàáìòêå  
ïðöîéïäè-åñêïäñ ñòàíäåðòà äëý ñèñòàí àáðòíàòè-åñêïé íáðàáìòêè òåéñòâå íà òàòàðñêïí  
ÿçûéå // Ñèñòåíüé áíàéèç è ñâìèòè-åñêïäñ íäåéëðíâáèå: ìàòåðòèåëü  
âñåðñïðñêéñêïé ëíñôððåíöèé ñ íàçäöíáðíäñ (SASM-2011). Êàçàíü, 2011. № 209&ndash;214