Ïîäõîäû ê ïîñòðîåíèþ äðåâíåðóññêîãî òåçàóðóñà (íà ìàòåðèàëàõ Ñëîâàðÿ ðóññêîãî ÿçûêà XI-

Àâòîð Îëüãà Âàñèëüåâíà Êîëîáåîâà
18.07.2008 ã.
Ïîñëåäíÿÿ ïðàâêà 20.07.2008 ã.

Ëþáàÿ ëåêñèêîãðàôè÷åñêàÿ ðàáîòà ñîïðÿæåíà ñ ïðèâëå÷åíèåì îãðîìíûõ ðåñóðñîâ âðåìåíè è òðóäà ëåêñèêîãðàôîâ. Ñîâðåìåííûå â
âûáîðà öèòàò äî ðåäàêòèðîâàíèÿ ñëîâàðÿ è åãî èçäàíèÿ. Â ýòîé ñâÿçè îñíîâíûì íàïðàâëåíèåì ïðåäñòîÿùåãî èññëåäîâàíèÿ áóäåò ïî
äðåâíåðóññêîãî ÿçûêà. Ðàáîòà áóäåò âåñòèñü íà áàçå ìàòåðèàëîâ Ñëîâàðÿ ðóññêîãî ÿçûêà XI–XVII ââ. (äàëåå ÑËÐß).

ÑËÐß ïî æàíðó èñòîðè÷åñêèé, îí îñíîâàí íà âûáîðêå ëåêñèêè èç ïàìÿòíèêîâ ïèñüìåííîñòè äðåâíå è ñòàðîðóññêîãî ïåðèîäà. Â ñïèñî
(ò. I–II. ÑÏá., 1858–1861 ãã.) è «Ìàòåðèàëû äëÿ Ñëîâàðÿ äðåâíåðóññêîãî ÿçûêà» È. È. Ñðåçíåâ-ñêîãî (ò. I–
õðîíîëîãè÷åñêèè ðàìêàìè, à äëÿ ëåêñèêè XV–XVII ââ. ïðèíàäëåæíîñòüþ ðóññêîìó ÿçûêó. Çà äâàäöàòü ëåò èçäàíî 27 âû
ëåêñè÷åñêîãî ìàòåðèàëà. Ñåé÷àñ ýòîò ñëîâàðü ïðåäñòàâëÿåòñÿ íàèáîëåå ïîëíûì èñòîðè÷åñêèì ñëîâàðåì ðóññêîãî ÿçûêà XI&nda
[1] 23–24: 3].

Ïîñêîëüêó ñåé÷àñ íå ñóùåñòâóåò òîìîâ, îõâàòûâàþùèõ îòðåçîê îò Ò äî ß, òî ïîñëåäóþùàÿ ðàçðàáîòêà äðåâíåðóññêîãî òåçàóðóñà

Âûáîð òåçàóðóñà êàê îñíîâíîãî ðåïðåçåíòàòîðà îáóñëîâëåí íåñêîëüêèìè ïðè÷èíàìè:

òåíäåíöèåé ê ëåêñèêîãðàôè÷åñêîìó ïðåäñòàâëåíèþ ëåêñè÷åñêîé èíôîðìàöèè;

àêòóàëüíîñòüþ èñïîëüçîâàíèÿ òåçàóðóñîâ â ïðàêòèêå;

íåîáõîäèìîñòüþ äîðàáîòêè ìàòåðèàëîâ íà íåäîñòàþùèå áóêâû.

Èòàê, îäíîé èç òåíäåíöèé ñîâðåìåííîé ëèíãâèñòèêè ÿâëÿåòñÿ ïðåäñòàâëåíèå äàííûõ î ÿçûêå â ëåêñèêîãðàôè÷åñêîé ôîðìå. Äóóãîé
ñòðåìëåíèå ê ôîðìàëèçàöèè ÿâëÿåòñÿ ñëåäñòâèåì âíóòðåííåé ëîãèêè ðàçâèòèÿ ëèíãâèñòèêè, ïðèøåäøåé ê íåîáõîäèìîñòè èíâåíò
âîñïðîèçâîäèìîñòè ðåçóëüòàòîâ. Ñ äóóãîé æå ñòîðîíû, òîëüêî ôîðìàëèçîâàííûé ìàòåðèàë ìîæåò áûòü îáðàáîòàí ìàøèííûìè ìåòîäàìè,
òåõíîêè, îòêðûâøèé äëÿ ëåêñèêîãðàôèè íîâóþ îáëàñòü êîðïóñíîé ëèíãâèñòèêè, ïîçâîëÿåò èçâëåêàòü îáøèðíûå ìàññèâû íåòðèâè

Òàêèì îáðàçîì, ìîæíî ïîðåäëèòü ìåòîäîëîãè÷åñêèå óñòàíîâêè èññëåäîâàíèÿ:

– îñíîâíîé åäèíèöåé èçó÷åíèÿ ÿâëÿåòñÿ ñëîâîôîðìà (ñëîâî), à îñíîâíîé ôîðìîé ïðåäñòàâëåíèÿ ïðàêòè÷åñêèõ ðåçóëüòàòîâ èñ

– ïðåäïî÷òèòåëüíûìè ìåòîäàìè èññëåäîâàíèÿ ÿâëÿþòñÿ òàêèå, êîòîðûå ïîäàþòñÿ ôîðìàëèçàöèè è àâòîìàòèçàöèè;

– èç âñåãî ìíîãîîáðàçèÿ ìàøèííûõ ìåòîäîâ ïðèîðèòåò ó ìåòîäîâ êîðïóñíîé ëèíãâèñòèêè.

Ðàññìîòðèì íåêîòîðûå ïîäõîäû ê ôîðìèðîâàíèþ ëèíãâèñòè÷åñêèõ áàç çíàíèé (äàëåå ËÁÇ), à èìåííî:

– ñòàòèñòè÷åñêèå;

– èñïîëüçóþùèå ýâðèñòèêè, ñïåöèô÷íûå äëÿ êîíêðåòíûõ ïðåäìåòíûõ îáëàñòåé;

– èñïîëüçóþùèå ÿâíóþ (ñîçäàííóþ àâòîðîì äîêóìåíòà) ñòðóêòóðó àíàëèçèðóåìîé èíôîðìàöèè;

– ëèíãâèñòè÷åñêèå (ãðàììàòè÷åñêèå).

Ñòàòèñòè÷åñêèå ìåòîäû îáû÷íî ïîäðàçóìåâàþò àíàëèç ñïåöèàëüíî ïîäîáðàííûõ òåêñòîâ ñ öåëüþ âûÿâëåíèÿ ñòàòèñòè÷åñêèõ õàðà
ìåòîäû èñïîëüçóþòñÿ â îñíîâíîì ïðè ôîðìèðîâàíèè ÷àñòîòíûõ êîìïîíåíòîâ ËÁÇ, òàêèõ, êàê ÷àñòîòíûå ñëîâàðè ñëîâîôîðì èëè ñëîâàðè ê

Ìåòîäû, êîòîðûå èñïîëüçóþò ðàçëè÷íûå ýâðèñòèêè, îòðàæàþùèå îñîáåííîñòè, ñïåöèô÷íûå äëÿ êîíêðåòíûõ ïðåäìåòíûõ îáëàñòåé è

Ìåòîäû, èñïîëüçóþùèå ñòðóêòóðíóþ ðàçìåòêó îáðàáàòûâàåìîãî òåêñòà, ìîæíî â ñâîþ î÷åðåäü ðàçäåëèòü íà äâå ïîäêàòåãîðèè. Âî-
òåêñòå (â çàãîëîâêå, â ïåðâîì àáçàöå è ò. ä.), ïðèìåñûâàÿ èì çàòåì ðàçëè÷íóþ çíà÷èìîñòü ïðè îáðàáîòêå. Âî-âòîðûõ, ýòî ìåòîäû, è

Ëèíãâèñòè÷åñêèå ìåòîäû ÿâëÿþòñÿ íàèáîëåå óíèâåðñàëüíûìè è ìîãóò èñïîëüçîâàòüñÿ ïðè ôîðìèðîâàíèè ñëîâàðíûõ êîìïîíåíòîâ ïðàêò
ïîëíîñòüþ àâòîìàòè÷åñêèõ ñèñòåìàõ, ïîýòîìó ñåãîäíÿ ó÷àñòèå ÷åëîâåêà-ýêñïåðòà ÿâëÿåòñÿ íåïðåìåííûì óñëîâèåì. Êðîìå òîãî, ïðèì

Ïðèìåíåíèå ëèíãâèñòè÷åñêèõ ìåòîäîâ ñîâìåñòíî ñî ñòàòèñòè÷åñêèìè ïîçâîëÿåò êîìïåíñèðîâàòü íåäîñòàòî÷íóþ ðåïðåçåíòàòèâíîñòü
ñòàòèñòè÷åñêèì ìåòîäàì, ñîõðàíèâ ïðè ýòîì èõ ïðåèìóùåñòâà (ýôôåêòèâíîñòü è ïðîñòîòó ðåàëèçàöèè), ÿâëÿþùèåñÿ åâñüìà âàæí

Ñôîðìóëèðóåì òðåáîâàíèÿ ê ïîäáîðíé ñèñòåìå.

Îáùèå òðåáîâàíèÿ. Ñèñòåìà äîëæíà ïîä÷èíÿòüñÿ ÿñíî ñôîðìóëèðîâàííé êîíöåïöèè, â êîòîðîé îòðàæàþòñÿ öåëè è çàäà÷è ñèñòåìû, è
óïðîùàþùèé äîñòóï ê èíôîðìàöèè. Â àðõèòåêòóðíîìêê ñèñòåìû ñëåäóåò ïðåäóñìîòðåòü äà çâèòèå åå èíôîðìàöèîííîé áàçû

Ê èíôîðìàöèîííûì òðåáîâàíèÿì ïðèíàäëåæàò: ñèñòåìíîñòü êîìïëåêòîâàíèÿ, èñêëþ÷àþùàÿ ïðîèçâîë è ñòèõèéíîñòü â îòáîðå èíôîðìàöèè
èäåíòèôèöèðóåìîñòü èíôîðìàöèîííûõ îáúåêòîâ; òî÷íîñòü âîñïðîèçâåäåíèÿ èíôîðìàöèè (ïðè ýëåêòðîííîì âîñïðîèçâåäåíèè ïå÷àòíîãî òåê
îôîðìëåíèå è ðÿä äóãèõ îñîáåííîñòåé).

Ôóíêöèîíàëüíûå òðåáîâàíèÿ ïðåäïîëàãàþò íàëè÷èå ðàçâèòûõ ôóíêöèîíàëüíûõ âîçìîæíîñòåé, ê ÷èñëó êîòîðûõ îòíîñÿòñÿ: íàâèãàöèÿ
àòðèáóòíûé ïîèñê (òî åñòü ïîèñê èíôîðìàöèîííûõ îáúåêòîâ ïî çíà÷åíèÿì èõ ôîðìàëüíûõ õàðàêòåðèñòèê); êîìáèíèðîâàííûå ïîèñêè; ñîðò
ïîñëåäîâàòåëüíûé (ñòðàíèöà çà ñòðàíèöåé), âûâîð÷íûé (ïåðåõîä íà çàäàííóþ ñòðàíèöó), ïàðàëëåëüíûé (îäíîâðåìåííûé ïðîñìîòð íåñ

Summary

The solutions of elaboration of thesauruses and linguistic knowledge bases are illustrated in the article. The texts of the Russian Language dictionary of XI–XVII centuries were used in the work as a source of terminological information. The system requirements and the methodological goals of investigation were also formulated.

Ñïèñîê ëèòåðàòóðû

Èíñòðóêöèÿ 1988 Èíñòðóêöèÿ äëÿ ñîñòàâèòåëåé Ñëîâàðÿ ðóññêîãî ÿçûêà XI–XVII ââ. Ì., 1988.

Ãåíäèíà 1991 Ãåíäèíà, Í. È. Ëèíãâèñòè÷åñêîå îáåñïå÷åíèå àâòîìàòèçèðîâàííûõ áèáëèîòå÷íûõ ñèñòåì / Í. È. Ãåíäèíà. Àëìà-Àòà,

Êðàòêèé 1995 Êðàòêèé ñëîâàðü ëèíãâèñòè÷åñêèõ òåðìèíîâ. — Ì., 1995.

Íèêèòèíà 1978 Íèêèòèíà, Ñ. Å. Òåçàóðóñ ïî òåîðåòè÷åñêîé è ïðèêëàäíîé ëèíãâèñòèêå: (Àâòîìàòè÷åñêàÿ îáðàáîòêà òåêñòà) / Ñ

Ñëîâàðü 1975 Ñëîâàðü ðóññêîãî ÿçûêà XI–XVII ââ. Âûï. 1 (À–Á) / Ñ. Ã. Áàðõóäàðîâ. 1975.

Ñëîâàðü 1976. Ñëîâàðü ðóññêîãî ÿçûêà XI–XVII ââ. Âûï. 3 (ÂÎËÎÄÅÍÜÅ–ÂßÙÜØÈÍÀ) / Ñ. Ã. Áàðõóäàðîâ. 1976

Äåðæàâèíà 1997 Äåðæàâèíà, Å. È. Èñòîðèêî-ëåêñèêîãðàôè÷åñêèå èññëåäîâàíèÿ è êîìïþòåð (íà ìàòåðèàëå Ñëîâàðÿ ðóññêîã
[1] 23–24.