

TRADUKESPLOROJ
per MALGRANDKAPACITA KOMPUTILO LERNEJA

Alberto García Fumero

1. Enkonduko.

Tiu ĉi artikolo celas montri kelkajn el la malfacilaĵoj kiujn devas fronti tiuj, kiuj deziras esplori la perkomputilajn tradukproblemojn pere de malgrandkapacita komputilo (kiel ekz. lerneja) kune kun ĝiaj eblaj solvoj. Dum la esplorado la aŭtoro uzis lernejan, 256-kbajtan komputilon, por provi tradukon de tekstoj redaktitaj en la hispana en la esperantan lingvon.

2. Elirpunkto.

La esperantan lingvon la aŭtoro elektis kiel cellingvon ĉar ĝi estas simpla. Cele al la esplorado de la tradukproblemoj el la hispana, la traduksistemo KOMENCANTO estis redaktita. Ĝis nun tiu ĉi poresplora sistemo konsistas el la jenaj programoj funkciaj sub la regado de la operacia sistemo MS-DOS:

- ĉefa tradukprogramo (LERNANTO.EXE),
- programo por eltrovi la eblajn komplementojn en frazo (MONITORO.EXE),
- vortaro faka (ekz. informadika vortaro) (VORTARO.INF),
- fakprogramo por krei vortarojn (DICCIONARIO.EXE).

Ĉiuj programoj estis redaktitaj pere de la programadlingvo TURBOPASCAL 5.0.

La tradukotaj tekstoj hispanaj povas esti redaktitaj pere de iu ajn tekstoprilaborilo kapabla eligi ASCII-tekstojn. La tekstoj estas poste legataj de la ĉefa programo, kiu transdonas krudan tradukon al MONITORO.EXE.

La ĉefa programo provas traduki laŭvorte, sen ŝanĝo de la hispana vortordo, laŭ la rekomendoj k konkludoj de antaŭa esploro (Fumero, 1992). Se vorto

restas neidentigita, la programo ĝin aperigas surekrane majuskligita. La atributoj de ĉiu identigita vorto estas konservataj dum la procezo en matrica variabla je signoĉeno, kiel proponis inter aliaj MAAS; tio iomete malrapidigas la analizon, sed ŝparas memor-spacon, kio estas la plej valora por sistemo funkcia en malgrandkapacita maŝino. La rezulton de la procezo la dua programo (MONITORO.EXE) skribas sur diskedo.

3. Malfacilaĵoj k problemoj.

Ĉar la maŝino estis malgrandkapacita estus nepenseble uzi diskedojn grandkapacitajn; ju pli simpla la sistemo des pli bona. Pro la samaj motivoj oni elpensis la vortaron radika, kio signifas, ke nur radikoj senfinaĵaj aperas en ĝi. Ĉiu artikolo el la vortaro kunportas la jenajn informojn:

- radiko hispana,
- plej ofta esperanto-ekvivalento,
- markilo indikanta, ĉu la vorto estas plursignifa,
- interpreto (1-loko, 2-planto, 3-landnomo, 4-objekto ktp),
- vortoklaso (substantivo, adjektivo ktp),
- ĉu verbo mova,
- ĉu verbo transitiva,
- longo de la hispana radiko,
- ebla dua signifo de la vorto,
- ebla tria signifo de la vorto.

Tiu ĉi estas ankoraŭ portempa aro da informoj, ĉar la aŭtoro ne estas plene konvinkita pri la taŭgeco inkluzivi kelkajn el ĉi tiuj informoj. Estas pripensinde inkluzivi la informon "ĉu vivanta" por solvi la problemon estigitan de frazoj hispanaj kiel ekz. *carnes comen perros*, kiu el la gramatika vidpunkto povas estis ĝuste interpretita kiel "viandon manĝas hundoj" k ankaŭ kiel "viando manĝas hundojn". En la hispana plej ofte troviĝas la objekto post la verbo transitiva, sed tio kelkfoje ne eblas, ĉefe pro stilefetoj. Jen tikla punkto.

La informo, "ĉu objekto" taugas ekz. por decidi, ĉu la regulo pri la genro devas esti aplikata (Fumero, 1992).

En la hispana oni kutime signas la inan genron per la morfemo -a: *perro* -

hundo, *perra* – hundino, kvankam kelkfoje tio ne validas, ĉar oni uzas alian morfemon: *emperador* – imperiestro, *emperatriz* – imperiestrino, au la du genroj postulas malsamajn radikojn. Se temas pri la sama radiko, la genron oni signas la antaŭlasta.

Kiam temas pri objekto, la regulo pri la genro ne validas: *mesa* – tablo ne estas ina formo de radiko; en tiaj kazoj la litero –a ne rolas genra, sed apartenas al la radiko. Kvankam temas pri ina vorto, kiu postulas inan formon de la difina artikolo, tio nenion signifas el la komputila vidpunkto: ne estas, ne povas esti vira formo de la radiko. Pro tio, por komputilaj celoj, en tiu kazo la litero –a ne rolas genre, kvankam tio povas soni hereze.

La informo "interpreto" utilas ankaŭ por decidi, kiu estas la ĝusta traduko, kiam temas pri vorto finiĝanta per sufikso. La sufiksoj hispanaj ne ĉiam portas la saman valoron. Ekz. en kazo de substantivoj la sufikso –al indikas, ke temas pri loko; ĉe adjektivoj ĝi indikas, ke temas pri rilato au aparteno, kelkfoje pri aro; konsideru la jenajn kazojn:

<i>plátan-al-</i>	radiko: banan–,
<i>person-al-</i>	radiko: person–.

En la unua kazo la sufikso signifas "ejo, kie io abundas", do "banan–uj–aro, ĉar banan–(*plátan–*) rilatas al plantoj.

En la dua kazo la sufikso povas egale signifi "person–a" au "person–aro". Tiukaze la ĝusta signifo de la vorto estas nur klarigebla de la kunteksto.

Alia ekzemplo estas la sufikso –aje; ĉe verboj ĝi esprimas agadon, au rezulton de ago, ankaŭ lokon. Ĉe substantivoj ĝi signifas aron, rajtojn pagendajn (*pupilaje, peaje*), ankaŭ tempon (*aprendizaje*).

La sufikso –ero ĉe substantivoj esprimas metion (*aduanero* – doganisto), ankaŭ signas la planton (*limonero* – citronujo), lokon (precipe la inaj formoj de la sufikso ekz. *leonera* – loĝejo de leono). Ĉe adjektivoj ĝi esprimas rilaton au apartenon.

Alia komunua sufikso hispana estas –ón. Kiam temas pri korp parto, ĝi estigas posedajn adjektivojn pligrandige funkciajn (*barrigón* – dikventrulo, ulo kun granda ventro). En la aliaj kazoj tiu sufikso rolas nur pligrandige, do estu pripensinde inkluzivi la informon "ĉu korp parto". Tamen, surprizoj estas: la vortoj *rabón, pekón* signifas respektive senvosta k senhara!

Cetere, la sufiksoj hispanaj multas (estas multe pli da sufiksoj ol da prefiksoj; ĉu tio konsistigas econ ĉe la aglutinaj lingvoj?) k kelkfoje la nuancoj esprimitaj de ili ne estas facile au rekte tradukeblaj. La vorto *reyezuelo* konsistas el la radiko *rey* – reĝo k el la sufikso –zuel; la aĉa sufikso –zuel tiukaze entenas ankaŭ malpligrandigan nuancon.

Analizo de morfemoj.

La analizon de la eblaj prefiksoj la sistemo komencu nur kaze de neidentigita radiko. Ekz. *preparar* – prepari ŝajne entenas la prefikson *pre-* – antaŭ, sed tio ne ĝustas. Se la sistemo, de la maldekstra dekstren, literon post litero, identigas radikon, tiukaze kompreneble ne estas prefikso k tiu parto de la analiz–procezo ne okazu; alikaze la sistemo estas devigata analizi la vorton, detranĉi la eblan sufikson k resendi tion, kio restas al la vortar–priserĉa fazo. Nur kaze de identigita vorto estu permesata la analizo de eblaj sufiksoj.

La ordo, lau kiu la sistemo analizu la morfemojn, ne gravos, kondiĉe ke ĝi estu kapabla reordigi ilin post ĉiu nova eltrovo k antaŭ ol eligi la esperantan ekvivalenton. Tamen, pli–malpli strikta ordo estas dezirinda, ĉar tempo–ŝpara.

Neregulaj formoj de radiko.

La neregulaj formoj de ekz. verboj estu enregistritaj en la vortaro kvazau apartaj vortoj (por komputilo ili ja estas apartaj radikoj) kune kun ilia "plena" traduko. Ne estas rekomendinde enregistri ilin kiel simplajn radikojn; tio riskus estigi problemojn. Konsideru la verban formon *son* – ili estas: ĝi estas neregula formo de la verbo *ser* – esti, sed nenio malhelpas al la sistemo pensi, ke temas pri la radiko *son-* apartenanta al la verbo *sonar* – soni. Pro tio la sistemo ĉiam antaŭmezuru la longon de vorto troviĝanta en la teksto; kaze de vorto samlonga kiel radiko vortara k samtempe identa kun ĝi, la sistemo ne serĉu finaĵon por eltrovi informojn pri kazo, tempo, nombro ktp, sed simple eltiru la stokatan ekvivalenton. Alikaze la sistemo enkonduku la analizon de la finaĵo por ĝuste traduki la vorton.

Frazstrukturo.

La sintaksa analizo devas okazi post la morfologia. Kiel trafe skribas JOHNSON, la bulea kriterio pri tio, ĉu la frazo estas bone strukturita, ne sufiĉas, sole uzata, por adekvata tradukado. La sistemo estu kapabla eltrovi, ke en frazo kiel ekz. *el escribio los programas* la vorto *programas*, kvankam identa kun verba formo de la radiko *programar* – programi ne rolas verbe, sed substantive, k do eligi la ekvivalentan "li skribis la programojn". Faka programero devas testi, ĉu la frazstrukturo estas allasebla au ne, k en tia kazo modifi la vortoklasojn de iu(j) vorto(j).

En la uzita modelo, la analizo sintaksa ekas el tio, ĉu la frazo entenas verbon. Ĉar oni konservas la hispanan vortordon, la sistemo provas eltrovi rektan komplementon post la verbo, kaze de verbo transitiva. Nur se ekzistas tauga kandidato, la sistemo serĉas antaŭ la verbo. Ĉar en la hispana same kiel en la esperanta estas akordo inter substantivo k adjektivo, oni ne bezonas detalegan analizon por eltrovi, kiuj paroj rilatas inter si.

Apartan malfacilon prezentas la frazoj enhavantaj kunordigajn konjunkciojn, ĉar ili ofte rilatas al elipso. Kaze de elipso tauga solvo povas eki el la fakto, ke en tiaj kazoj la konjunkcio kunordigas "kompletan" branĉon kun "nekompleta" k do la mankantajn informojn oni povas eltiri el la unua k sendi al la dua dum la analizo. Tiu ĉi vojo estas ankoraŭ traierota de KOMENCANTO.

4. Rezultoj.

Kiel oni povas vidi, eĉ simpla, poresplora sistemo funkcia en malgrandkapacita komputilo taugas por prilumi la aferojn tradukajn k por eltiri interesajn konkludojn. Multaj el la problemoj povas esti solvitaj rekte en la komenco de la procezo kondiĉe ke la elirpunkto estu zorge elektita. La plursignifo de vortoj ekz. ne konsistigas tiom da grandaj problemoj, se oni uzas vortaron fakan k se oni provas traduki nur tekston fakan; nur frazoj ĝustaj lau la gramatika vidpunkto estu analizataj; la rezulto de la tradukprocezo tiam estas ofte kontentiga.

5. Literaturo.

Fumero, Alberto García, Skiza komparo inter la hispana kaj la esperanta lingvoj cele al perkomputila traduko de tekstoj hispanaj, Scienca Revuo 43, (1992)(2)-161.

Maas, Heinz Dieter, Persona Komunikado k artikoloj pri la traduksistemoj SUSY k SUSY-II: Perkomputila tekstoprilaboro, Scienca Eldona Centro, 1985.

Johnson-Parsing, R. L., MT perspektive, Automatic Natural Language Parsing, Ellis Horwood Series, 1983.

Schubert, Klaus, Syntactic Structure in DTL, BSO/Research, Utrecht, 1986.

Adreso de la aŭtoro:

Alberto García Fumero

Teknika Lernejo pri Manĝaĵoj "Ejército Rebelde"

Persona adreso:

Calle 26 # 3313 entre 33 y 35 Playa

C 11300 Ciudad de la Habana

CUBA.