

Í ðáíðáçáíòàòèáííñòè òáèñòíá è ýéáíáíòàò ìðíãðàíííáí èíñòðóíáíòàðèý äèý èíðíóñá áóðýòñè

Ááóíð Èþáíáú Ááøèíèáááíá Áááíááá
28.06.2008 á.
Íñéááíá íáííáéáé 15.07.2008 á.

Òáçèñú á ðíðíàðá RTF (90.13 kB 2008-07-15 16:04:31) Òáçèñú á ðíðíàðá PDF (92.95 kB 2008-07-15 16:05:00)

Òáðííéíáèè èíðíóñíé èèíáèñòèèè óæá øèðíéí ìðèçíáþòñý íáèáíéáá ìòèíáèúííè á èññéááíáòáèúñéíé ìðáèòèèá ñíáðáíáííé ìèñúíáííó ìáíýóíèèáð, ñááíáíý ìæáò áúòú ìíááðáíóð «èíðíðèçáòèè», -òí ááçíáíáíðí-íí ìðááíñòáèèýáð çáíáí-éáúá áíçííæííñòè è òáðííéíáèè ñ òáèúþ ñíçááíèý òñéíáèè äèý ááí áàèúíáéøááí òóíèòèíèðíáíáíèý á ýéáèòðíííó ñèñòáíàð èáè á èññéááíáòáèúñéíé,

Íáóíáíéíáèý ðáçáèááþáèñý èíðíóñíé èèíáèñòèèè ìçáíèýáð ìíèó-áòú ìáèñèíáèúíí íáááæíúá ýçúèíáúá ðáèòú ðááèúííáí òóíèò èíðíóñá áóðýòñèèè òáèñòíá, íá èíòíðíí áíðíáèðòþòñý èííèíðááíñáðú [Grievous; Watt], ìíçáíèýþáèè ìíèó-áòú ìíèííèíáèñòíóþ èíðíóñá ýçúèíáúð ááèíèò.

Òðááèòèíííúá ìíáðíáú íá ìíçáíèýþò ìðíñíàððèááòú è áíáèèèèðíáòú áñá òáèñòú á áíñòáòí-íí ìíèíé ìáðá. Ááæá ìðè áúííéíáèè ñòàòúýð) ñòðááááò ìíðáááèáííé íáðáíè-áíííñòúþ (ááííáý íáðáíè-áíííñòú íè á èíáí ñèó-áá íá ýáèýáòñý òíðáèíí á áððáñ èññéááíá èííè-áñòááííí íèáíá, òáè è á èá-áñòááííí. Ííá ìíñéááíèíí ìú ìíðáçáíòáááá ðáíðáçáíòàòèáííñòú ñáííáí ñíèñéá èñòí-íéèíá, èíòíðúè è èðòííóð èññéááíáòáèèè ìí áóðýòñèííó è øèðá – ìííáíèúñéííó ýçúèíáíèþ, íáíðèíáð, á ðááíòáð Ò. Á. Òúááíááíáááá, Á. Á. Ý. Ð. Ðááíáááá, Ò.-Æ. Ø. Áííáóèíá, Ò. Á. Áóááááá, Ì. Í. Íðèíáñéí, Ñ. Ø. xáááòðíáá, Á. Á. Áèáèñáááá, Á. Á. Áííáíèííáá, Ø-Í. Ð. íáðèíáíá èáè ýáèýþáèñý íáèáíéáá ìíðááèýþáèèè ñòáííéáíéá è ìíñòóíáòáèúííá ðáçáèòèá ìííáíèíáááíèý á òáèíí è áóðýòíáááíèè

Íðè áñáè òóíááíáíòáèúííñòè òðóáíá ááííóð ááòíðíá ðáññíàððèáááíúé ñíèñíè èññèúçíááííóð èñòí-íéèíá äèý ýííèðè-áñéíáí íáðá áèáèáèòíèíáè-áñèèèè, èèáí èíúíè. Ññíí, -òí ááííúá ñíèñèè íá áóáòò ìíèííòúþ ñíáíááòú. Íáèííðíúá èñòí-íéèè áóáòò áñòðá-áòúñ èáèíí-òí íáííí ñíèñéá. Íáèè-éá áí áñáò íáðá-íýð èáèíáí-òí èñòí-íéèá ìíæáò ñòáòú èðèòáðèáí íáýçáòáèúííáí ááí áèèþ-áíèý á èíðíóñ. Áñèè íáúááèíèòú áñá íáðá-íè òáèñòíá, òéàçáííóð á ááííóð èññéááíáíèýð, òí, ááðíýòíí, òáèñòíáúá íáòáðèáèú òáèíáí íáíáúáííáí ñíèñéáò á òáèíí çá ñ-áò ááèíè-ííñòè «ìíèðúááþòñý» á òíè èèè èíé ñòáíáíè áñá ìíèý òóíèòèíèðíáíáíèý èññéááíáíáí á íèò áóðýòñ òíèúèíí ááííúí íáíáúáííúí ñíèñéíí òáèñòíá, ìíñèíèúèò á íáñòíýúáá áðáíý ìíñòíýííí ìíðíæááþòñý ííáúá òáèñòú, -òí ñáýçáíí áúá è ñ òáí, ááðñèè òáèñòíá íá áóðýòñèíí ýçúèá ááç ìðíáæòóí-íúð èð íóáèèèáòèè íá áóíáæíúð ìíñèòáèýð. Èñòíáý èç ýòíáí, áóíáæíúá ááðñè ýéáèòðíííí áèáá. Íááí ìòáòèòú, -òí ýéáèòðíííúá áóðýòñèèá òáèñòú ñòáèè ìíýáèýòúñý è íá èíííáèò-áèñéáò, ìðáááá, òáèíè òèí íóá ñèò ìíð íáèçááñòíóþ äèý íèò, ýéáèòðíííóþ æèçíú. È ííá íòèðúáááò ííáúá áíðèçííóð äèý íáó-ííáí ìíèñéá íá òíèúèíí èèíáèñòáí, íí è èí

Ííááíáíèáííúé íáíè ýéáèòðíííúé áððèá áóðýòñèèè òáèñòíá ýáèýáòñý á ñííáííí ìíèííóáèñòíáúí, -òí íáòñèíáèáíí òáððèòíðèáèúíú èíðíóñá, èíòíðúè áíèæáí ìððáæáòú ñííáííúá ñòèèè ñíáðáíáííáí áóðýòñèíáí èèòáðáòðííáí ýçúèá – òóáíæáñòááííí-èèòáðá. Ðáíðáçáíòàòèáííñòú èíðíóñá ñííáííáááòñý íá ìðááñòáèáííñòè ñíáðáíáííáí áóðýòñèíáí ýçúèá ðáçíúíè ìí æáíðó è ñòèèþ òáèñòáí ýéáèòðíííáí èíðíóñá áúííéíáíú á íáúáíá íáííáí íèèèííá èáèñè-áñèèò áòíæááíèè - ñèíáíòíðí. Òáðííéíáèè èíðíóñíé èèíáèñòèèè è çáí-èííñòè òðááèòèíííó èññéááíáíáíèè è ñíííáíñòáóý ðáñèððáíèþ èññéááíáòáèúñèèò áíçííæííñòáè, -òí á ñáíþ í-áðááú, ááç ñí èèíáèñòè-áñèèá ìíèñáíèý, ìíñòðíáííúá íá òáèòè-áñèíí íáðáðèáèá, áíáèèç òóíèòèíèðíáíáíèý èíòíðíáí ííñèò áèíááèúíúé òáðáèòáð

Òáðííéíáèè èíðíóñíé èèíáèñòèèè ìáíðááèáíú íá ñáíáíí ðíáá «áèíááèèèçáòèþ» èññéááíáíáíèý á ìðáááèò òíáí èèè èíáí èíðíóñ ñíííáíñòáòáò ìáèñèíáèúíííó áúýáèáíèþ ááí òáðáèòáðíèíáè-áñèèò -áðò, -òí ìíçáíèýýáð ñíòúáñòáèýòú íáèáíéáá áíñòíááðííá ááí ìíè

Íñíáííúè íáíðááèáíèýíè ðááíòú äèý ìðááíèçáòèè èíðíóñá áóðýòñèíáí ýçúèá á íáñòíýúáá áðáíý ýáèýþòñý ñèááòþúéá: ýèñòðáè ýèñòðáèèíáèñòè-áñèíè èíòíðíáòèè (íáòáááííúá: òíðíáèúííá ìíèñáíèá òáèñòá – ááóíð, íáçááíèá, áðáíý è íáñòí ñíçááíèý, CorpusTool. Íðè ðáðáíèè ç çááá-è áðáòáíàðè-áñèíáí áíáèèçá ìðááííèááááòñý ðáçðááíòèá ñííòááòñòáòþúááí ìðíáðáíííáí èíñòðóíá òíððááèáíèá á áóðýòñèíí ýçúèá íáðíáí ñííáííèíáíèý èáè íáííáí èç ñèíáííáðáçíáòáèúííó ñðááñòá. Áèý ìððíèíáè-áñèíè ðáçíáò ðáçíáòèè. Ñèááòáò ìòáòèòú, -òí èññéááíáíáíèý á ýòíí íáíðááèáíèè òíèúèíí íá-èíáþòñý. Íáíè èç ìðáááððèòáèúííó ðáçòèúòáòáí ý ñíçááíèý ìáðñáðá èíáííó ñèíáíòíðí. Íðáááèáí èíáííáòáðú áðáííáòè-áñèèò ìíáò (áðáííáí), èíòíðúè ìðááíáçíá-áí äèý èññèúçíááíèý òáèñòíá TEI. Áèý ñíçááíèý ñíáñòááííáí ìáíííæáñòáá TEI èññèúçóáòñý web-ìðíáðáííá TEI Pizza Chef.

Ááèáò òíáí, -òí ìðíáðáííú-èííèíðááíñáðú ìáðáíè-áíú á ñáíèò áíçííæííñòýð, äèý áíèáá ìíèííáí èññèúçíááíèý ááííóð èíðíóñá áóááò Manatee / Bonito, IMS Corpus Workbench (CWB), CQP/CQP-WS, Xaira (SARA).

