

Ebloj por elektronika prilaborado de grandaj vortaroj de Esperanto

Unu el kondiĉoj por plua disvolvo de Esperanto, precipe por ĝia uzado en faka sfero estas i.a. ekzisto de bonaj kaj sufiĉe enhavaj vortaroj difinaj kaj ankaŭ tradukaj esperanta-nacilingvaj kaj nacilingva-esperantaj. Ili estas gravaj ankaŭ por pruvo, ke Esperanto estas uzebla kiel pontlingvo en unuiĝanta Eŭropo kaj ke ĝi povas esti solvo de la lingva problemaro en Eŭropa Unio. Pro tio estas granda handikapo de Esperanto, ke ĝi ĝis nun ne disponis per granda vortaro, t.e. la vortaro enhavanta minimume 100 mil da vortoj.

Per kalkulo de la vortoj en ĉiu dudeka paĝo mi difinis proksimumajn nombrojn da vortoj en kelkaj signifaj esperantaj vortaroj. Ĉeĥa-esperanta vortaro de Karel Kraft enhavas proksimume 52 mil da vortoj, Nova PIV ĉirkaŭ 50 mil da vortoj kaj Esperanto-germana vortaro de Prof. Krause proksimume 80 mil da vortoj. Esperanta-rusa vortaro de E. Bokarev enhavas ĉ. 35 mil da vortoj.

Esperanto bedaŭrinde ne havas oficialan difinan vortaron, per kiu disponas plimulto da naciaj lingvoj. Oficiala esperanta leksikografia dokumento estas Akademia vortaro. Sed, kiel vi scias, tiu estas nur listo de elektitaj vortradikoj al kiu tute mankas difina parto. Ununura nova mezgranda difina vortaro de Esperanto estas NPIV, kiu pro tio estas de plimulto da esperantistoj konsiderata kaj prezentata oficiala, malgaŭ ĝi tian statuton ne havas kaj kiu en kelkaj kazoj eĉ ne estas identa al Akademia vortaro. Ankaŭ per la enhavo de menciitaj 50 mil da vortoj ĝi ne estas komparebla kun oficialaj grandaj nacilingvaj vortaroj. Ekzemple la plej granda ĉeĥa difina vortaro, t.n. Ĉemana Vortaro de la Lingvo Ĉeĥa aperadis de 1935 kaj estis finita de Instituto por la Lingvo Ĉeĥa en 1957. Ĝi havas 9 volumojn kaj enhavas 250 mil da vortoj. La aŭtora kolekto estis post apero de la vortaro aprezita per la ŝtata prezo. La difinaj vortaroj de iuj eŭropaj nacioj estas onidire ankoraŭ pli grandaj. Tamen unuiĝanta signifo de NPIV por Esperanto estas nedubebla kaj ni ĉiuj certe aprezas ĝian ekziston.

Estas vero, ke la mezgrandaj vortaroj por kutima laboro en Esperanto sufiĉas. Sed tio ne validas por laboro scienca kaj faka entute, por preciza eksplikado de juraj kaj politikaj rilatoj. En ĉi tiuj kazoj oni nepre bezonas, ke ĉiu vorto, ĉiu faka termino havu sian ununuran oficialan signifon. Do ne sufiĉas, ke ekzemple en kazo de la internacia jura konflikto oni decidus laŭ vortaro de la plimulto da esperantistoj konsiderata oficiala.

Kiaj estas kaŭzoj de tiu ĉi nekonsolebla stato? Eble ili estas pli multaj, sed ni nomu almenaŭ tiujn ĉi:

1. Esperanto ne havas oficialan leksikografian instituton, komisiitan per prilaboro de oficiala difina vortaro. La Akademio, kiu estas al plenumdevo de tiu ĉi tasko plej proksima, ne plenumas ĝin, eĉ ne konsideras sin

kompetenta por plenumado (*pri tiu ĉi temo vidu artikolon de Wera Blanke: Terminologia Esperanto-Centro: spertoj, problemoj, perspektivoj*).

2. Laboro de la vortaristoj kaj terminaristoj ne estas iel gvidata, oni ne organizas iujn internaciajn renkontiĝojn kaj interŝanĝoj de la spertoj praktike ne ekzistas.
3. Metodoj de la laboro pri la vortaroj kaj terminaroj estas malnovigitaj. La komputilo estas uzata plejparte nur kiel skribomaŝino kaj tipografa maŝino. Libre eldonataj vortaroj estas multekostaj, grandaj tradukaj vortaroj ne povas esti pagataj de esperantistoj de unu nacio. Libraj vortaroj ankaŭ ne ebligas efikan vastan diskuton kaj interŝanĝon de opinioj pri la enhavo, pri diskutindaj esprimoj kaj terminoj, ĉar eventualaj ŝanĝoj aŭ korektoj de la vortaro laŭ rezultoj de la diskuto estas praktike neeblaj. Ni vidas tion ekzemple je NPIV. La notojn oni publikas je interretaj paĝoj, eble kelkaj esperantistoj laŭ ili kaj laŭ sia opinio korektas siajn ekzemplerojn, sed amaso de la vortaroj restas tia kiaj ili estis presitaj.

Granda espero por solvo de tiu ĉi problemaro estas vortaroj elektronikaj. Elspezoj por aĉeto de la datportilo, kopiado kaj distribuo de la elektronika vortaro estas nekompareble malpligrandaj ol elspezoj por aĉeto de papero, preso, bindo kaj distribuo de la papera vortaro. Krome elektronika vortaro povas esti eldonata en malgrandaj eldonkvantoj iom post iom laŭ forvendado. Ĉe tio oni povas senpere fari korektojn kaj kompletaĵojn iniciatitaj aŭ de la aŭtoro mem aŭ laŭ atentigoj de la uzantoj. Tio ebligas i.a. ĝis nun nekutiman kunlaboron de la uzantoj kun la aŭtoro aŭ aŭtoro kolektivo.

Por kreado kaj uzado de la elektronikaj vortaroj estas grava elekto de taŭga programaro. Pro tio mi ŝatus prezenti al vi ĝisnunajn rezultojn de mia laboro pri elektronika vortaro, ĉefe el vidpunkto de uzata programaro.

Kiel multaj el vi scias, mi jam dum kelkaj jaroj laboras pri elektronika „Granda Vortaro Esperanta-Ĉeĥa“ en ĝenerale konata tabela programo Microsoft-Excel. Por kreado de la vortaro mi elektis tiun ĉi programon pro tio, ke mi bone konis ĝin el mia okupo, kie mi en ĝi prilaboradis diversajn grandajn analizajn kaj planumajn databazajn tabelojn. Komence mi el leksikografia vidpunkto je la programo aprezadis ĉefe supervideblon de la tabela aranĝo kaj eblecon de alfanumera vicordigado de la enmetitaj esperantaj kaj ĉeĥaj vortoj. Sed sinsekve mi aperigadis kaj ĝis nun mi aperigas pluajn avantaĝojn de tiu ĉi programo.

Por kreado de la Esperanta-ĉeĥa vortaro mi starigis al si celon kolekti kiel eble plej grandan provizon de la esperantaj vortoj, inkluzive de fakaj terminoj kaj alporti ilian tradukon al la ĉeĥa. La vortaro estas destinita prefere por ĉeĥoj, do ĝi ne enhavas gramatikajn datojn por ĉeĥaj vortoj, sed male enhavas gramatikajn datojn por esperantaj vortoj, nome por ĉiu vorto la vortradikon kaj al la verboj iliajn transitivecojn. La vortaro penas pri detala signado de aparto de la vortetoj al unuopaj fakoj, por ke la vortaro povus esti laŭ tiuj signoj ordigata kaj povus

servi kiel bazo por kreado de la fakvortaroj al koncernaj fakuloj. Al malmulte konataj vortoj estas alligitaj signifklarigoj en la ĉeĥa. La radikvortoj estas presitaj per grasaj literoj kaj oni povas lokigi ilin ĉiam al la komenco de la koncerna vortnesto. Al multaj vortoj estas alligitaj atentigoj je aliradikaj sinonimoj kaj parencaj vortoj. Dum kreado de la vortaro mi uzas, nur por mia labora bezono, ankaŭ pluajn datojn, ekzemple pri fontoj de kiuj mi prenis unuopajn vortetojn, atentigoj je gramatiko, en kiu estas la vorto pli detale pritraktata, signo ĉu la radikoj de la vorto estas en Akademia vortaro k.s. Esperantaj kaj ĉeĥaj vortetoj kaj kompletigaj datoj estas en la tabelo metitaj tiel, ke ĉiu dato okupas unu ĉelon de la tabelo. La ĉeloj apartenataj al unu esperanta vorto estas ordigitaj en unu linio kaj ĉiu dato-tipo en unu kolumno. Por versio preparata al prezentado do estas kelkaj dat-kolumnoj forigitaj kaj male kelkaj kolumnoj aldonataj, ekzemple por facila komforta kaj rapida serĉado de la vortoj estas aldonitaj kolumnoj kun esperantaj vortetoj en X-sistemo kaj kun antaŭmetitaj kaj postmetitaj punktoj kaj kolumnoj kun ĉeĥaj vortoj kun enmetitaj antaŭ- kaj post-komoj. Kiam ni metas la punktojn aŭ komojn en la serĉilon, la serĉilo proponas nur la vortojn kun traduko kaj ne tiujn uzitajn en aliaj nekonvenaj ĉeloj, ekzemple kun uzekzemploj, kiujn ni povas serĉi helpe de antaŭmetita kaj postmetita punktokomo.

Al ĉiu vorteto do apartenas kelkaj datoj, respektive ĉeloj. Per la vortaro, kies vortprovizo jam proksimiĝas al 130 mil da vortoj estas okupita pli ol miliono da ĉeloj. La programo kaj mia ne tro moderna komputilo regas tiun ĉi kvanton da datoj sen problemoj. Mi eĉ provis enmeti ĉ. 200 mil da vortetoj, la programo funkciis ankaŭ sen problemoj.

Por pli detala priskribo de la laboro pri la vortaro ne estas ĉi tie sufiĉa loko kaj kompreneble ankaŭ ne ĉiujn aŭskultantojn la detaloj interesus. Principe oni povas diri, ke mi je la vortaro laboras per kutimaj datumbazaj metodoj. Mi do mencias nur kelkajn eblecojn por eluzo de tiuj ĉi metodoj por leksokografia laboro.

Krome jam menciitaj avantaĝoj de la programo Excel, do facila regado de grandaj kvantoj da datoj, mi jam menciis ankaŭ pri senproblema alfanumera vicigado de datoj. Se estas enmetitaj bezonataj klasigaj datoj, oni povas la vortordon laŭbezone ŝanĝi, ekzemple al

- simpla alfabeto vicordo de la vortoj,
- alfabeto vicordo laŭ vortradikoj kaj interne de ĉiu radiknesto ankaŭ la vortojn ordigi alfabeto, eventuale kun la radikvorto je la unua loko,
- vicordigo de la terminoj laŭ la fakoj kaj interne de ĉiu fakoj ree alfabeto,
- kunigi kelkajn fakojn, ekzemple anatomion, medicinon, farmacion kaj krei pli vastan fakvortaron.

Pluaj ebloj, kiujn Excel proponas al uzantoj estas ekzemple

- facila kopiado de datoj el teksta formato Word kaj de datoj el databazaj sistemoj,
- facila kopiado de datoj al teksta formato Word kaj al databazaj sistemoj,
- ebloj de diversaj grafikaj aranĝoj de la teksto en la ĉeloj, nome elekto de tiparo, stilo, granda, koloro, lokigo kaj aranĝo de la teksto kaj elekto de koloro de surfaco de la ĉeloj,
- ebloj starigi larĝon de la kolumnoj kaj altecon de la linioj, eventuale povu tute kaŝi kelkajn el ili,
- tekston el unu ĉelo oni povas dividi diversmaniere en pli da ĉeloj kaj male enhavon de pli da ĉeloj eblas kunigi en unu ĉelon k.t.p.

Kun eluzo de la nomitaj paŝoj eblas kompletigi diversajn procedurojn por aŭtomata laboro aŭ por faciligo de neaŭtomata laboro. Eblas ekzemple:

1. prijuĝi kolekton de esperantaj vortoj, ĉu unuopaj vortoj, kiujn ĝi ehhavas jam estas aŭ ne estas en la kreata vortaro kaj laŭdezire signigi ilin, eventuale enmeti ilin aŭ ne en la vortaron,
2. fari facile rekonstruon de la vortaro je vortaro de alia speco aŭ alia formo, ekzemple, kiel mi jam menciis, ne estas granda problemo el la vortaro krei fakan vortaron aŭ specialan kolekton de la vortoj, ekzemple por kontrolado de gramatika korekto de iu esperanta teksto en programo Word,
3. relative facile oni povas fari ekscerpcion de la vortoj, kiuj kreas certan tekston. En naciaj lingvoj estis ĝis nun pro granda kaj enuiga laborpostulo faritaj ekscerpcoj nur el malmultaj tekstoj. En Esperanto pro ĝia reguleco per uzo de komputila tekniko estas ekscerpco relative facila. Ekzemple ekscerpcion de la uzitaj vortoj el esperanta traduko de la romano de Karel Čapek „Milito kontraŭ salamandroj“ mi per programoj Word kaj Excel faris dum ĉ. tri horoj kaj trovis, ke en la romano estis uzitaj ĉ. 70 mil da vortoj kaj el tiuj ĉ. 13 mil da diversaj vortoj. Provizore por kreado de la GVEĈ ne minacas elĉerpo de vortprovizo jam prilaborita en diversaj aliaj vortaroj kaj terminaroj. Tamen eble foje la vortprovizo de la vortaro kreskos tiel, ke serĉado de pluaj vortoj eblos nur per ekscerpcado de esperantaj tekstoj. En tiu ĉi tempo ekscerpcado de fakaj tekstoj povas esti signifa fonto por terminologia laboro,
4. fari eventualan rekonstruon de la esperanta-ĉeĥa vortaro je vortaro ĉeĥa-esperanta mi ankoraŭ ne provis, sed mi kredas, ke en programo Excel ĝi estas farebla multe pli facile ol per tradiciaj metodoj

Kompreneble Excel ne estis kreita por prilaborado de la vortaroj kaj des malpli de la vortaroj esperantaj. Pro tio por prilaborado de la vortaro ni devas en iuj kazoj ankaŭ iom akomodiĝi al la programo. Tiuj akomodiĝoj tamen ne estas nesupereblaj rilate al avantaĝoj kiujn la programo proponas. Ekzemple

- Excel ankoraŭ ne estas prilaborita por rekta laboro per esperantaj literoj. Ni povas por enskribado de esperantaj literoj uzi iun helpprogramon, ekzemple programon „Ek“. Bedaŭrinde la vicordigado kun supersignitaj literoj ne funkcias bone. Pro tio dum laboro je la vortaro ni devas uzi x-dismetitajn literojn kaj nur post la laboro je la vortaro ni laŭbezone povas facile ŝanĝi la dismetitajn literojn al esperantaj,
- alia problemo estas, ke Excel enhavas nur 65 mil da linioj. Por granda vortaro ni bezonus kelkoble pli da ili. La problemon ni povas ĉirkaŭiri per tio, ke ni laboros en pli da dosieroj, kaj en definitivan vortaron ni metos pli grupojn de la kolumnoj apude,
- same, kiel en vortaroj prilaboritaj en aliaj programoj, por ke serĉado de la vortoj estus korekta, oni ne povas uzadi diversajn mallongaĵojn, kutimaj en paperaj vortaroj. Tio estas kompreneble komplikaĵo por la vortaristo. Sed por la uzantoj de la vortaro tio tamen estas avantaĝo, ĉar ili ne devas pensi pri senco de la mallongigo.

En esperanta medio per eluzo de la elektronika formo, speciale de Excel, oni povas krei kondiĉojn por laboro je granda internacia difina kaj traduka vortarkolekto. Konkrete tion oni povas fari tiamaniere, ke surbaze de ĝis nun prilaboritaj esperanto-naciaj elektronikaj ĝeneralaj kaj fakaj vortaroj estus prilaborita komuna maldekstra, esperanta parto de la granda vortaro. Al tiu diversnaciaj esperantistoj kompletigus laŭ siaj ebloj sinsekve la dekstrajn partojn en siaj naciaj lingvoj. Per tio estus plenumita principo de prilaboro de la vortaro per denaska parolanto.

Krome povus esti prilaborata kaj iom post iom kompletata difina parto de la difina vortaro al unuopaj esperantaj vortoj kaj terminoj. Eblus komenci, se la aŭtoroj kaj la eldonisto konsentus, per eluzo de NPIV. Estus tiel metita bazo por kreado de oficiala difina vortaro de Esperanto.

Per tia internacia kunlaboro ekestus kolekto de interligitaj elektronikaj vortaroj, kiu super sia normala uzovaloro ankaŭ pruvus eksterordinaran signifon de Esperanto kiel pontlingvo. La vortaroj povus esti plu kompletigataj kaj korektataj laŭ evoluo de la lingvoj. Ekestus tiel ne nur la vortaro, sed ankaŭ vasta internacia esperantista aŭtora kolektivo, kiu povus fariĝi bazo por konstituo de la oficiala leksikografia esperanto-instituto. Kompreneble estus plu bonvenata ankaŭ kunlaboro de pliaj uzantoj, ĉefe atentigado je mankoj kaj en la vortaro mankantaj vortoj.

Por tia kreado de la grandaj esperantaj vortaroj, eventuale de la granda esperanta internacia vortarkolekto, mi konsideras granda avantaĝo de Excel ĝian disvastecon en la tuta mondo. Ĉar Excel estas vendata kaj kutime instalata kiel parto de programara pakego Microsoft-Office, kiel ankaŭ konata teksta programo Word, havas Excelon multaj esperantistoj jam instalitan en sia komputilo, malgraŭ multaj el ili eble ne uzas ĝin. Krome la programara pakego

Office estas en la tuta mondo vendata jam priaranĝita por koncernaj naciaj kondiĉoj, respektive eblas ĝin por koncernaj kondiĉoj starigi. Excel estas ĉiam plu evoluigata, ĝiaj funkcioj estas plibonigataj. Ekzistas ankaŭ rigarda versio de la programo Excel, libere senpage atingebla el interreto, kompreneble nur por pasiva uzo.

Do ni povas veni al konkludo:

1. Esperanto nepre bezonas, ke estu kreitaj oficialaj grandaj vortaroj, difina esperanta kaj tradukaj esperanta-nacilingvaj kaj nacilingva-esperantaj,
2. per eluzo de ebleco krei elektronikajn vortarojn oni povas superigi ĉefajn ĝisnunajn problemojn, pro kiuj ne estis grandaj esperantaj vortaroj ĝis nun kreitaj,
3. per taŭga programo oni povas krei kondiĉojn por prilaboro de unika internacie kunigita sistemo de la grandaj esperantaj vortaroj,
4. pro ĝenerala disvasteco kaj teknikaj eblecoj oni rekomendas uzi por tiu ĉi laboro tabelan programon Microsoft-Excel.

Ing. Josef Hron, teknikisto por akvomastumado ĉe regiona oficejo kaj ĉe Entrepreno de akvokondukiloj kaj kanalizacioj en Hradec Králové kaj Pardubice, nun emeritulo.
Esperantisto de 1989, vicprezidanto de Asocio de Esperantistoj Handikapuloj, prezidanto de Inspekta Komisiono de ĈEA.