

## STROJNO PREVOĐENJE I NAŠI NEPOSREDNI ZADACI

*Božidar Finka i Bulcsú László*

Obavijesne teškoće. U 50 milijuna znanstvenih tiskovina čovječanstva leži zakopano golemo blago. U njemu je nagomilana vjekovita energija čovječjega duha, usmjerena da preobliči prirodu u kojoj živimo, i nas same. Tu uskladištenu energiju zovemo informacija ili obavijest. Ali ona je tako duboko zakopana na dnu knjižnica i arhiva, knjiga, časopisa i rukopisa da je njeno iskapanje mnogo teže i skuplje od iskapanja ikojeg zemnog blaga. Nije bez razloga što znanstvenici tvrde da je danas lakše izumjeti nešto posve novo nego pronaći je li već izumljeno i gdje god opisano. Kakva li besmislena traćenja vremena i rasipanja snage!

Znanstvenu obavijest, najveće blago čovječanstva, neustrašivo čuvaju tri čudovišta, i dok njih ne svladamo, i blago će ostati ležeći na ugaru. To su prostorna, strukovna i jezična nepristupačnost podataka.

Prostorna nepristupačnost. Otkako je izumljen tisak, štampa je omogućila širenje i snaženje znanosti i množenje nauka te je danas, rastući zajedno s njima, prerasla samu sebe i postala kočnica njihova razvoja. Jedini je izlaz prekodiranje (tj. preznakovljenje) svekolike literature u takvo osobito pismo da je može stroj strelovitom brzinom čitati, na malenu prostoru uskladištavati i po potrebi odmah pronalaziti. Postoje brzočitači koji između podataka snimljenih na obrtjive koture odabiru željenu obavijest brzinom od 200.000 znakova u sekundi (npr. pri prevođenju riječi iz rječnika). Postoje uređaji koji za tren oka zapamte i za tren se oka prisjete milijunâ riječi (a to su sve riječi našega jezika u svim njihovim oblicima). Ipak to nije dosta.

Čovjeku koji provede 50 godina u živoj i mnogostranoj umnoj djelatnosti prođe kroz pamet milijarda i po riječi. Da bi se stroj mogao mjeriti s takvim istaknutim intelektualcem ne samo po brzini nego i po kakvoći izvršivanja zadataka, treba da ima pamćenje i jezično iskustvo istoga kapaciteta, a za to dakako — to je i najvažnije — treba da ima ugrađene vodove za združivanje pojmova i za logičko provjeravanje i zaključivanje, jednom riječju sposobnost brza učenja.

Strukovna nepristupačnost. Druga je kočnica suvremenoga napretka takozvana difuznost informacije, tj. rasutost obavijesti po strukama. Nekad jedinstvena i cjelovita znanost, obuhvaćena pod nazivom »philosophia« (ljubav prema mudrosti), počela se s vremenom, najprije polako a zatim sve brže, raspadati na bezbroj struka i stručica sa strogo razgraničenim predmetom i područjem istraživanja, tako da smo danas došli do apsurdne činjenice da se stručnjaci različitih struka ne razumiju. Kao reakciju na to, naše je doba nametnulo potrebu zbliživanja najrazličitijih

stručnjaka, kolektivizaciju naučnog rada. Donedavna su se na primjer jezikom bavili jezičari, sami samcati — bespomoćno; danas su se na nesagledljivome poslu proučavanja i modeliranja jezika pomoću stroja i za stroj najednom našli uz jezičare i njima najbliže, stručnjake za književnost još i filozofi, psiholozi, logičari, neurolozi, matematičari, telekomunikacijski stručnjaci (priopćajci, »vezisti«), elektroničari, kibernetičari. Da bismo dakle mogli istražiti zajednički jezik, treba da najprije tražimo zajednički jezik sa svima tim našim suposlenicima.

**J e z i č n a n e p r i s t u p a č n o s t.** Da ukloni treću, najveću teškoću u stjecanju stručnih obavještenja, jezičnu rasparčanost i izoliranost stručnjaka i stručne građe, javilo se kao jedino današnje rješenje — strojno prevođenje. Kad bi sva stručna literatura i bila pristupačna stručnjacima svijeta — npr. skupljena u pamćenju džinovskog stroja koji bi na zahtjev za čas pronalazio sve podatke iz sviju struka o kakvu pitanju — polovina bi naučnih spoznaja opet bila ograđena od njih neprobojnim jezičnim zidom. Ta koji stručnjak ima vremena, mogućnosti i snage naučiti pedesetak jezika da bi — barem u svojoj struci — pratio u korak znanstvena zbivanja današnjice? Nema druge, treba što prije omasoviti i mehanizirati proces prevođenja, a to je moguće samo ako se stvori sposoban, brz i neumoran stroj koji bi mogao nasljedovati prevodilački tok, pa bilo to i šablonski. Stroj treba da misli za nas. Ako već pomaže čovjeku u fizičkom poslu, zašto mu ne bi pomogao i u umnom poslu — svojim mehaničkim pamćenjem i automatskom logikom.

Kada već uočavamo teškoće oko prenošenja obavijesti i nedostupnosti podataka, valja, koliko mognemo, da im i sami doskočimo. Zato ćemo sad ovdje najkraće izvijestiti što se oko strojnoga prevođenja uradilo u nas, kakvo je stanje u svijetu i koji su naši neposredni zadaci.

**S t a n j e u z e m l j i.** Organiziran i sistematski rad u vezi sa strojnim prevođenjem počinje se u nas u Zagrebu 1959. godine. Pojedinačna zanimanja za taj problem datiraju nekoliko godina prije. Prof. dr. Vladimir Matković, pročelnik Zavoda za telekomunikacije u Zagrebu, već je 1957. godine obranio doktorsku disertaciju o entropiji hrvatskoga jezika s područja obavijesne teorije, koja omogućuje pristup k strojnom prevođenju. 30. IV 1958. godine asistent Filozofskoga fakulteta u Zagrebu Bulcsú László na sastanku Kruga mladih lingvista (sada Zagrebačkog lingvističkoga kruga) izlaže na praktičkim primjerima primjenljivost redigiranja teksta pomoću stroja. Nezavisno od toga u jesen 1958. godine, zanimajući se za teoriju prevođenja, Svetozar Petrović, asistent Filozofskoga fakulteta u Zagrebu, počinje se baviti problemima strojnog prevođenja. 6. I 1959. B. László, kao rukovodilac Kruga, i 9. VI 1959. S. Petrović, kao stručnjak za komparativnu književnost, upoznaju članove Kruga s mogućnostima i vidicima primjene strojnoga prevođenja i prihvataju se uredništva budućega Zbornika. Uzajamnom suradnjom desetera zainteresiranih članova Kruga pokrenut je studij

teoretskih postavaka strojnoga prevođenja. Rezultat je toga bio niz predavanja, u kojima su sudjelovali profesori tehničkih fakulteta dr Vladimir Vranić, pročelnik Centra za numerička istraživanja JAZU, dr Vladimir Muljević, pročelnik Zavoda za regulacione i signalne uređaje, i dr Vladimir Matković, zatim članovi Zagrebačkoga lingvističkog kruga: Stjepan Babić i Krunoslav Pranjić sa serbokroatistike, Željko Bujas i Leonardo Spalatin s anglistike, Božidar Finka iz Instituta za jezik JAZU, Malik Mulić s rusistike i Bulcsú László s poredbene slavistike. Ta je djelatnost članova Kruga stvorila zbornik

### *STROJNO PREUODENJE I STATISTIKA U JEZIKU*

koji je izašao 1959. godine kao posebni prilog 6. broju časopisa »Naše teme«, na dvjestotinjak strana s bibliografijom o strojnom prevođenju — prvom na svijetu.

Osvrćući se na taj zbornik u sovjetskom vodećem lingvističkom časopisu »Voprosy jazykoznanija«, Moskva 1960/4, istaknuti je sovjetski lingvist N. D. Andrejev, profesor Lenjingradskog sveučilišta, naglasio da po tom zborniku Jugoslavija kao deseta zemlja na svijetu (poslije SAD, Velike Britanije, SSSR, Kine, Rumunjske, Čehoslovačke, Japana, Italije i Francuske) ulazi u red onih zemalja u kojima se započela konkretna organizirana djelatnost u proučavanju matematičke lingvistike i strojnoga prevođenja. Pozdravljajući pojavu našega zbornika prof. Andrejev je izrazio nadu da će autorski kolektiv uskoro objelodaniti s toga područja i praktičke rezultate. Da bismo takve rezultate doista mogli postići, potrebna nam je društvena pomoć.

Stanje u svijetu. Jezik je najvažnije sredstvo ljudskoga priopćavanja, opće sredstvo prenošenja obavijesti. Suvremeno proučavanje jezika zahtijeva suradnju, međusobno poređenje i dopunjanje rezultata stručnjaka u raznim znanstvenim disciplinama u svrhu neposredne društveno korisne upotrebe. Tako nastaje novo istraživačko područje — primijenjena lingvistika. Strojno je prevođenje grana primijenjene lingvistike koja proučava modele jezičnoga ustrojstva prikladne za preobličavanje tekstova iz jednoga jezika automatski, mehanički, pomoću stroja u drugi jezik.

Prvi glas u javnosti čuo se o strojnom prevođenju 1951. godine. Prvo je pokusno strojno prevođenje izvršeno u SAD 1954. godine. 1955. godine izlazi prva knjiga radova o tom velikom znanstvenom problemu. Iduće se godine priključuju Velika Britanija i SSSR, a poslije i druge zemlje. Među njih se 1959. godine uvršćuje Jugoslavija kao deseta zemlja, a poslije nje DR Njemačka, Poljska, Švedska, Brazilijska, Meksiko, Mađarska, Nizozemska, Indija, Izrael i dr. Tako se širi napredna znanost, koja revolucionira proučavanje jezika. U zemljama okupljenim oko Sovjetskog Saveza zajedničkim

se silama proučavaju, među mnogim drugim jezicima, i svi slavenski jezici za svrhe strojnoga prevođenja s *izuzetkom našega, hrvatskosrpskoga jezika*.

Kako se strojno prevođenje ne može ni zamisliti prije nego što se dobije potpun, temeljit i precizan opis svih jezika s kojih i na koje želimo strojem prevoditi, potrebno je izvršiti goleme predradnje primjenjujući tekovine egzaktnih nauka: statistike, vjerojatnosne i obavijesne teorije, matematičke logike, teorije skupova i grafova i kibernetike. Osim toga što ima neprocjenljivu teoretsku važnost, jezično modeliranje neposredno zadire u pitanja automatizacije, uskladištavanja i nalaženja znanstvene obavijesti, govornog upravljanja proizvodnim postrojenjima, konstruiranja čitajućih i pišućih naprava, povećanja propusne sposobnosti kanala žičanog i bežičanog priopćavanja i mnogih drugih područja. Matematiziranje je jezikoznanstva omogućilo da su suvremeni istraživači jezika mogli postaviti sebi konkretan zadatak: obrađivanje jezika radi stvaranja ekonomičnoga prevodilačkoga stroja. Ostvarivanje toga cilja neće imati samo izuzetno društveno, ekonomsko i kulturno značenje već će pomoći da se i mnoga pitanja samoga jezika, uključujući ovamo i »tajanstvo« čovjekova mišljenja, shvate i egzaktno opišu. Nema nikakve sumnje da će studij strojnoga prevođenja donijeti znanosti koristi koje će možda biti važnije i od njegove — same po sebi velike — neposredne praktične primjene kao robota prevodioca.

Velika očekivanja od prevodilačkog stroja iziskuju i velika ulaganja i napore. Izrada ni jednoga drugog stroja u povijesti nije tražila toliko napora i tolike pripreme radove. Primjerom može poslužiti SSSR. Tamo ministar Višega obrazovanja već je 28. XI 1958. izdao naredbu br. 1228 o razvoju znanstvenih istraživanja sa područja strojnoga prevođenja koja obavezuje velik broj jezičnih i tehničkih ustanova da pruže »svemjernu« pomoć uspostavljanju i razvijanju znanstvenih kolektiva za matematizaciju i tehnizaciju jezičnih proučavanja i posebno za automatizaciju i za strojno prevođenje. Rezultat je toga da se danas preko osamdeset sovjetskih državnih jezičnih i tehničkih ustanova bavi problemima strojnog prevođenja, i to u vezi s najrazličitijim i najegzotičnijim jezicima.

Naši neposredni zadaci. Očito je da je za tako velik i obuhvatan zadatak potrebno udružiti stručnjake svih zainteresiranih znanstvenih grana. Stoga smo 22. II. 1961. godine organizirali

### *SKUP STRUČNJAKA ZA STROJNO PREVOĐENJE.*

Prvenstveni je lingvistički zadatak Skupa da izradi:

1. hrvatskosrpski odostražni rječnik (odòstrāžnik)
2. hrvatskosrpski čestotni rječnik (čestòtnik)
3. rječnik hrvatskosrpskih značenja poklapanja (tezaurs).

Tek će se nakon toga moći načeti pitanje temeljnih konfiguracija (sklopova riječi) i slično.

Odostražnik će omogućiti da se dobiju i sistematiziraju svi tipovi nastavaka, da se odredi broj i utvrdi čestotni i hijerarhijski (važnosni) red. Napokon će se dobiti podaci za ekonomično kodiranje (uznakovljavanje) svake riječi s obzirom na strojno prevođenje.

Čestotnik će pokazati kojih se nekoliko tisuća riječi najčešće pojavljuje u jeziku, a to je od velike važnosti i za određivanje obujma pamćenja prevodilačkog stroja.

Tezaurus bi obuhvatio oko 1.000 riječi natuknica, jer se svi pojmovi mogu svesti na kojih 1.000 značenjskih skupova. Sve bi se ostale riječi svrstale, svaka prema svom značenju, u te značenjske skupove — nosioce glavnoga značenja. Pri prevođenju bi izbor prevodilačkoga stroja time svaki put bio sužen na određeno semantičko područje. To bi omogućilo ekonomično korištenje prijenosnih vodova i brzinu prevođenja.\*

## TEČAJ HRVATSKOSRPSKOG JEZIKA PO GLOBALNO-STRUKTURALNOJ AUDIO-VIZUELNOJ METODI<sup>1</sup>

*Emia Leskovar — Krunoslav Pranjić*

Uz engleski, francuski, njemački, talijanski i ruski, hrvatskosrpski je šesti jezični tečaj što su ga, prema načelima (AV) metode, izradili suradnici Zavoda za fonetiku Filozofskog fakulteta Sveučilišta u Zagrebu. Namijenjen je svakome tko hrvatskosrpski uči kao strani jezik. Autori metode su: P. Guberina, sveučilišni profesor, ujedno i pročelnik pomenutog Zavoda i P. Rivenc, profesor na Ecole Normale Supérieure u Saint-Cloudu, Francuska.

Hrvatskosrpski, dakle, posljednji je, zasad, u nizu takvih tečajeva. ali bi se prema dobro poznatoj engleskoj izreci, moglo gotovo poslovično reći: posljednji, ali ne i najmanji, što će reći: ne i najvažniji.

Potpun naziv ove metode zapravo je širi: globalno-strukturalna audio-vizuelna (AGVS). Smatramo potrebnim, prije negoli išta kažemo o specifičnim teorijski relevantnim problemima koji su iskrslili u toku pisanja i tonskog snimanja hrvatskosrpskoga jezičnog tečaja, da u minimalnom sažetku kažemo nešto i o općim lingvističkim koncepcijama i teorijskim temeljima AGSV metode same.

\* U idućim brojevima »Jezika« nastojat ćemo potanje opisati načela izrade svakoga od tih rječnika.

<sup>1</sup> Autori: dr Lj. Jonke, E. Leskovar, K. Pranjić; izdanje: Zavod za fonetiku Filozofskog fakulteta Sveučilišta u Zagrebu, Zagreb 1962.