

Esperanto en komputa lingvistiko – retrorigardo

Ilona Koutny

Interlingvistikaj Studoj, UAM

Poznan, Pollando

Resumo

Trakti lingvon per komputilo aperis jam komence de komputiko (en la 50-aj jaroj) inter la celoj de la komputilaplikoj. Ankaŭ esperantistoj proponis jam frue (en la 70-aj jaroj) esperanton kiel modelon por la komputila lingvoprilaboro. La aŭtorino, kiu mem partoprenis en kelkaj projektoj kaj redaktis libron pri la temo antaŭ 25 jaroj (*Perkomputila Tekstoprilaboro* 1985), trarigardas la komencojn de esperanta komputa lingvistiko, prezentas ĝiajn diversajn terenojn i.a.: parolsintezo kaj -rekono, sintaksaj analizo kaj genero, maŝina tradukado, komputa leksikografio, komputile asistata lingvolernado laŭ la skemo: taskoj, problemoj, specifeco de esperanto kaj projektoj. La evolutendencoj de la tereno estos skizataj.

- 1. Enkonduke**
- 2. Komputa Lingvistiko (KL)**
- 3. Terenoj de komputa lingvistiko**
- 4. Esperanto (kiel modelo) en komputa lingvistiko**
- 5. De ortografia kontrolo al gramatika kontrolo**
- 6. E-o en la parolsintezo**
- 7. E-o en la parolrekono**
- 8. E-o kaj la Komputile Asistata LingvoLernado (KALL)**
- 9. Maŝina Tradukado (MT)**
- 10. Komputa leksikografio**
- 11. Resume**

1. Enkonduke

La unuaj komputiloj en la 50-aj jaroj estis grandaj kaj malfacile manipuleblaj, tamen la homa menso jam planis la ekscitan taskon prilabori homan lingvon, eĉ komputile traduki tekstojn. Centra persono inter la fruaj evoluigistoj de maŝina tradukado estis Yehoshua Bar-Hillel, kiu poste konvinkigis, ke strukturalisma lingvoprilaboro sen semantiko ne povos komprenigi tekston al komputilo (Nirenburg 1997). La pesismisma ALPAC raporto en 1966 bremsis la entusiasman esploron pri aŭtomata tradukado, sed la naturlingva prilaborado (NLP) surbaze de datumbazoj daŭris en diversaj terenoj. La evoluo de komputotekniko ebligis traktadon de multaj datumoj, de grandaj datumbazoj en la 70-aj

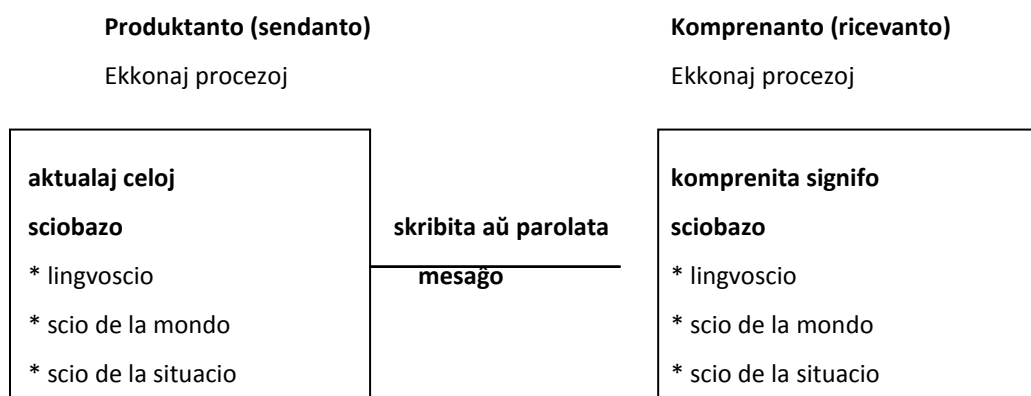
jaroj. Tiuj ebloj pli kaj pli kreskis kaj poste kondukis al la nuntempe ofte aplikataj statistikaj prilaboroj en NLP.

Por esperanto la „heroa epoko“ estis la 70-80-aj jaroj. Esperantistaj fakuloj pledis por la rolo de e-o kiel modelo por NLP, ĉefe en maŝina tradukado (en 1978 okazis diskuto, iniciatita far Ian Kelly pri tiu rolo de e-o en la konata semajna revuo *Computer Weekly*). Aperis e-rilataj aplikoj de pluraj fakuloj kiel la parolsintezo de Sherwood, la maŝintraduka sistemo de Maas kaj la vaste konata DLT instigita de Witkam, kiu uzis esperanton kiel pontolingvon de plurlingva reta tradukprojekto. Mi mem kunmetis trarigardan libron pri la temo: *Perkomputila tekstoprilaboro* (Koutny 1985) kaj eklaboris pri parolsintezo.

En la sekvoj mi skizos la terenon de komputa lingvistiko kaj la kontribuon de e-o.

2. Komputa Lingvistiko (KL)

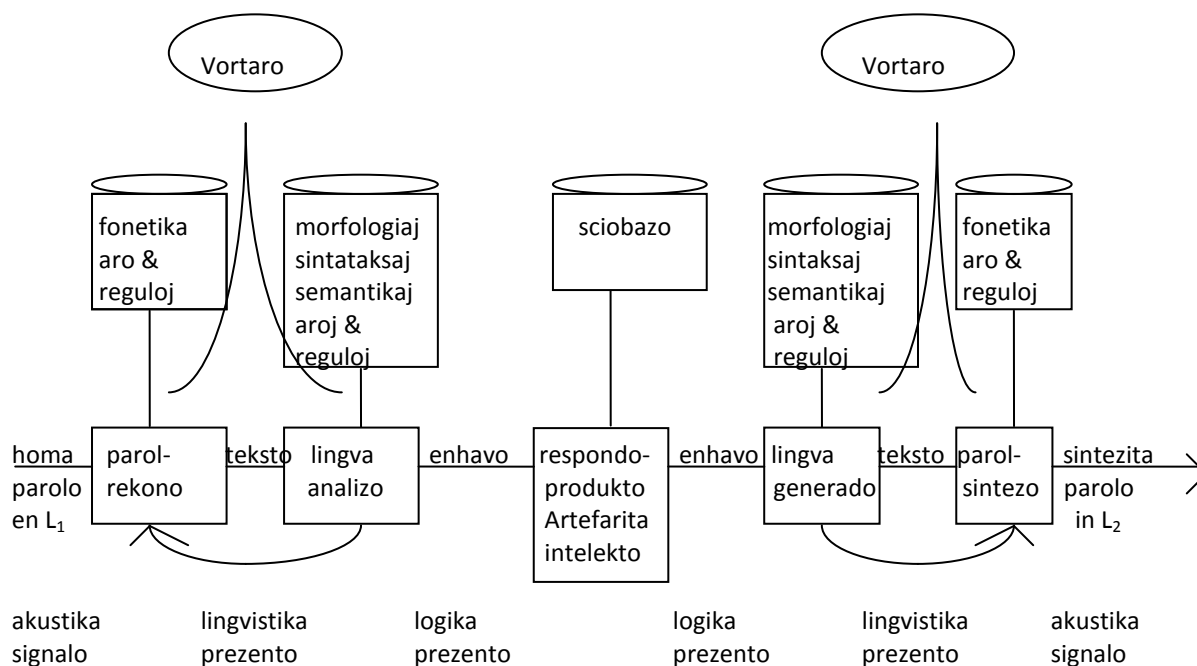
La apero de komputiloj en la komunika procezo kiel ebla partnero postulis pliajn precizigojn kaj modifojn en la komunika modelo. Ekestis nova **komputa paradigmo** (Winograd 1982). Winograd difinas la komunikadon kiel informinterŝanĝon inter inteligentaj proceziloj (la produktanto aŭ la komprenanto povas esti ankaŭ komputiloj), kiuj plenumas ekkonajn (kognitivajn) procezojn surbaze de sciobazo, kiel la Figuro 1 montras tion.



Figuro 1: Komunikado laŭ la komputa paradigmo de Winograd

La komunikadan procezon influas la aktualaj celoj de la komunikanto. La sciobazo entenas scion pri la mondo, en kiu la komunikantoj vivas, kaj scion pri la aktuala situacio, en kiu ili partoprenas. Por la lingva komunikado necesas scio de la sama lingvo, ĉu la mesaĝo estas parolata aŭ skribata.

La ĉefa celo de **komputa lingvistiko** estas la komputila traktado de homa lingvo por realigi la komunikadon inter homo kaj komputilo. Tio envolas diversnivelan traktadon de naturaj lingvoj kiel morfologiajn, sintaksajn, semantikajn analizojn kaj generojn. La komplika tasko koncerne scion pri la mondo kaj situacio jam transiras al la tereno de artefarita intelekto. Por dialoga sistemo tio ne eviteblas. Se konsideri la parolan komunikadon, kiu estas baza en la homa komunikado, necesas solvi la transformon de homa parolo (akustika signalo) al skribita teksto (lingvistika prezentaĵo) kaj en la alia fino de la procezo la inverson, t.e. la transformon de la skribita teksto al parolo. La Figuro 2 resume montras la taskojn de komunikado inter homo kaj maŝino.



Dialoga sistemo: $L_1 = L_2$, enira kaj elira lingvoj estas la samaj;

Maŝina tradukado: $L_1 \neq L_2$

Figuro 2: Modelo de komunikado homo-maŝino (Koutny 1988d)

3. Terenoj de komputa lingvistiko

Striktasence komputa lingvistiko signifas la **modeligon de la lingva komunikado per komputilo**.

Vastasence KL involvas ĉiuspecan **traktadon de lingvo per komputilo** kaj tiuj abundas:

- statistikaj prilaboroj (frekvenco de literoj, vortoj ktp.)
- tekstoredaktado
- ortografia kontrolo (literumkontrolo kaj propono por ŝanĝo)
- gramatika kontrolo (ekz. agordo inter adjektivo kaj substantivo)
- analizo de naturlingva teksto en morfologia, sintaksa, semantika niveloj
- genero de naturlingva teksto en diversaj niveloj
- dokumentoprilaboro (indeksado, resumado ktp.)
- maŝina tradukado
- parolsintezo
- parolrekono
- dialogaj sistemoj
- komputa leksikografio
- KALL: Komputile asistata lingvolernado

La terenoj interplektiĝas ofte, ekz. maŝina tradukado aŭ parolrekono necesigas diversnivelaĵajn lingvajn analizojn. Eĉ la ortografia kontrolo ne povas eviti morfologian analizon kaze de aglutinaj lingvoj kiel la hungara kaj kaze de fleksiaj lingvoj kiel la pola.

4. Esperanto (kiel modelo) en komputa lingvistiko

Esperanto estas aplikata ne nur paroli pri KL, sed ĝi – kiel lingvo – estas studtereno de KL, kaj pro ĝia reguleco proponiĝis kiel modelo, resumante ĝian rolon:

- **proprarola analizo** kiel lingvo (ekz. sintezo de esperanta parolo)
- **modelo**
 - pli simpla kaj regula strukturo ol tiuj de naturaj lingvoj
 - malpli da ambigueco, kvankam ekzistas:

Li legis la libron la tutan tagon veturante Parizon.

malambiguigo de tiaj sintaksaj strukturoj okazis en DLT

tablo, tabelo, tabulo (restrikto de signifo „table“ jam origine far Zamenhof)

tamen plursignifeco neeviteblas: *piedo, muso*.

- **pontolingvo en MT** ĉe DLT.

Inter la unuaj la sveda Siv Sjøgren (1970) proponis esperanton por perkomputila sintaksa analizo en dialoga sistemo kaj enkondukis kelkajn modifojn. La hungara komputisto Antal Münnich ellaboris la programlingvon PROGRESO destinita por tekstaj prilaboroj en la 70-aj jaroj (ekz. Münnich – Arnold – Hilgers 1975). La ordonojn eblis formi per simplaj esperantaj frazoj.

La celo de PREDIS (Frank et al. 1979) estis retrovi tekstajn dokumentojn ankaŭ per e-o surbaze de ŝlosilvortoj. La datumbazo entenis tekstojn pri klerigkibernetiko. En tiuj fruaj jaroj okazis la enkomputiligo de la kapvortoj de PIV, inversigo kaj selektado laŭ etiketoj (Gy. Mandler).

La sekvaj elektitaj terenoj estos prezentataj laŭ la skemo: Tasko, Problemoj, Specifeco de e-o, Projektoj kaj Tendencoj.

5. De ortografia kontrolo al gramatika kontrolo

Tasko:

Teksto → literumkontrolo → propono por ŝanĝoj → ortografie ĝusta teksto

Problemoj:

Diversaj interpretebloj de vortostrukturo, ekz. vortoj kun afiksoj aŭ radikoj; vortoj havas diversajn formojn, necesas trovi la radikojn kaj poste proponi solvon kun la ĝusta finaĵo.

Specifeco de e-o:

Diversaj interpretoj ekzistas kiel *kolego*: kol-ego aŭ kolego, *sentema*: sen-tema aŭ sent-ema, eĉ se ne ĉiam sencohava *turbino*: turb-ino.

La eblaj finaĵoj estas limigitaj kaj regulaj (-n, -j, -jn), tiel la retrovo de radiko estas simpla. La problemo de ĉapelitaj literoj komence kaŭzis problemojn kaj uzatis diversaj kodadoj, definitiva solvo ankaŭ nun ne ekzistas, sed proksimas.

Projektoj:

En la fruaj tempoj Gábor Deák Jahn (HU) laboris pri literumkontrolo, sed la rezulto ne estis enkonstruita en tekstoredaktilon kaj disvastigita. **Ĉapelilo** estis tekstoredakta programo kun literumilo kaj dividilo sufiĉe vaste uzata (ellaborita far Simono Pejno). Ĝi estis aktualigata ĝis Windows 2000.

Tendenco: Kunligo kun gramatika kontrolilo → Esperantilo

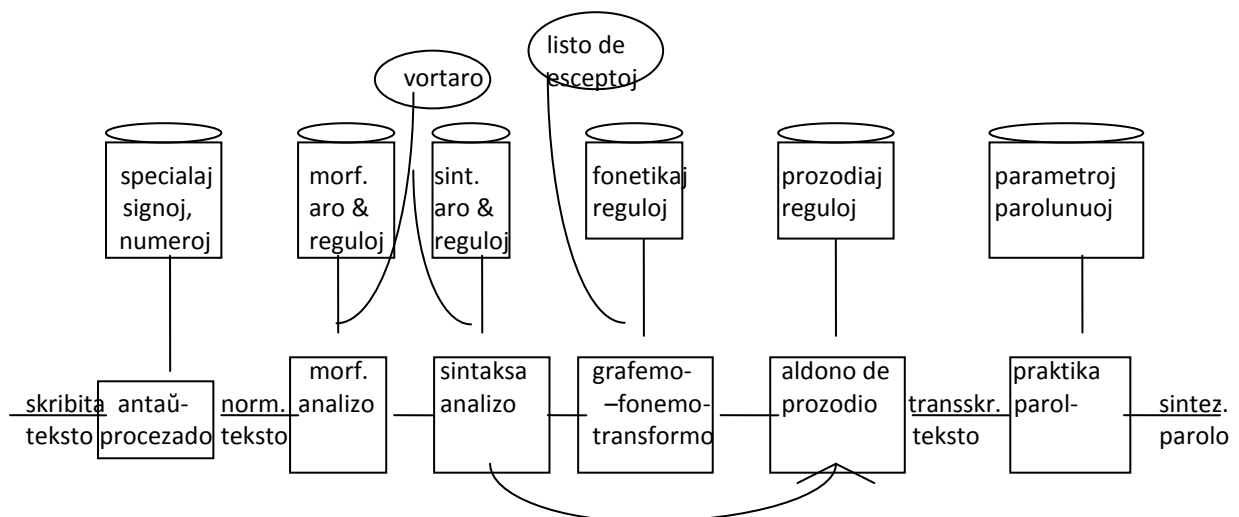
Necesas certa tekstoanalizo (kontrolo de agordoj, ev. verbaj rekcioj ktp.) kaj interfaco amika al la uzantoj (aplikeblo en diversaj operaciaj sistemo kiel Vindozo kaj Unikso, uzo de e-literoj).

6. Esperanto en la parolsintezo

Tasko:

Teksto skribita → transformo al elparolaĵo + elparoligo → parolo

La sintezo komenciĝas per la analizo de la skribita teksto, kiu povas enteni ne nur vortojn. Post ilia transformo al elparoleblaj vortoj – depende de la donita lingvo – pli-malpli kompleksa morfologia analizo sekvas (ĉe morfemlimoj necesas atenti). La sintaksa analizo gravas por la la konvena intonacio (en la angla ankaŭ por la ĝusta prononco de vortoj). Komence la sintezo realiĝis per reguloj kaj hardware, poste pli kaj pli samplaj el homa parolo estis uzataj kaj la tuto estis realigata kutime software. La sekva figuro montras la funkciadon de klasika parolsintezo.



Figuro 3: Procezo de parolsintezo (Koutny 1988d, 2001)

Problemoj:

La neregula ortografio kaŭzas problemojn (ekz. en la angla multaj vortoj kun siaj prononcoj devas esti storataj), diversaj elparolebloj depende de la ĉirkaŭaĵo necesigas analizojn (same en la angla *record* havas du prononcojn depende, ĉu ĝi estas verbo aŭ substantivo). La konverto de teksto al elparoleblaĵo dependas de la reguleco de ortografio.

Sen ĝusta prozodio la elparolaĵo ne estas natura kaj kelkfoje eĉ ne bone komprenebla. Sintaksa analizo povas helpi por trovi la frazinternajn limojn, kie paŭzo povas aperi kaj la krei la konvenan intonacion.

Specifeco de e-o:

La transformo de teksto al elparolaĵo estas sufiĉe regula pro la sistema esperanta ortografio, t.e. fareblas per reguloj, kaj pro tio ĝi estas pli facila ol la sintezo de multaj etnaj lingvoj. Tamen la elparola normo ne estas tiom konkreta kiel ĉe etnaj lingvoj, kaj la prozodio tute ne estas difinita. Tiel eblas pli libere difini la ecojn de la sintezita parolo. Certajn aferojn kiel la akcenton, necesis reрпиensi pro la sintezo (detala priskribo en Koutny 1994, 2001).

Projektoj:

Bruce Sherwood (1982a,b, 1985) surbaze de la angla VOTRAX kaj analizo de parolata e-o konstruis esperatan parolsintezon kaj aplikis la sintezitan parolon en la instruado.

Esparol estis ellaborita en Hungario en la Budapeŝta Teknika Unversitato far Ilona Koutny kaj Gábor Olasz (Koutny et al., 1987, Koutny 1988d, 2001) surbaze de la hungara parolsintezilo Multivox (kiu uzis sonsegmentojn, tn. *dufonojn*). Aparta sonaro estis difinita surbaze de la hungaraj sonoj, reguloj por akcentado estis formulitaj kaj intonacio aldoniĝis. Interesaĵo, ke la regulbaza tn. *formanta sintezo* estis hardware realigita (komence por Commodore (!), poste por personaj komputiloj), la apartan sintezilon necesis kunligi kun la komputilo.

Tendencoj:

Programara sintezo sekvis surbaze de parolsegmentoj ekstraktitaj el parolataj tekstoj. Ne estas konata por mi sintezo de e-o, kiu realigus la ekstraktadon de parolsegmentoj el registrita esperanta parolo, nur provoj apliki ekzistantan sintezilon de alia lingvo por esperanto, ekz. „Parolu“ (<http://traduku.net/parolu/#parole/>) surbaze de la franca MBROLA.

7. E-o en la parolrekono

Tasko: Rekono de vortoj kaj rekono de flua parolo

Parolo → fizika rekono + lingva identigo → skribita teksto

Problemoj:

Malsamaj fizikaj signaloj havas la saman lingvan enhavon. Ni ĉiuj parolas diversmaniere, ja la parolo portas signojn ne nur de la enhavo, sed ankaŭ de la parolanto. Tio, kio helpas al la homo rekonante la voĉon de lia/ŝia interparolanto, kaŭzas problemojn al la komputilo. Krome la sama homo parolas diverse, depende de lia/ŝia animstato, eventuala malsano, ktp.

Specifeco de e-o:

Esperanto ne havas grandajn avantaĝojn ĉe parolrekono kompare kun tiu de etnaj lingvoj, pro la kompleksa konsonanta sistemo, kiu kaŭzas miksadon de similaj vortoj, ekz. *nano – mano – mamo, ses – sep* aŭ *dono – tono*. Ankaŭ la similaj finaĵoj kontribuas al la falsa rekono de vortoj. Sed lingva modelo, kiu estus pli regula kaj travidebla por esperanto, povus helpi (vd Koutny 1991, 2001).

Projektoj:

Ekzistis nur provoj por esperanto en rekonilo de izolitaj vortoj, sed ne ellaborita parolrekonsistemo (Koutny 1985).

Tendencoj:

Por rekono de flua parolo necesas registri sufiĉe longan trejnparolon. Kaze de fremdaj parolantoj la sistemoj pli malbone plenumas la rekonon. Ofte temas pri lernantaj sistemoj, kiuj uzas neŭronajn retojn (por simuli la homan heuristikon). Esperantistoj ne engaĝiĝis en tiu ĉi tereno.

8. Esperanto kaj la Komputile Asistata Lingvolernado (KALL)

Tasko: helpi al la lernanto en la lingvolernado per la ebloj de komputilo

Ekzistas simplaj taskoj, kiuj plejparte imitas la tipajn lernolibrajn taskojn: kompletigo de teksto, anstataŭigo de certaj elementoj, gramatikaj transformoj, ortografia kontrolo, simplaj demandoj kun difinitaj respondoj. La komputila sistemo per sia plurmedia fono kiel registrita voĉo, bildoj, lastatempe ankaŭ filmetoj, igas la taskojn pli allogaj. La kontrolkapablo kaj la sumigo de rezultoj en komputila lernprogramo certe superas la eblojn de tradiciaj libroj.

Problemoj:

La ĉefa tasko de lingvolernado estas kapabligi la lernantojn komunikadi, stimuli ilin al interparolo. Por tio necesas plena lingva analizo kaj ankaŭ kompreno de la situacio! Tiun taskon komputiloj ankoraŭ ne kapablas plene realigi, sed la inteligentaj programoj evoluas en tiu direkto. Por la parolkomunikado ankaŭ parolsinteza kaj -rekona modulo devus esti integritaj al la instrusistemo. La rekono de ĝusta flua parolo estas malfacila, la diversmaniere erara eĉ pli. Ankaŭ kontrolo de tradukado kaŭzas problemojn al la komputilo, ja tie eblas pluraj bonaj solvoj. Pro tio KALL transprenas unuavice la taskon de mekanika ekzercado de gramatiko kaj vortoprovizo.

Specifeco de e-o:

Pro la travidebla gramatika strukturo de e-o pli da taskoj estas kontroleblaj komputile, sed la inteligentaj taskoj ankaŭ en e-o kaŭzas malfacilaĵojn, des pli se la komunikado okazas parole.

Projektoj:

Parolern estis aŭtora instruprogramaro uzante la parolsintezan sistemon ESPAROL (Koutny 1991), kiun karakterizis:

- Simplaj taskoj kun kelkaj antaŭvideblaj solvoj (trutekstoj, transformoj, plurelektaj taskoj, ktp.),
- Prezento kaj ekzerco de vortoj,
- Eluzo de sintezita voĉo (Esparol), ekz. por diktado, demandado, anonco de rezultoj,
- Provo de primitiva interparolo (simile al la kapabloj de Eliza),
- Sistema de ekzercoskemoj, kiun instruistoj povis plenigi per siaj materialoj (temis pri tn. aŭtoraj programoj),
- Komenco de kursmaterialo por baza esperanta kurso.

Tendencoj:

Teknike pli kaj pli altnivelaj kursoj aperas sur KD kaj en la reto kun bildoj, voĉo (malofte kun rekono), filmetoj. En la reto troviĝas pluraj e-kursoj, inter kiuj Lernu.net estas tre populara. Sed ofte la tradiciaj taskoj aperas en plurmedia prezento.

9. Maŝina tradukado (MT)

Tasko: Teksto en fontolingvo → traduko → teksto en cellingvo

Analizoj morfologia, sintaksa, semantika
 Generoj morfologia, sintaksa, semantika

Tipoj:

- rekta (tipo Systran): traduko preskaŭ laŭvorte;
- transfera (kiel EUROTRA): la fontolingvo estas analizata ĝis certa nivelo (profundeco varias de sistemo kaj lingvoparoj) kaj poste okazas transiro al la cellingvo;
- interlingva (abstrakta prezentaĵo aŭ uzo de funkcia interlingvo kiel e-o en DLT).

Problemoj:

Lingvoj diferencas inter si en strukturo, vortoprovizo kaj laŭ la lingva bildo de la mondo. Ambigueco aperas en diversaj lingvaj niveloj, kio malfaciligas la komplete aŭtomatan tradukadon. Unuecigon helpas la kunteksto kaj la scio pri la mondo se temas pri homoj. La tn. *sciobankoj* estas tre limigitaj, ĉar nur parton de la scio eblas stori.

Specifeco de e-o:

Esperanto estas malpli ambigua (plursenca) ol la etnaj lingvoj pro sia planiteco, tamen ĝi ne povas eviti la ambiguecon pro sia normala funkciado en homa komunikado. La sintaksa plursignifeco (kiel la esprimo de tempo kaj loko same per akuzativo) solveblas, sed la ambigueco de diverssignifaj vortoj kaŭzas problemojn.

Esperanto povas plenumi tre konvene la rolon de interlingvo (IL) en MT. Pro sia sufiĉe granda reguleco, manko de esceptoj ĝi taŭgas por la komputila traktado, kaj samtempe povas koncize transigi informojn simile al etnaj lingvoj (kp. Schubert 1992). Abstrakta IL, kiu entenus ĉiujn ecojn de la fontoteksto, estus tre komplika.

IL-traduko estas kost-efika, ĉe aliaj tipoj necesas pli multe da tradukiloj (tradukdirektoj). Kaze de n lingvoj necesas nur $2n$ tradukiloj en IL-sistemo kaj kaze de traduko el ĉiuj eblaj lingvoj de la sistemo al la aliaj: $n(n-1)$. Se konsideri nur la 6 oficialajn lingvojn de la Unuiĝintaj Nacioj, la diferenco tiel aspektas: $2 \times 6 = 12$ kompare al $6 \times 5 = 30$.

Projektoj:

Heinz Dieter Maas (1985) laboris pri la maŝintraduka sistemo **SUSY**, kiu celis traduki inter e-o kaj la germana.

Distribuita Lingvo-Tradukado (T. Witkam, K. Schubert, V. Sadler – 80-aj jaroj)

La plej prestiĝa projekto estis DLT, kiun iniciatis Toon Witkam en la nederlanda firmao BSO. En tempo, kiam ni ankoraŭ eĉ ne konjektis, ke la interreto penetros nian tutan vivon, li planis sistemon, kiu estas tempe kaj loke distribuita pro la bezonoj de retuzantoj internaciskale. La sistemo estis destinita plenumi tradukadon el la fontolingvo tie, kie la fontoteksto estis enigata kaj en la momento de enigo, al la interlingvo esperanto. La uzanto povis peti tradukon el IL tiam, kiam li volis legi la informojn kaj tie, kie li estis. La problemojn leviĝantajn ĉe la traduko de FL al IL, eblis solvi dum la tn. *malambiguiga dialogo* inter la komputilo kaj la eniganto.

La interne uzata e-o estis sintakse reguligita por eviti problemojn ĉe la traduko el IL al la cellingvo. La prototipo inter la angla kaj franca (1987-88) estis prezentata dum internaciaj konferencoj pri KL (i.a. dum la speciala konferenco pri Maŝina Tradukado en Budapeŝto en 1988). Ankaŭ aliaj lingvoj estis pretigataj por la alligo al la sistemo, ekz. la hungara, kiu havas alian strukturon ol la indoeŭropaj lingvoj. Tiuj lingvoj estis priskribataj laŭ la dependogramatiko (en Maxwell & Schubert 1989).

Komence temis pri regulbaza traduko, sed pli poste ankaŭ paralelaj tekstoj en la donitaj lingvoj estis uzataj, kaj okazis statistikaj decidoj surbaze de la komunaj aperoj de diversaj lingvaj elementoj (ekz. Sadler & Vendelmans 1990). Priskribo de DLT: en Witkam 1985, Schubert 1989, 1996, Maxwell 1992 ktp.

Ekzistis kelkaj aliaj provoj de maŝina tradukado en la 80-aj jaroj en diversaj landoj de Nederlando (de Kat) ĝis Ĉinio kaj Japanio (Katsumori, Saato, vd. Koutny 1985).

Tendenco:

Perstatistika aŭtomata tradukado (Witkam 2005).

10. Komputa leksikografio

Taskoj: tekstoredaktado, ordigo laŭ alfabeto, unuecigo, selektado laŭ etiketoj

Tekstaroj servas kiel bazo por decidi pri kapvortoj (oftaj vortoj), gramatikaj ecoj (rekcioj), kutimaj esprimoj, frazeologio (komuna apero).

Problemoj:

Vere fidindajn datumojn eblas akiri nur el grandegaj tekstaroj, kiuj abunde reprezentas la donitan lingvon, ties kreo kostas multe. Necesas ankaŭ aŭtomata ekstrakto de bazaj vortoj el diversformaj tekstaj vortoj.

Specifeco de e-o:

Por e-o, kiu ne ne havas signifan kaj unuecan nedenskan parolantaron, nur parolantojn kun diversaj lingvaj kaj kulturaj fonoj, estas esenca havi grandan tekstaron kaj surbaze de tio pretigitajn vortarojn por superi la lokajn variantojn, diferencojn en la uzo de vortoprovizo. La plej signifa vortaro PIV ankoraŭ ne eluzas tiun eblecon.

Projektoj:

Esperanta Tekstaro (ellaborita far Bertilo Wennergren) entenas pli ol 4 milionojn da vortoj (i.a. zamenhofaĵojn kaj revutekstojn).

Retaj vortaroj ebligas aktualigon konstante, ofte per komuna evoluigo kaj postulas novtipan leksikografian laboron. **ReVo** estas disvolviĝanta projekto kun difinoj en e-o kaj ekvivalentoj en pluraj lingvoj, ankaŭ temindekso.

Tezaŭro estas vortaro surbaze de sistemigita sciaro. La projekto **Strukturita Esperanta Temvortaro** havas similan strukturon, ĝi prezentas la lingvan bildon de la mondo en laŭtemaj grupoj en pluraj lingvoj uzante e-o kiel mapilon (Koutny 2005).

work, labour, US labor (konkreta) job	laboro	munka
• mental ~	• intelekta/mensa/ spirita ~	• szellemi ~
creative ~	krea ~	alkotó~
physical ~	korpa/fizika ~	fizikai ~
➤ <i>i</i> work, do the ~	➤ labori, plenumi ~n	➤ ~t végez, dolgozik
look/hunt for a job	serĉi ~n	~t keres
job seeker/hunter	laborserĉanto	munkakereső
classified ads	anoncetoj	apróhirdetések
o permanent, constant	o konstanta	o állandó
seasonal	sezona	szezonális
varied	varia	változatos
monotonous	monotona	egyhangú, monoton
♣ working	♣ laborado	♣ munkavégzés

work discipline **labordisciplino** munkafegyelem

working method **labormetodo** munkamódszer

Reduktita ekzemplo de la Angla-Esperanta-Hungara Temvortaro

(Koutny 2002)

11. Resume

Komputa lingvistiko – surbaze de analizo de homa lingvo – provas provizi la komputilojn per similaj komunikkapabloj kiujn la homoj manifestas interparolante. Tiu tasko estas sufiĉe ambicia, sed ĝuste pro tio defias kaj defias la esploristojn de KL. Eblas instrui al la maŝino bone difinitajn taskerojn, pro tio necesas algoritmigi la pli kompleksajn taskojn al iuj difinitaj unuoj. Problemoj aperas tiam, kiam la tasko eskapas de la algoritmigeblo, ja la homa intuicio ne programeblas, kvankam la neŭronaj retoj provas imiti la homan funkciadon. Anstataŭ la plena lingvistika analizo, regulbazaj solvoj, pli kaj pli kreskas la rolo de grandegaj tekstaroj kaj statistikaj konsideroj anstataŭas la lingvistikajn regulojn. La interreto disponigas novajn eblojn kaj donas alireblon al grandaj amasoj. La taskoj de KL estas tre aktualaj por ke ĉiuj uzantoj povu havi aliron al informoj en sia gepatra lingvo aŭ en lingvo, kiun li/ŝi konas.

E-o ekde la komenco de KL havis rolon en ĝi. Kiel pli regula, senescepta lingvo ol la etnaj, e-o havas pli da ŝancoj por la algoritmigo, ellaboreblas pli travidebla lingva modelo. Tiu rolo destinis ĝin por la rolo de interlingvo en MT, kaj tio realiĝis en DLT. Ankaŭ en aliaj terenoj de KL e-o estis traktata simile al etnaj lingvoj, ekz. en parolsintezo (ESPAROL). En la retaj aplikoj, kiel retaj vortaroj kaj instrusistemoj e-o jam havas bone establitan lokon (ReVo, Lernu.net). Ĉi tie mi provis prezenti la komencojn kaj la tendencojn de apliko de e-o en KL kun elektitaj ekzemploj, pro la limigoj nur tre surface. La tuta tereno bezonus multe pli grandan studon.

Bibliografio

Esperanta tekstaro. <http://tekstaro.com/>

Gledhill, Christopher 1998: *The Grammar of Esperanto. A corpus-based description.* München: Lincom Europa. 151 p.

De Kat, J. O.: Traduko el Internacia Lingvo al Naciaj. En: Koutny 1985: 259-266

Katsumori, H. – Makino, S. – HIRATA, M. 1987: *A Machine Translation System by the Use of Esperanto as a Bridge Language.* Universitato de Cubu

Koutny I. 1984: Diversaj aspektoj de perkomputila tekstoprilaboro. En: *Internacia Komputado* 1984, n-ro 2, Budapeŝt: MESZ, 4-7.

Koutny I. 1985: Problemoj de esperanto en la aŭtomata tekstoprilaboro. En: *Planlingvistiko* 1985, n-ro 3, La-Chaux-de-Fonds: LF-Koop., 7-13.

Koutny I. (red.) 1985: *Perkomputila Tekstoprilaboro.* Budapeŝt: MESZ - SEC.

- Koutny I. 1988a: Automatic Synthesis of Esperanto, kun: G. Olaszy, L. Czap En: *Budapesti Akusztikai Kollokvium* 1988, nr 9, 132-137.
- Koutny I. 1988b: Esperanto Speech Synthesis and its Application in Language Learning, kun: G. Olaszy, K. Kisfaludy, En: *Magyar Fonetikai Füzetek*.1988, n-ro 19, Budapest: Nyelvtud. Int., 47-54.
- Koutny I. 1988c: La rolo de komputilo en la instruprocezo. En: *Internacia Pedagogia Revuo* 1988, n-ro 4, 1-6.
- Koutny I. 1988d: Komputila parolgenero. En: *Fokuso* 1988, n-ro 4, 48-54.
- Koutny I. 1991: Komputila parolrekono. En: *Scienca Revuo* 1991, n-ro 1, 6-13.
- Koutny I. 1991: PAROLERN: A Speaking Authoring Program Package, En: *EUROCALL 1991*, Int. Conference on Computer Aissisted Language Learning. Helsinki: Helsinki School of Economics. 62-69.
- Koutny I. 1994: Fonetika fono kaj realigo de esperantlingva elektronika parolgenero kun: G.Olaszy. En: H.M.Maitzen, H.Mayer, Z.Tisljar (red.) 1994: *Aktoj de Internacia Scienca Simpozio: Esperanto 100-jara. 1987*. Vieno: ELA, Pro Esperanto, Maribor: Inter Kulturo. 187-202.
- Koutny I. 1997: Maŝintraduka modelo. En: Koutny I (red.) 1997: *Struktura kaj Socilingvistika Esploro de Esperanto*. Budapest: Steleto & ILEI. 153-161.
- Koutny I. 2001: Defioj de moderna leksikografio por Esperanto. En: Fiedler, Sabine - Liu Haitao (red.) 2001: *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. 660-674.
- Koutny I. 2001: Speech Processing and Esperanto. En: *Interface. Journal of Applied Linguistics* 2000/1, 99-120. kaj En: Schubert, Klaus (red.) 2001: *Planned Languages: From Concept to Reality*. Brussel: Hogeschool voor Wetenschap en Kunst.
- Koutny I. 2002: Lexikography und die Bedeutung eines Esperanto-korpus. En: Detlev Blanke (red) 2003: *Plansprachen und elektronische Medien*. Interlinguistische Informationen. Beiheft 9 – Beiträge der 12. Jahrestagung der Gesellschaft für Interlinguistik, 2002, Berlin. Berlin: GIL. 77-97.
- Koutny I. 2005: Computational Tools for Elaboration and Functioning of a Multilingual Thematic Dictionary. En: Vetulani, Z. (red.) 2005: *Human Language Technologies as a Challenge for Computer Science and Linguistics*. Proc. of the 2nd Language & Technology Conference. Poznan: Wydawnictwo Poznanskie Sp. z.o.o. 41-44.
- Koutny I. 2005: A Corpora Assisted Multilingual Thematic Dictionary. En: *Archives of Control Sciences* 2005/3. 251-260.
- Koutny I. 2008: *Natural Language Processing for Hungarian Speech Synthesis*. Poznan: Wydawnictwo Naukowe UAM. 200 p.
- Koutny I. 2009: Esperanta leksikografio en la epoko de komputiloj. En: Koutny I. (red.) 2009: *Abunda fonto*. Memolibro omaĝe al prof. István Szerdahelyi. Poznań: ProDruk & Steleto. 402-407.
- Liu Haitao 2001: Informadika Aspekto de Interlingvistiko. En: Fiedler, Sabine & Liu Haitao (red.) 2001: *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.
- Maas, H.D. 1985: La traduksistemo "SUSY" kaj ĝiaj perspektivoj. En: Koutny 1985: 163-174
- Maas, H.D. 1985: Pri Kelkaj Strategioj por Fraz-Analizo. En: Koutny 1985: 175-206
- Maxwell, Dan 1992: *Perkomputila tradukado: la revo kaj la realo*. „Esperanto-Dokumentoj“ n-ro 30E. Rotterdam: UEA. 20 p.
- Maxwell, Dan & Schubert, Klaus & Witkam, Toon (red.) 1988: *New Directions in Machine Translation*. DLT4 Dordrecht: Foris
- Münnich Antal & Arnold, E. & Hilgers, R. 1975: *Prinzipen der Programmiersprache PROGRESO*. Paderborn: Institut für Kybernetische Pädagogik.
- Nirenburg, S. 1997: [Bar Hillel on Machine Translation: Then and Now](#). In: M. Caspi and E. Shamir (Eds.), In Memoriam Yehoshua Bar Hillel. <http://ilit.umbc.edu/SergeiPub/bar-hillel.pdf>

- Ouyang Wendao 1985: Enkonduko de 5-dimensia modelo de ĉin-lingva informo kaj algoritmaj studoj pri vort-identigo, fraz-analizo kaj semantika prezentado. En: Koutny 1985: 145-152.
- Sadler, Victor & Vendelmans, Ronald 1990: Pilot Implementation of a Bilingual Knowledge Bank. <http://acl.ldc.upenn.edu/C/C90/C90-3101.pdf>
- Sherwood, Bruce A. 1982a: Spertoj pri sintezo de Esperanta parolado. In: Frank, H. & Yashovardan (red.) 1982: *Lingvo-kibernetiko*. Tübingen: Günter Narr Verlag. 63-74.
- Sherwood, Bruce A. 1982b: Statistical Analysis of Conversational Esperanto, with Discussion of the Accusative. En: *Studies in the Linguistic Sciences*. Vol 12. Spring/1982. 165-182.; <http://www4.ncsu.edu/~basherwo/docs/EsperantoParolata.pdf>
- Sherwood, Bruce A. 1985: Raporto pri sintezo de Esperanta parolado. / Sintezo de Esperanto kaj de diversaj naturaj lingvoj En: Koutny 1985: 49-56.
- Schubert, Klaus 1989: A Dependency Syntax of Esperanto. En: Maxwell, D. & Schubert, K. (red.) 1989: *Metataxis in Practice. Dependency Syntax for Multilingual Machine Translation*. Dordrecht/Providence: Foris, 207-232.
- Schubert, Klaus 1992: Esperanto as an Intermediate Language for Machine Translation. En: Newton, J. (red.) 1992: *Computers in Translation*. London/New York: Routledge. 78-95.
- Schubert, Klaus 1996: DLT – faktoj pri finita projekto. <http://www.esperantic.org/ced/HLT.htm>
- Tišljar, Zlatko 1982: Pri la morfemfrekvencaro en la parolata Internacia Lingvo. En: *Grkg/Humankybernetik* 231982/1: 40-47.
- Winograd, Terry 1982: *Language As A Cognitive Process, Volume 1, Syntax* Addison-Wesley.
- Witkam, A.P.M. 1985: Distribuita Lingvo-Tradukado. En: Koutny 1985: 207-228.
- Witkam, A.P.M. 2005: Nova vojo al aŭtomata tradukado. En: IKU 2005. Amri Wandel (red.). UEA