

Î ðåïðåçåíòàðèâíñòè òåêñòïâ è ýëåìåíòåõ îðïäðàííñâí èíñòðóïåíòàðèÿ äëÿ êîðíóñà áóðÿøñê

Áâòîð Ëþáîâü Áàøèíèìàååáíà Áàäìàååà
28.06.2008 á.
Íñëååíå ïáíîâéåíèå 15.07.2008 á.

Ôâçèñû à ôïðìàòå RTF (90.13 kB 2008-07-15 16:04:31) Ôâçèñû à ôïðìàòå PDF (92.95 kB 2008-07-15 16:05:00)

Óåðñíïèïâèè êîðíóñííé ëëíââèñòèêè óæå øðíñíè íðèçíàþþñý íàéâèå á ðòèðàëüíûé á èññéåâîáàðåðéüññéíé íðåèòèå ñíâðåðàíåííé íèñúìåííûô íàìþòééåô, ñåñíäíÿ ííæåò áûòü ííäâåðäíó «êîðñíðèçåöèè», -òî áâçíñâîðñí ðòèðåñòâåéÿåò çàìàí-èåûå áíçííæíñòè è òåðñíïèïâèè ñ óåðñíè ñíçâàíèÿ óñéïâèé äëÿ åâí åâéüíåéòåâåí ðòèðåñòâåíèÿ á yéâéòðñííûô ñèñòâåàò êàè á èññéåâîáàðåðéüññéíé,

Iåòïâiæïæy ðàçâèåþùåñy êïðïóññé eëíåâèñòeëe iïçâiæyåò iïéó-àòü lâéñèiæüí fâäåæíûå ýçûéiâûå ôàéòû ðåàëüífâi ôóíéðêïðïñá áóðyöñèò òåéñòiâ, íà êïòïðì àïðiáeðóþöñy êïéïðåáíñåðû [Grieves; Watt], iïçâiæyþùèå iïéó-àòü iïéíiøåéñòiþ ìéòiðiá ýçûéiâûå ååeïéò.

Øåðøíîèñàèè êíðíóñííé ëëíáâæñòéèé íàïðåâæåíû íà ñâïåâñí ðíäà «åéñâæèçàðéþ» èññíèåáñâæåò ðíññíèåéèåò ðíññíèåéèåò ðíññíèåéèåò ðíññíèåéèåò ðíññíèåéèåò

Ââèäó òïäî, -òí iöñäðàìû-êíïéëäáñåðû ïäðàíè÷åíû â ñâïèõ áïçïíæíñòþ, äëÿ áïéåå ìïéïäà èñïïëüçâàéÿ äàííûõ êïðïóñà áóäåð Manatee / Bonito, IMS Corpus Workbench (CWB), CQP/CQP-WS, Xaira (SARA).

Daáíóö aúüííéíáá íöðe öéíáíñíáíé ííäááðæéå ÐÔÓÈ (íöðláéò 08-06-00151-à) è Áiáðééåáíñéíáí nííááðåá láóó-íúó nííááñåñòå (ACLS).

Nïëñîê ëèòåðàòóðû

Grieves – Grieves, Chris. ConcApp Concordancing Programs [Ýéâéôðííûé ðâñóðñ]. – Ðâæèì äïñòóïà : www.edict.com.hk./PUB/concapp, ñâïáíäíûé. – Çäãë. ñ òèööé. ñòðàíèöü.

Watt – Watt, Robert. Concordance [Ýéâéòðííûé ðâñóðñ]. – Ðâæèì äïñòóïà : www.concordancesoftware.co.uk, ñâí – Çäää. ñ òèòóé. ñòðàíèöû.

On the representativity of texts and software development for a corpus of the Buryat language

Lubov’ D. Badmaeva, Zhargal B. Badagarov

Institute for Mongolian, Buddhist, and Tibetan Studies of the Siberian Branch of the Russian Academy of Sciences,
Buryat State University, Ulan-Ude, Russia

A corpus of Buryat texts is being constructed, mostly by scanning. Concordance software is used on these electronic texts, which include samples of literary, public, and scholarly works. Extralinguistic analysis, morphological annotation, and graphemic analysis are discussed.

[2] Çàääà÷à ñeíòàêñè÷åñêîé è ñäìaíòè÷åñêîé ðàçìåòòè ïîëà îòëîæåíà íà áóöóùåå.