

**KOUTNY Iona**

## **Esperanta leksikografio en la epoko de komputiloj**

Antaŭ 25 jaroj mi komencis verki artikolon pri Esperanto kaj komputiko por la memorlibro, kiu devintus aperi por la 60-jariĝo de prof. Szerdahelyi. Tio estis adekvata temo por mi, ĉar Esperanto jam tiam plene trovis sian lokon en komputiko kaj ankoraŭ eblis trarigardi la esplorojn, aplikojn, cetere tiam mi estis aktiva komputisto (vd. parolsintezan sistemon Esperol, partoprenon en maŝina tradukado de DLT, Koutny 1985, ktp.). Tio estis la heroa tempo de la hungara komputista fakgrupo (Péter Brockó, Iván Bujdosó, Mandler György, Szőke László por mencii nur kelkajn), kiu regule kunvenis kaj okupiĝis serioze ankaŭ pri terminologio.

Nuntempe, la komputiko tiomgrade enpenetris nian vivon, estas tiom da terenoj, ke ne eblas en unu mallonga studo prezenti la lokon de Esperanto, krome ankaŭ mi el aktiva komputisto iĝis uzanto de komputilo. Mi elektis skize prezenti la rolon de komputiloj en la esperanta leksikografio.

### **Historie – pretigo de la Hungara-Esperanta meza vortaro**

La ellaboro de la Hungara-Esperanta meza vortaro – kiu ŝajnis esti pli urĝa, kvankam estintus pli motivita komenci per la alia direkto – komenciĝis kun la gvido de prof. Szerdahelyi en 1983. Kiel tiutempe kutime, ni – kvideko da kunlaborantoj – komencis la laboron uzante slipojn por la unuopaj vortartikoloj. La ĉefredaktoro ricevis fine grandan amason da slipoj, kiujn li kunigis, ordigis kaj redaktis. Temis pri ĉ. 40 skatoloj!

Personaj komputiloj ne estis ankoraŭ disvastiĝintaj, tamen la Hungara Esperanto-Asocio jam ekhavis sian. La slipoj atendis enkomputiligon kaj prilaboron (pluraj laboris pri tio, i.a. Iona Sirkó, Lin Wang, Sándor Hideg, Iván Bujdosó), mi kiel sekretario de la vortara komisiono organizis la laboron. Estis longa vojo kaj multaj praktikaj problemoj ĝis la enigo kaj prilaboro de la tuta materialo, ni devis pagi la prezon de moderneco. La Hungara-Esperanta meza vortaro povintus esti la unua komputile pretigita vortaro en Hungario, se ĝi estus preta en 1987 kiel planite.

Sed en la lasta fazo de la redaktado István Szerdahelyi mortis, mi ekokupiĝis pri la taskoj de la Esperanto-fako... Multo ŝaĝiĝis ĉirkaŭ ni en 1989: La Akademia eldonejo, kiu ĝis tiam eldonis ĉiujn vortarojn, kaj kun kiu entreprenis la eldonon de la e-vortaro, ne plu pretis eldoni ĝin sen monfonto. La vortaro, preta komence de la 90-aj jaroj, povis aperi fine nur en 1996 ĉe HEA. (Pri principoj de la vortaro vd. Koutny 1997.)

Grupigi vortarojn eblas laŭ pluraj kriterioj kiel enhavo, amplekso, ordigo, nombro da lingvoj, speco da informoj, celgrupo aŭ formato (koncerne esperantaj vortarojn vd. Blanke 1996, Koutny 2001). Ĉi tie mi grupigos la vortarojn laŭ la formo: papera, elektronika aŭ reta kaj prezentos la komputikajn eblojn por la uzantoj kaj vortaristoj.

## Paperaj vortaroj

Vortaroj, kiel ĉiuj libroj, pretiĝas nuntempe per komputilo. Krom la tesktoredaktaj funkcioj, ili ebligas alfabetan ordigon, unuecigon de la materialo en have kaj forme. Dum la ellaborfazo diversaj listoj (ekz. laŭ etiketoj fakaj aŭ gramatikaj) estas fareblaj por helpi la laboron de leksikografoj. Povas okazi ankaŭ ortografia kontrolo de la materialo, se la konvenaj programoj ekzistas. La plejmulto de e-vortaroj estas paperaj.

Dum la pretigo de miaj laŭtemaj etvortaroj (ekz. Koutny 2002), kiu konstruiĝas hierarkie, la aperigo de strukturitaj listoj kontribuis al la elekto de konvenaj ĉefaj kapvortoj kaj ilia proporcia kompletigo per aldonaj vortoj, esprimoj (Koutny 2005).

## Elektronikaj (komputile legeblaj) vortaroj sur kompakta disko

Jam pli kaj pli ofte vortaroj sur KD estas aldonaj al la papera vortaro aŭ eĉ anstataŭas ĝin. Kvankam ili baziĝas principe sur papera vortaro, tamen la komputila / elektronika versio ebligas multe pli ol la papera vortaro al la uzanto:

- La **retrovo** de serĉata vorto dum uzo de komputilo, ekz. tradukado, pli rapidigas la laboron, des pli se eblas transpreni la trovitan vorton per musa indiko.
- Eblas trovi serĉatan elementon ankaŭ **en la teksto** de la vortaro, tiel la konvencia alfabeto limigo malaperas, retroveblas la vorto en diversaj kombinaĵoj, ekzemploj, kiel parto de alia esprimo.
- Iusence komputila vortaro funkcias kiel kontraŭ-direkta vortaro pro la retrovebleco de la celvorto ne nur kiel kapvorto, sed ie en la teksto. Kompreneble, tio ne anstataŭigas la inversdirektan leksikografian laboron, sed donas ideon al rutina uzanto.
- Eblas trovi **laŭ fakaj aŭ aliaj etiketoj** grupon de vortoj.
- Se la vortoj estas ligitaj, tiam eblas per muŝokliko tuj transiri al la vortartikolo de la indikita vorto alie en la vortaro.
- La komputila vortaro estas pli simple ŝanĝebla, **aktualigebla** kompare kun papera vortaro kaj la produktadkostoj estas malpli altaj.
- **Aldona materialo** povas troviĝi sur la disko (ne plu estas spacolimigo), kelkfoje programo por ekzerci, eventuala ebleco por la uzanto krei sian propran aldonon vortaron.
- Realigeblas **plurmedia informaro**: ekz. voĉa elparolo de la vortoj (ofta por la angla lingvo, male en esperanto), prezento de bildoj.
- Ev. ebleco ligiĝi per ĝi al la reto kaj akiri novajn informojn.

Ankaŭ la *Hungara-Esperanta meza vortaro* havas sian version sur disko kun aldonaj tekstaj materialoj, sed la aktualigo forestas.

Fakvortaroj sur KD *MultiDic* de Christian Bertin el 2000 (<http://bertin.ifrance.com/catal/multidic/multidic.htm>) entenas terminojn el 56 terenoj, parto estas havebla rete.

Tamen ankaŭ multaj aliaj vortaroj aperas daŭre paperforme, nur menciis la grandajn vortarojn germanan de Krause, italan de Minnaja, hispanan De Diego.

## Retaj vortaroj – kiel produktoj

Al tiu ĉi kategorio mi klasifikas vortarojn, kiuj havas sian aŭtoron aŭ kompilan stabon, kaj ofte post leksikografia laboro en matura stato estas disponigitaj al la uzantoj. La limo inter la antaŭa kategorio kaj la nuna ne estas akra. Se la vortaro jam ekzistas elektronike kun funkcioj por la uzantoj, tiam ĝi povas aperi ankaŭ en la reto kaj laŭdeziro uzeblas tie aŭ elŝuteblas ofte. Tio povas tamen nur tiam okazi, se la uzo de la vortaro estas senpaga aŭ estas publikigita tiunmaniere, ke nur per pasvorto kaj pago estas alirebla. La retaj vortaroj disvastiĝas pli kaj pli.

- Tiu tipo ebligas oftan aktualigon, kompletigon de la materialo.

Al tiu kategorio apartenas *La Vortaro* kun siaj 29 lingvoj (<http://pagesperso-orange.fr/la-vortaro/>), kie oni konas la ellaboranton de diverslingvaj variantoj kaj la uzantoj ricevas informojn pri la novaj vortaroj aŭ la aktualigo de iu aŭ alia vortaro. Pri tiu skribas Jozefo Horvath en sia artikolo ĉi tie. Li estas la aŭtoro de la Esperanta-Hungara vortaro ene de *La Vortaro*, kaj ankaŭ de la pli kompleksa propra E-H-a vortaro, kiu estas rete alirebla kaj elŝutebla.

En la reto troveblas fakvortaroj kaj terminaroj, ekz. en la grava kolekto de STEB (Scienca kaj teknika Esperanto-Biblioteko: <http://www.amfibi.com/cached?idx=0&id=896367>) de asecura terminaro tra insekta kaj sporta ĝis vina terminaro, *Komputika Leksikono* de Pokrovskij kaj multaj aliaj.

La Baza radikaro (ĉ. 2500 radikoj kaj 25000 vortoj) troviĝas ĉe Wouter F. Pilger, la Akademia Vortaro ([http://akademio-de-esperanto.org/akademia\\_vortaro/](http://akademio-de-esperanto.org/akademia_vortaro/)) entenas la komencan staton de tiu vortaro. ktp.

En la reto povas aperi – krom la utilaj kaj fidindaj vortaroj – ankaŭ nefidindaj aŭ malgrandaj vortlistoj. Estas vere – kiel koncerne ankaŭ aliajn retajn informojn – ke la reto estas samtempe trezorejo kaj rubujo, oni devas esti singarda. Estas multe pli da nematuraj verkoj ol paperforme, ĉar la aperigo en la reto ne kostas.

## Reta Vortaro – disvolviĝanta

Reta vortaro laŭ la tipo de **ReVo** (<http://www.reta-vortaro.de/revo/>) kaj ties terminologia versio ReTo estas alia kategorio, ĉar tie la ellaboro okazas kolektive kaj anonime, do baze ŝanĝiĝas la leksikografia laboro – temas pri kunlaboro, pli demokratia kun siaj avantaĝoj kaj malavantaĝoj.

Tamen povas aperi nematuraj aferoj aŭ ne proporciaj, ekzemple la tradukoj de leksikaj unuoj al diversaj lingvoj dependas de la disponebleco de la koncerna nacilingvulo. La daŭra kontrolo de diverslingvanoj ĉiam plibonigas la vortaron. Tia vortaro estas en konstanta ŝanĝiĝo.

ReVo entenas 77 fakajn terenojn, simile etikeditajn kiel en PIV (NPIV 65). Lastatempe ĝi proponas ankaŭ laŭfakajn retrovojn kaj ĉefajn nociojn por la donita tereno. Ekz. pedagogio entenas 84 elementojn kaj tri ĉefnociojn: *instruisto*, *kurso* kaj *lernejo*. Por komparo: la *Etvortaro pri Lernado kaj Laboro* (Koutny 2002) enhavas ĉ. 1200 leksikajn unuojn el tiu tereno en 4 subĉapitroj kaj 60 ĉefnocioj kiel kapvortoj.

## Uzo de grandaj Tekstaroj

Nuntempe leksikografia laboro devas baziĝi sur granda tekstaro, t.n. *korpuso*, kiu kompletigas la personan scion, inventemon kaj plurjaran vortokolektadon de leksikografo per utilaj informoj pri la uzkonteksto de leksikaj elementoj. Tio speciale estas vera kaze de esperanto, kiu kutime estas lernita lingvo de la leksikografo, tial la vortostoko estas pli limigita ol la gepatralingva, la signifokampojn kaj la uzon de vortoj certagrade influas la gepatra lingvo.

Nur tre granda tekstaro aŭ en limigita kampo, reprezente elektitaj tekstoj povas doni fidindajn rezultojn, ĉiukaze la leksikografo decidas (pri postuloj kaj problemoj por e-tekstaro kaj entute pri uzo de korpusoj vd Koutny 2003). Menciendas, ke jam en la 90-aj jaroj Hitosi Gotoo (1994) kunmetis 10-milionan korpuson el rete troveblaj materialoj (ĉefe literaturaĵoj kaj gazetartikoloj), kiu servas la ellaboron de Esperanto-Japana vortaro.

La **Tekstaro de Esperanto** (<http://tekstaro.com/tekstaro.html>) – kvankam kun siaj 4,7 milionoj da tekstvortoj multe postrestas la naciajn korpusojn, jam donas uzeblajn rezultojn koncerne la fruan lingvuzon (proportie multaj Zamenhofaj tekstoj estas en ĝi) kaj nuntempan gazetaran lingvaĵon (pro Monato kaj La Ondo). Ĝi estas kreita surbaze de moderna tekniko kaj ebligas diversajn retrovojn. Ĝi estas rete alirebla por ĉiuj. Ĝia pliampleksigo laŭ tekstotipoj kaj vortonombro estas dezirinda ne nur por leksikografia laboro, sed por verkado de lernolibroj kaj gramatikoj.

Korpuso disponigas informojn pri la:

- strukturo kaj uzo de vortprovizo (ofteco de vortoj, iliaj diversaj formoj),
- pri la kunteksta signifo de vortoj,
- pri la gramatika konduto de vortoj (gramatikaj valentoj),
- komuna apero de vortoj (semantikaj emoj) – kolokaĵoj, frazeologiaĵoj.
- realaj uzkontekstoj (depende de tekstotipo).

## Temvortaroj kaj korpuso

Por temvortaro korpuso povas doni novajn ideojn pri vortkombinoj, se sufiĉe grandas. Por specialkorpuso pri instruado nur paŝetoj estas faritaj (La Torre 1997). La jam menciita temvortareto pri lernado el la serio Celtrafo (Koutny 2002) havas specialan strukturon, kie en la vortartikolo aperas la tipoj de la kapvorto, verbaj kaj adjektivaj esprimoj kaj aliaj rilataj vortoj, esprimoj.

Korpuso povas helpi (ekzemploj el Instruado) en:

- elekto de kapvorto (ekz. inter *instruado* – *klerigado* – *edukado*),
- elekto de konkurencaj formoj kiel *kvazaŭlernejo* – *kvazaŭa lernejo* – *virtuala lernejo*,
- precizigo de signifo, ekz. ĉe *bazlernejo*, *mezlernejo*, *altlernejo*, *kolegio*,
- kiuj lernejotipoj aperas en la uzo: *baza*, *elementa*, *kvazaŭa*, *meza*, *porinstruista*, *privata*, *somera*, *supera*, *virtuala*...
- Instalado de gramatika konduto: *partopreni ion* aŭ *en io*, *helpi iun* aŭ *al iu*, *leciono de* aŭ *pri historio* (Pluaj ekzemploj en Koutny 2003).

## Perspektivoj

La elektronikaj vortaroj integras pli kaj pli da informoj por la favoro de uzantoj. La semantikaj rilatoj, kiuj perdiĝas grandparte en tradiciaj unu- aŭ dulingvaj alfabetaj vortaroj trovas lokon en tezaŭroj (kiel la abonebla reta *Visualthesaurus*, kiu montras la rilatojn per arbodesegnaĵo), en la kompleksa rilataro de *Wordnet* (difino, hiponimoj, hiperonimoj, meronimoj, sinonimoj, antonimoj) aŭ *Framenet* (faktoroj de situacio).

Minnaja komencis ellabori Wordnet-version por matematikaj terminoj (<http://www.math.unipd.it/~minnaja/WORDNETESP/vortaro.html>). Miaj etvortaroj – se pli ampleksiĝos – eniros elektronikan datumaron, kie la rolantoj de situacio kaj iliaj karakteriziloj estos ligitaj kaj prezentablaj. Tio multe helpas la lingvolernanton.

Novgeneracia plena vortaro devos baziĝi sur grandega elektronika tekstaro, kiu siavice devas realiĝi. Tiel ni havos pli realan bildon pri la uzateco de vortoj.

## Bibliografio

- Blanke, Detlev (1996): Tipologio de nia leksikografio. En: *Literatura Foiro*, n-ro 161.
- Gotoo, Hitosi (1994): Esperantologii aziece. Raporto pri la Esperantologia konferenco dum UK en Seulo. In: Esperanto 1994/Oktobro kaj en la aktoj de "La 17-a Esperantologia Konferenco". Seoul: UEA
- Koutny Ilona (red. 1985): *Perkomputila Tekstoprilaboro*. Budapest: MESZ – SEC.
- (1998): Esperanta leksikografio kaj la Hungara-Esperanta vortaro. En: Gecső T. – Varga-Haszonits Zs. (red.): *Memorlibro*. Budapest: ELTE
- (2001): Defioj de moderna leksikografio por Esperanto. In: Fiedler, Sabine – Liu Haitao (red.): *Studoj pri interlingvistiko. Festlibro omaĝe al la 60-jariĝo de Detlev Blanke*. Studien zur Interlinguistik. Festschrift für Detlev Blanke zum 60. Geburtstag. Dobrichovice ĉe Prago: KAVA-PECH
- (2002): *Angla-Esperanta-Hungara etvortaro pri Lernado kaj Laboro*. Poznan: ProDruk
- (2003): Lexikography und die Bedeutung eines Esperanto-korpus. In: Detlev Blanke (red.): *Plansprachen und elektronische Medien*. Interlinguistische Informationen. Beiheft 9 – Beiträge der 12. Jahrestagung der Gesellschaft für Interlinguistik, 2002, Berlin. Berlin: GIL. 77-97
- (2005): Computational Tools for Elaboration and Functioning of a Multilingual Thematic Dictionary. En: Vetulani, Z. (red.): *Human Language Technologies as a Challenge for Computer Science and Linguistics*. Proc. of the 2<sup>nd</sup> Language & Technology Conference. Poznan: Wydawnictwo Poznanskie Sp. z.o.o. 41-44.
- (2008): Instruado de vortprovizo. En: *IPR* 2008/1
- La Torre, Mauro (1997): Desuprismo kaj desubismo en la terminara laboro. IKU-Referat während der UK in Adelaide. Manuskript
- Szerdahelyi István (1983-84): La leksikografiaj problemoj de la terminologio. En: *Planlingvistiko* n-ro 6 1983 somero, n-ro 9 1984 printempo, n-ro 12 1984 vintro