

LINGVISTIKA I INFORMATIKA

U članku se sažeto razmatra predmet i područje istraživanja lingvistike i informatike i ukazuje se na neka dodirna područja istraživanja tih znanosti. Dodirnim se područjima smatraju sažimanje i indeksiranje, izrada tezaurusa, terminološka istraživanja i oblikovanje terminoloških baza podataka, umjetna inteligencija, automatsko prepoznavanje i sinteza govora, strojno razumijevanje prirodnih jezika i prevođenje pomoću računala.

Lingvistika; informatika; informacijska znanost

1.

1.1. Lingvistika se obično definira kao "znanstveno proučavanje jezika" (Crystal 1985:76). Odrednica "znanstven" podrazumijeva da se istraživanje zasniva na empirijski provjerljivim podacima u okviru pojedine opće teorije jezične strukture kao sustava načela i kategorija kojima se objašnjavaju jezične činjenice. U spomenutoj definiciji "jezik" označava opće svojstvo čovjekovo, jezik uopće, za razliku od drugih sustava komuniciranja, ali isto tako i lingvistički sustav kojim komunicira određena ljudska zajednica, bilo da se radi o apstraktnom sistemu ili njegovim očitovanjima u govoru ili pismu pojedinaca ili skupine pojedinaca. Predmet je lingvističkog istraživanja, dakle, svaki pojedini jezik i jezik uopće, potraga za spoznajama o zajedničkim svojstvima jezika kao općeljudskog fenomena, na temelju izučavanja svih jezičnih pojava u svim ljudskim jezicima. Lingvistika je zapravo osobita među znanostima jer je jezik i predmet i sredstvo njene analize: jezikoslovci raščlanjuju jezik služeći se jezikom.

1.2. Suvremeno jezikoslovlje poklanja veliku pozornost teorijskim pitanjima. Da bi uspješnije rješavala brojne praktične zadatke i unaprijedila izučavanje svojega vlastitog predmeta, lingvistika izgrađuje svoju teoriju kao skup teorijskih načela, opisnih kategorija i metodoloških postupaka pomoću kojih se može odrediti ljudski jezik kao pojava, opisati sličnosti i razlike među jezicima, te objasniti prostorne, vremenske i društvene varijacije jezika (Bugarski 1985:3). Dakako, bez obzira na zrelu dob lingvistike kao znanosti, i danas se može reći da je lingvistička teorija još uvijek u stanju burnog previranja. Iznose se različite teorije jezika, neke u okviru istog osnovnog pogleda na jezične činjenice, a druge opet s različitog ili, čak, suprotnog polazišta. Iako se u izgradnji lingvističke teorije hipoteze obično uspostavljaju induktivnim postupkom, na temelju promatranja jezičnih događaja, a zatim se empirijski provjeravaju i ugrađuju u teorijski sustav, u lingvistici je također proširen i tzv. hipotetičko-deduktivni pristup. U ovom drugom slučaju polazi se od stanovitih teorijskih pretpostavki, koje čine okvir za

odabir i opis jezičnih događaja, pri čemu se predviđaju sve logičke mogućnosti koje proizlaze iz zadanih pretpostavki i provjeravaju na analizi stvarnog jezičnog materijala (Katičić 1967:4-5), na temelju čega se stvara nova teorija i opet provjerava na materijalu koji se istražuje. U okviru ta dva osnovna pristupa razvile su se brojne jezične teorije i modeli koje čine teorijski fundus lingvistike: tagmatika, glosematika, funkcionalna gramatika, stratifikacijska gramatika, generativno-transformacijska teorija, generativna semantika i dr.

1.3. Proučavanju jezičnih pojava može se pristupiti s različitim svrhom i u okviru različitih teorijskih opredjeljenja. S obzirom na vremensku dimenziju možemo govoriti o opisu pojedinog jezičnog sistema u danom vremenskom trenutku (sinkronijski pristup) ili o opisu njegova mijenjanja u vremenskom toku (dijakronijski pristup). Ili, pristup jezičnim činjenicama može biti pretežno usmjeren u pravcu razvijanja lingvističke teorije, ali se proučavanje, isto tako, može pretežno vršiti sa svrhom da se riješi ili pokuša riješiti kakav praktičan zadatak koji suvremeni život stavlja pred znanost o jeziku, u kojem slučaju se radi o primijenjenom istraživanju. Ako se pak istraživanje vrši u svrhu uspoređivanja dvaju ili više jezičnih sustava, govori se o poredbenim ili komparativnim istraživanjima, pri čemu se istraživanje može vršiti na sinkronijskoj ili dijakronijskoj osnovi. Uz to, s obzirom na jezičnu teoriju koja se primjenjuje, možemo govoriti o generativno-transformacijskom, tagmemičkom, stratifikacijskom ili kakvom drugom teorijskom pristupu jezičnim činjenicama, tako da se različiti kriteriji i pristupi često međusobno ukrštavaju i isprepliću.

1.4. Kako je jezik veoma složen sustav, nemoguće ga je u isto vrijeme opisati u cjelini, već se u znanstvenom raščlanjivanju pažnja usmjerava na različite, mada usko povezane, aspekte jezika. Različiti aspekti nazivaju se razine jezične analize (Robins 1964:12), a razlikovanje pojedinih razina i njihov broj uvelike ovisi o teorijskom pristupu koji se primjenjuje. U evropskoj lingvističkoj tradiciji uobičajeno je razlikovanje fonetske, gramatičke, leksičke i semantičke razine jezične analize. Tim se razinama jezičnog proučavanja bave zasebne lingvističke discipline: fonetika, gramatika, leksikologija i semantika. Unutar tih područja izučavanja moguće je izdvojiti i uža područja koja su u dovoljnoj mjeri izgradila svoju teoriju i metodologiju i smatraju se zasebnim lingvističkim disciplinama.

1.4.1. Fonetika izučava zvukovnu stranu jezika i razmatra glasovne i prozodičke elemente (naglasak, ritam, intonacija) bez obzira na njihovu funkciju u pojedinom jeziku. To je proučavanje usmjereno na opis i klasifikaciju artikulacijskih (fizioloških), akustičkih (fizičkih) i perceptivnih osobina glasova i prozodičkih elemenata, pa su se, u vezi s tim, razvila tri izražena područja fonetskih istraživanja: artikulacijska fonetika, akustička fonetika i auditivna fonetika. S tog je razloga fonetika usko povezana s fiziologijom, akustikom i psihologijom, znanostima koje joj pružaju metode i postupke za različite aspekte fonetske analize. Kako fonetika proučava zvukovnu stranu jezika bez obzira na funkciju i značenje, funkcijskom stranom glasova i prozodičkih elemenata bavi se posebna disciplina koja se naziva fonologijom. Ona proučava foneme kao razlikovne glasovne jedinice koje mogu mijenjati značenje jezičnih elemenata. S obzirom na važnost toga vida jezičnog izučavanja u suvremenoj lingvistici, fonološki se aspekt često smatra zasebnom razinom jezične analize, mada treba naglasiti da su fonetski i fonološki vid proučavanja jezika usko povezani i praktički neodvojivi.

1.4.2. Gramatika opisuje promjene u oblicima riječi i načine na koje se ti oblici slažu u skupove riječi i rečenice. Tako unutar gramatike postoje dva izraženja na vida jezičnih izučavanja: morfološki i sintaktički. Tim dvama vidovima proučavanja bave se i zasebne gramatičke discipline: morfologija i sintaksa. Morfologija

se bavi promjenom oblika riječi radi izražavanja gramatičkih odnosa i svrstavanjem riječi u gramatičko-leksičke razrede. Ponekad se unutar gramatike, kao dio morfologije, uključuje i tvorba riječi (Babić 1986:11-12), mada je uobičajenije da se to područje izučavanja uvrsti unutar leksikologije. Sintaksa je, s druge strane, usredotočena na odnose i procese koji se očituju u stvaranju onih gramatičkih jedinica koje su veće od riječi, što znači da se bavi skupovima riječi i rečenicama i njihovim unutrašnjim ustrojstvom. U novije vrijeme razvila se i posebna grana jezikoslovlja, usko povezana sa sintaksom, koja se bavi nadrečeničnim jedinicama i poznata je pod imenom analiza diskursa ili, kad je u pitanju pisani jezik, lingvistika teksta (Crystal 1985:247).

1.4.3. Leksikologija je grana jezikoslovlja koja sustavno izučava rječnički sa stav jezika. U njeno područje izučavanja ulazi pitanje određenja riječi kao temeljne jezične jedinice, struktura značenja riječi, morfološki sastav riječi, tvorba riječi, leksički skupovi riječi i frazeološke jedinice, klasifikacija leksičkih elemenata s obzirom na značenje i upotrebu i dr. Dakako, svi se istraživači jezika ne slažu o tome koja područja treba izučavati unutar leksikologije, pa tako neki lingvisti tvorbu riječi uključuju u područje gramatike (1.4.2), a neki opet proučavanje skupa riječi i frazeoloških jedinica izdvajaju u posebnu lingvističku disciplinu koja se naziva frazeologijom (Kunin 1970:7-10). U vezi s područjem izučavanja leksikologije treba posebice naglasiti da se u novije vrijeme unutar leksikoloških istraživanja znatna pažnja posvećuje proučavanju terminoloških sustava i terminološkog leksika (Fromkina 1978:176-182). S leksikologijom je usko povezana i leksikografija, lingvistička disciplina koja se bavi teorijom i praksom izrade rječnika.

1.4.4. Semantika je grana lingvistike koja se bavi proučavanjem značenja što ga izražavaju različite jezične jedinice kao što su morfemi, riječi, skupovi riječi, klauzule i rečenice. Iako se semantika shvaća kao zasebna lingvistička disciplina, pitanja značenja prisutna su na svim razinama jezične analize pa se tako semantičkim aspektima neminovno bave sve grane jezikoslovlja. Značenje pojedine jezične jedinice može se proučavati s tri osnovna stajališta: odnos prema predmetima i pojavama izvanjezičnog svijeta, odnos prema drugim jezičnim elementima i odnos prema govornicima koji se jezikom služe. Unutar tih odnosa moguće je razlikovati više tipova značenja među kojima središnje mjesto zauzima tzv. denotativno ili kognitivno značenje (Leech 1974:10-13). Dakako, proučavanjem značenja bave se i druge znanstvene discipline kao što su filozofija, logika, teorija komunikacije i dr. (Palmer 1976:15), pa tako semantika predstavlja područje izučavanja i unutar drugih znanosti. S tog se razloga grana lingvistike koja se bavi izučavanjem značenja može preciznije označiti nazivom lingvistička semantika.

1.5. Budući da je jezik povezan sa svim područjima ljudske djelatnosti, neminovno je da se lingvistika kao znanost o jeziku dodiruje i preklapa s mnogim drugim znanstvenim disciplinama. U toku osamostaljivanja lingvistike kao znanosti, međutim, jezikoslovci su uvelike inzistirali na njoj neovisnosti o drugim znanostima. Razlog je u tome što je proučavanje jezika u prošlosti bilo podložno snažnom utjecaju drugih znanstvenih disciplina, osobito filozofije i logike, koje su često stvarale iskrivljenu sliku o jeziku i jezičnim pojavama. Takav stav pogodovao je razvoju lingvistike kao znanosti i omogućio je studij jezika kao formalnog sistema. No, nakon što se potvrdila kao samosvojna znanstvena disciplina s razvijenim teorijskim i metodološkim aparatom, nije više toliko izražena potreba da se ustraje na načelu autonomije u lingvističkim istraživanjima, a zanimanje za pristup jeziku sa stajališta drugih znanosti uzelo je šireg zamaha (Lyons, ed. 1970:8-9) i otvorilo nova područja istraživanja jezičnih pojava. Tako je u suvremenom jezikoslovlju veoma izražena težnja prema interdisciplinarnim istraživanjima, a neka područja istraživanja jezika na razmeđu između lingvistike i drugih znanosti, npr. so-

ciolingvistika i psiholingvistika, izrasla su u relativno samostalna i veoma aktivna područja suvremenih istraživanja jezika.

2.

2.1. Iako je izraz informatika široko prihvaćen kao naziv znanosti koja se bavi proučavanjem informacija i informacijskih procesa, još smo uvijek daleko od toga da postoji široka suglasnost o njenom predmetu i području izučavanja. Dapače, s obzirom na univerzalnost informacijskih procesa, teško je odrediti i granice te znanosti u odnosu na druga područja izučavanja (Harvey 1967:421). To nije ni čudo jer je informatika razmjerno mlada znanost koja se značajnije razvila od pedesetih godina ovog stoljeća pa stoga još uvijek nema dovoljno razradenu teorijsku osnovu i metodologiju istraživanja. Štoviše, ne samo što postoji neslaganje u pogledu predmeta kojim se ona bavi već ni sam naziv te znanosti nije podjednako rasprostranjen u svijetu. Dok je naziv informatika općenito prihvaćen na evropskom kontinentu, u anglosaksonskoj literaturi taj se naziv rijetko spominje (Brookes 1972: 138) jer se obično upotrebljava izraz informacijska znanost. Pod utjecajem anglosaksonskih izvora naziv informacijska znanost počinje se šire koristiti i u hrvatskom ili srpskom jeziku (Kržak 1975, Tudman 1986), pa se ta dvojnost naziva već očituje i u našim uvjetima.

2.2. Iako se u stručnoj literaturi nazivi informatika (informatics) i informacijska znanost (information science) uglavnom smatraju istoznačnim s obzirom na predmet i područje istraživanja (Ralston i Meek 1976:656), definicije informatike (informacijske znanosti) pokazuju da u poimanju tog područja izučavanja mišljenja nisu uvijek podudarna. Tako se npr. u Sovjetskom Savezu pod informatikom prvenstveno razumijeva znanstvena disciplina "... koja proučava strukturu i opća svojstva znanstvenih informacija kao i osnovne zakonitosti svih procesa znanstvene komunikacije" (Mihajlov i Giljarevskij 1983:1). U vezi s takvim određenjem treba ipak dodati da se u novije vrijeme u Sovjetskom Savezu unutar informatike uključuje i cjelokupno područje informacijske tehnike i tehnologije, a postoje i manje proširena shvaćanja informatike kao opće znanosti o informacijama (usp. Semenjuk 1984:2-7). Nasuprot tome, u SR Njemačkoj naziv informatika (Informatik) označava računalnu znanost (Bauer i Goos 1973: VII, Ralston i Meek 1976:656). Prema definiciji koja je nastala 1962. godine u Americi, a koja je i danas široko prihvaćena, informatika (information science) se određuje na slijedeći način: "Znanost koja istražuje svojstva i ponašanje informacija, sile koje upravljaju informacijskim tokovima te sredstva za obradu informacija radi njihove optimalne pristupačnosti i iskoristivosti. Ti procesi obuhvaćaju nastajanje, diseminaciju, prikupljanje, organiziranje, pohranu, dohvat, tumačenje i upotrebu informacija" (Shera i Cleveland 1977:265). Ili, nešto drukčijim riječima, a prema definiciji koju navode Zunde i Gehl (1979:68), "...informacijska znanost je proučavanje prirode informacije kako se ona očituje u različitim pojavama koje su povezane s generiranjem, prijenosom, pretvorbom, skupljanjem, pohranom i drugim sličnim procesima". Ni kod nas, isto tako, ne postoji potpuno suglasje o predmetu informatike i njenim područjima izučavanja. Tako jedni autori informatiku definiraju "... kao znanstvenu disciplinu koja istražuje sastav, funkcije, oblikovanje, provedbu i rad informacijskih sistema s kompjutorskom podrškom" (Ferišak, ur. 1981:6), dok drugi smatraju (Han i Balaban 1983:16) da "... ova mlada disciplina u osnovi nije vezana za elektronske sisteme za organizovanje podataka..." Informatika se, dakle, s jedne strane shvaća kao opća znanost o informacijama i informacijskim procesima, dok se, s druge strane, predmet njena istraživanja ograničava na bavljenje određenom vrstom informacija ili na obradu informacija pomoću elektroničkog računala. U tom pogledu situacija se, izgleda, nije bitnije izmijenila za posljednjih desetak godina, od vre-

mena kada je pregled shvaćanja informatike sustavno razmotren u veoma korisnom Kržakovom članku (Kržak 1975). Stoga, da bi se izbjegli mogući nesporazumi, u daljnjem izlaganju informatike će se shvaćati u širem značenju, a u skladu s definicijom koju prenose Shera i Cleveland (1977:265).

2.3. Ako se informatika poima kao znanost koja se bavi informacijama i informacijskim procesima, onda je očigledno da je podrobnije njena istraživanja veoma široko. Ona se u prvom redu mora pozabaviti pojmom informacije te njenom prirodom i svojstvima kao osnovnim fenomenom svojega izučavanja (Zunde 1981:341). To se izučavanje ne može ograničiti samo na vjerojatnostno-statistička svojstva informacije, u smislu Shannonove teorije, već se mora šire pozabaviti i sadržajem informacije, tj. njenim semantičkim svojstvima. Ovakvo široki pristup proučavanju informacije, istina, gdje kad se stavlja u zadatak općoj znanosti o informacijama, koja nije istovjetna s informatikom, i za koju se predlaže naziv informatologija (Ralston i Meek 1976:657, Težak 1971:10) ili informologija (Semenjuk (1984:6-8). Kako se u takvim razmatranjima informatika mora neminovno shvaćati u užem smislu, postavlja se pitanje da li proučavanje općih svojstava informacije ulazi unutar predmeta njena istraživanja i kako treba sagledavati informatiku unutar šireg područja informatologije. Budući da je ovakva podjela tek nagoviještena kao teorijska mogućnost i nije šire prihvaćena, proučavanje općih svojstava informacije ostaje jedan od temeljnih zadataka informatike kao znanosti, bez obzira shvaćali je u širem ili užem smislu.

2.4. Proučavajući informacijske pojave i procese, informatika teži da dopre do dubljih spoznaja o prirodi informacije i da uspostavi opća teorijska načela za proučavanje, objašnjavanje i predviđanje različitih pojava povezanih s pojmom informacije. U tim proučavanjima često se polazi od postavke da informacija nije izravno dostupna promatranju niti je izravno mjerljiva, već da se može spoznati preko drugih pojava, od kojih su neke mjerljive. Tako s npr. Shannonova teorija zasniva na uzajamnoj vezi između vjerojatnosti događaja i informacije koja se njome prenosi: što je događaj manje vjerojatan, to se njegovim očitovanjem prenosi više informacija. Takvo shvaćanje informacije pogodno je za tehničke svrhe, za reprodukciju poslanih poruka na određeno mjesto. Poruka se u ovom smislu uzima kao skup znakova koji je istovjetan odaslanim znakovima, a to što ti znakovi nose određeno značenje, tj. stoje u određenom odnosu sa stanovitim fizičkim ili misaonim entitetima, nije relevantno sa stajališta tehnike (Zunde 1981: 346). Prema mišljenju informatičkih teoretičara (Shera i Cleveland 1977: 261, Brookes 1972:139) ova teorija, koja se često smatra granom matematike, nije za sada primjenjiva na bilo koju vrstu informacijskih procesa i nije izrazitije utjecala na informatiku kao znanost. Ona obuhvaća jedan aspekt informacije i samo djelomično osvjetljava taj složeni fenomen, što ne znači da u sklopu drugih empirijskih načela i zakona ne daje svoj doprinos informatici kao znanosti. Informaciju kojom se bavi informatika treba svakako shvaćati znatno šire, mada je i širina shvaćanja informacije problem sam po sebi. Ima mišljenja, naime, da je informacija utjelovljena u svim predmetima i procesima univerzuma, i prirodnim i umjetnim, i da se iz njih može izlučiti ako je to potrebno i kada je to potrebno (Rzevski 1985:201). Informacija mijenja naše znanje o univerzumu na taj način što smanjuje našu nesigurnost o predmetima i pojavama koje ga tvore. Stanoviti artefakti upravo se i stvaraju s ciljem da budu izvori ili nositelji informacija. Sposobnost izlučivanja informacija iz pojedinog izvora ili nosioca povezana je s inteligencijom. U tom smislu informacija je relativna jer će različiti umovi, s obzirom na opseg i strukturu znanja koje posjeduju i sposobnost kojim raspolažu, izlučiti različite informacije iz istog izvora. Obrada se informacija stoga može promatrati kao postupak obavljanja operacija na nosiocima informacija sa svrhom da se sadržaj informacija tako prilagodi da ga mogu optimalno izlučiti oni kojima je namijenjen. Predmet proučavanja informa-

tike usredotočen je na formalne sisteme prikazivanja i obrade informacija, mada se veza s prirodnim sistemima ne smije i ne može prekidati. I u okviru informatike, smatra se (Stamper 1985:198), informaciju treba proučavati sa semiotičkog stajališta i pažnju posvećivati trima osnovnim vidovima informacije: pragmatičkom, semantičkom i sintaktičkom. Pragmatika se bavi upotrebom znakova od strane ljudi, semantika proučava odnose između znakova i entiteta na koje znakovi upućuju, a sintaksa istražuje oblik i strukturiranje znakova. Naravno, ovo je samo moguće učiniti za potrebe analize s obzirom na činjenicu da su stvarni sintaktički, semantički i pragmatički aspekti informacijskih procesa međusobno usko povezani: nema stvarnih znakovnih procesa koji bi po svojoj prirodi bili čisto sintaktički, semantički ili pragmatički (Zunde 1981:344). Izučavanje tih aspekata i njihovih međusobnih veza i utjecaja važan je dio izučavanja informatike kao empirijske discipline.

2.5. Unutar informatike mogu se razlikovati tri vida proučavanja informacijskih procesa kojima se, u različitim njenim poljima, daje više ili manje istaknuto mjesto. To su: tehničko-tehnološki, društveni i biološki (Ralston i Meek 1976:657). Na današnjem stupnju razvoja informatika se ponajprije dovodi u usku vezu s tehnikom i tehnologijom. Ta je veza neizbježna jer se operacije na nosiocima informacija obavljaju pomoću brojnih uređaja od kojih je svakako najvažnije elektroničko računalo. Tehnološki aspekt izučavanja postaje u suvremenoj informatici sve izraženiji i širi jer je, osim obrade, prenošenje informacija danas postalo neodvojiv i izvanredno važan dio informacijskih procesa. U ovom aspektu informatika se u velikoj mjeri preklapa s računalnom znanostju i telekomunikacijama, područjima koja se smatraju samostalnim područjima izučavanja. Društveni aspekti informacijskih procesa također čine važno područje proučavanja unutar informatike. S obzirom da informacija tek postoji u odnosu na pojedinca ili skupinu pojedinaca, informacije i informacijski procesi moraju se sagledavati i proučavati sa stajališta njihovih potencijalnih korisnika. Budući da je informatika svojom djelatnošću povezana sa sistemom informiranja na svim područjima privrednog i društvenog života, ona se neizbježno mora baviti djelotvornošću informacija i njihovim prilagodavanjem pojedinim kategorijama korisnika. I, konačno, informatika u novije vrijeme pokazuje živo zanimanje za proučavanje informacijskih procesa u živim organizmima, osobito u čovjeku, jer i takva proučavanja mogu značajno osvijetliti svojstva i prirodu informacije kao općeljudskog fenomena. Takve spoznaje o prirodi informacijskih procesa stvaraju bolju podlogu za oblikovanje umjetnih sistema koji će jednom moći obrađivati informacije slično ljudima.

2.6. Informatika je u ovom času interdisciplinarna znanost. Njeno područje izučavanja preklapa se u većoj ili manjoj mjeri s brojnim drugim područjima i disciplinama, osobito s računalnom znanostju (Ralston i Meek 1976:657). Širina područja kojim se informatika u današnje vrijeme bavi može se potpunije promotriti ako se prelistaju pojedini brojevi veoma korisne godišnje publikacije Annual Review of Information Science and Technology koju izdaje Američko društvo za informacijsku znanost od 1966. godine. Tu se, između ostalog, pregledno obrađuju slijedeća područja: teorijske osnove informatike, povijesni razvoj informatike, informacijske potrebe, projektiranje i izgradnja informacijskih sistema, vrednovanje informacijskih sistema, primjena informacijskih sistema na različitim područjima ljudske djelatnosti, opis i prikazivanje podataka, upravljanje informacijama, stvaranje, održavanje i pretraživanje datoteka, zaštita informacija, baze podataka, informacijski centri i službe, informacijske mreže, automatizacija knjižnica, automatska obrada jezičnih podataka, analiza sadržaja, metode indeksiranja dokumenata, mikrografija i reprografija, tehnike publiciranja, građa računalnih sistema, prenošenje podataka, telekomunikacije, komuniciranje između čovjeka i stroja, umjetna inteligencija, učenje pomoću računala i dr. Ovako širok raspon tema kojim se informatika bavi zapravo je u skladu s ranije spomenutom definicijom informatike (2.2) koja je uobičajena u Sjedinjenim

Američkim Državama. Sudeći prema ovim temama, može se ustvari reći da se informatika bavi svim oblicima informacija, bilo da su one zabilježene na papirnim, magnetskim, optičkim ili drugim medijima. Iako je, preko papirnih nosilaca informacije, informatika u uskoj vezi s područjem dokumentalistike iz kojeg je, u neku ruku, i ponikla (Shera i Cleveland 1977:257), središnje područje njena proučavanja jesu informacije koje se prikupljaju, obrađuju, pohranjuju, dohvaćaju i prenoše elektroničkim putem, područje što ga je poznati sovjetski teoretičar Gluškov (1987:10-13) nazvao bespapirnom informatikom.

3.

3.1. Informatički teoretičari u više navrata ukazuju na činjenicu da je, među ostalim disciplinama, informatika usko povezana s lingvistikom (Taylor 1966:19, Mihajlov i Giljarevskij 1971:17). U manjoj, se mjeri, međutim, govori gdje se ta dva područja izučavnja sreću i u kojem su vidu jedno drugom korisni ili mogu biti od koristi. Sa stajališta lingvistike, a to je slučaj, može se reći, sa svim ostalim znanostima, veza se uspostavlja upotrebom informacijske tehnologije na području proučavanja jezika. Računalo se već podavno koristi u jezikoslovnim istraživanjima, a postoji i posebna grana lingvistike koja se zove računalna ili kompjutorska lingvistika. Upotreba računala dala je značajan doprinos metodologiji leksikografske djelatnosti (Tosh 1969:198-200), stilističkoj analizi te stvaranju i prikupljanju strojno pohranjenih jezičnih korpusa koji služe kao izvor podataka za sustavna i obuhvatna lingvistička istraživanja (Raben i Burton 1981:252-257). U Jugoslaviji, međutim, bez obzira na stanovit napredak u računalnoj lingvistici, još se uvijek odveć mali broj jezikoslovaca uključuje u kompjutorska lingvistička proučavanja, a nosioci takve djelatnosti za sada su pretežno informatičke ustanove (Tancig i Vitas 1985:10).

3.2. Kao disciplina koja se bavi jezičnim oblikom i jezičnim značenjem i lingvistika, sa svoje strane, istraživanjem svojega predmeta neposredno doprinosi informatici, kao, uostalom, i mnogim drugim znanstvenim disciplinama. Budući da je jedno od temeljnih obilježja svake znanosti da ima svoje znanstveno nazivlje, da kle da jezičnim jedinicama sustavno i precizno izražava sistem svojih znanstvenih pojmova, i lingvistika se tim putem izravno povezuje s informatikom jer je proučavanje nazivlja dobrim dijelom i predmet lingvistike, bolje reći leksikologije u širokom smislu (1.4.3). Dakako, veza putem terminologije samo je jedan vid povezanosti tih dviju disciplina. Rezultati jezikoslovnih istraživanja potencijalno su relevantni za informatiku i na drugim područjima (Gilchrist 1972:330), mada je u nekoliko navrata ocijenjeno (Sparck Jones i Kay 1977:183) da je do sedamdesetih godina doprinos lingvistike informatici bio zapravo neznatan jer su se lingvisti bavili pitanjima koja su dosta udaljena od praktične djelatnosti dohvata i preoblikovanja informacija. U današnje vrijeme, međutim, mada rezultati nisu bitno drukčiji u odnosu na ranije stanje, uočljivo je značajno približavanje lingvističkih i informatičkih istraživanja na pojedinim područjima.

3.3. Kao znanost koja se bavi svim vidovima informacije, informatika mora koristiti i produbljivati znanstvene spoznaje o prirodi jezične informacije. Ako ni zbog čega drugo, a ono zbog činjenice što se informatika mora neizbježno baviti pretvorbom jezičnih u nejezične informacije, i obratno. Jer obrada informacija prvenstveno se obavlja zbog povećanja njihove djelotvornosti, a najveći dio informacija prenosi se jezičnim oblicima. Izučavanje jezičnog vida informacije osobito je izraženo na području korištenja informacija i njihove djelotvornosti u širem sklopu komunikacijskih postupaka i procesa. Lingvistička proučavanja na semantičkoj, sintaktičkoj i pragmatičkoj razini otvaraju mogućnost širih uopćavanja o informacijskim procesima kojim se informatika bavi. Tu su izrazito potencijalno relevantna proučavanja unutar graničnih lingvističkih disciplina psiholingvistike i sociolingvistike.

3.4. Razvoj suvremenih informacijskih sredstava i uređaja pokazuje nedvojbena smjer: od strojnog jezika ka prirodnom jeziku. Krajnja je svrha, zapravo, da se umjetni programski jezici eliminiraju kao posrednici između čovjeka i stroja i da se takva komunikacija ostvaruje prirodnim jezikom (Rumjančev 1986:47). Za sada, međutim, još nije zadovoljavajuće riješeno pitanje prepoznavanja ljudskog govora i sinteze govora radi prihvatljivog i razumljivog govornog izlaza iz stroja. Strojno prepoznavanje govora, među ostalim, nužno se mora zasnivati na fonetsko-fonološkom proučavanju svojstava glasova (1.4.1) i njihovog vezivanja u govorne cjeline. Istraživanja akustične i auditivne fonetike ovdje su od velikog značenja, mada se ona, očigledno, moraju usmjeriti na rješavanje ovog specifičnog problema. U tim proučavanjima osobito je važno istražiti pojave koje se očituju u vezanom govoru. Za sada, naime, osobna svojstva govora pojedinog govornika i brzina govora predstavlja za stroj ogroman problem (Poulton 1983:34). Problem je uostalom još i širi i složeniji jer se percipiranje govora ne svodi isključivo na prepoznavanje zvuka (Rumjančev 1986:47). Prepoznavanje govora usko je povezano s "razumijevanjem" govora, područjem kojim se intenzivno bave stručnjaci na području umjetne inteligencije, što otvara niz dodatnih i veoma složenih pitanja. Slični su problemi prisutni i na području sinteze govora. Radi se zapravo o tome da se stvore mogućnosti za izgradnju strojeva koji će bez teškoća moći pretvarati pisani ili digitalizirani tekst u govor, i obratno. U današnje vrijeme govorni izlaz iz stroja ostvaruje se na dva načina: zapisivanjem čitavih riječi i oblika i njihovim spajanjem u rečenice radi reproduciranja, i registriranjem osnovnih glasovnih jedinica, fonema, i njihovim spajanjem u više jezične jedinice. Za sada druga metoda dovodi do zvukovno neprirodnih tvorevina jer se ne postiže prilagodljivost glasova u govornom toku, dok se kod prve radi o dosta ograničenom opsegu jezičnih jedinica. Rad na tom području veoma je aktualan i lingvistika, zajedno s informatikom i akustikom, treba da doprinese ostvarenju tog dragocjenog tehnološkog napretka.

3.5. U vezi sa sintezom i prepoznavanjem govora stoje i istraživanja na području umjetne inteligencije. Iako se to područje smatra dijelom računalne znanosti (Barr i Feigenbaum, eds. 1981:3, Simons 1984:44), ono sigurno isto tako pripada i informatici (Mihajlov i Giljarevskij 1983:8) i nemoguće ga je iz nje izdvojiti. Naime, kao djelatnost koja se bavi "... oblikovanjem inteligentnih računalnih sustava, tj. sustava koji pokazuju značajke što ih povezujemo s inteligencijom u ljudskom ponašanju - razumijevanje jezika, učenje, zaključivanje, rješavanje problema itd.", ona je po prirodu istvari povezana s jezikom koji je zapravo prozor u ljudski um. Stoga je jedan od važnih zadataka i ciljeva na području umjetne inteligencije razvoj računalnih sistema koji će biti u stanju da analiziraju i "razumiju" širok raspon značenja prirodnog jezika, što je nemoguće ostvariti bez podataka koje pruža lingvistika, osobito sintaksa i semantika (Simons 1984:52). Razumijevanje prirodnog jezika također je usko povezano s prikazivanjem znanja, jer se inteligentno vladanje računala, uključujući i razumijevanje jezika, mora zasnivati na pohranjenom znanju. Svako "razumijevanje" jezika pretpostavlja raščlanjivanje jezičnog ulaza, prepoznavanje određenih sintaktičkih jedinica i pridavanje tim jedinicama određenog značenja. Teorijske postavke i postupci generativno-transformacijske gramatike, bez obzira na kritički odnos pojedinih proučavatelja umjetne inteligencije (Winograd 1983:186-189), u značajnoj se mjeri danas primjenjuju u sistemima umjetne inteligencije koji se bave razumijevanjem prirodnog jezika. U vezi s tim nazire se i novi vid proučavanja jezika sa stajališta umjetne inteligencije, a javlja se pod imenom računalna semantika (Charniak i Wilks, eds. 1976:1). Na taj način pitanja koja se javljaju na području umjetne inteligencije već usmjeravaju i lingvistička istraživanja u pravcu davanja odgovora na takva pitanja. Kako na tom području nisu samo bitne lingvističke spoznaje već i kognitivni procesi i ljudske akcije i interakcije, jezična djelatnost mora se sagledavati i proučavati u širokom sklopu ljudskog ponašanja i djelovanja.

3.6. Izrada tekstovnih baza podataka također je područje na kojem rezultati jezikoslovnih istraživanja mogu potencijalno koristiti informatici. Za indeksiranje i pretraživanje takvih baza podataka upotrebljavaju se posebni sustavi naziva koji, u sklopu utvrđenih pravila i postupaka, tvore odgovarajuće jezike za indeksiranje i pretraživanje. Postupak prikupljanja i odabiranja takvih naziva djelomično je i lingvističko pitanje, jer se preferiranje i odabir pojedinog naziva često zasniva na jezikoslovnim kriterijima. Korištenje rezultata lingvističkih istraživanja, osobito onih koje pružaju semantika, morfologija i leksikologija (Gilchrist 1971:18), posebno dolazi do izražaja pri izradi tezaurusa kao strukturiranog rječnika koji pojmovno pokriva određeno područje znanja, a sadržava deskriptore i druge nazive koji upućuju na deskriptore. Uostalom, metodologija odabira naziva koji bi ušli u takve jezike za indeksiranje i pretraživanje polazi od prikupljanja pojedinih riječi i izraza iz stručne literature i postojećih leksikografskih priručnika (Vickery 1975:13-19), pa lingvistička proučavanja jezičnih pojava na području pojedinih znanosti i struka, u okviru jezikoslovnog izučavanja terminološkog leksika, može svakako doprinijeti boljem razumijevanju funkcioniranja terminoloških sistema i jezinica koje ulaze u takve sisteme. To je zapravo područje gdje se ranija dokumentalistika sjedinjuje s informatikom i unutar nje se očituje kao područje obrade znanstveno-stručnih informacija, što se, prema prevladajućem shvaćanju u Sovjetskom Savezu (2.2), i smatra središnjim područjem izučavanja informatike kao znanosti. U vezi s tim značajna su i kvantitativna lingvistička istraživanja sa stajališta pokušaja automatskog indeksiranja i sažimanja dokumenata. Pokušaji automatskog pripisivanja dokumentu skupa deskriptora koji su pokazatelj njegova sadržaja i sredstvo za dohvrat takvog dokumenta zasnivaju se na pretpostavci da je učešće stala pojava određene riječi u dokumentu povezana s vjerojatnošću da je ona značajan pokazatelj njegova sadržaja. Na taj način statistička obrada dokumenta daje popis sadržajnih riječi poredanih po učestalosti javljanja, iz kojeg je, na osnovi utvrđenih pravila, moguće automatski izdvojiti indeksne nazive. Automatsko sažimanje, s druge strane, znatno je složeniji zadatak i njegovo uspješno ostvarenje usko je povezano s razvojem spoznaja na području umjetne inteligencije i razumijevanja govora. Kako je sažimanje postupak izrade skraćenog sadržaja dokumenta, pred stroj se postavlja složen zadatak da "pročita" dokument i da načini njegov sažetak koji je i jezično i sadržajno prihvatljiv. Dosadašnji pokušaji izrade takvih sažetaka (Wylls 1968:136) zasnivaju se na izdvajanju najznačajnijih rečenica iz teksta i njihovu slaganju prema redosljedu pojavljivanja da bi se na taj način stvorio automatski sažetak. Naravno, ovakav postupak ostavlja brojna neriješena pitanja, pa bi lingvistička analiza sažetaka koje su izradili kvalificirani sažimачi, osobito u okviru lingvistike teksta, mogla vjerojatno pružiti relevantne pokazatelje koji bi se dali korisno primijeniti u daljnjim istraživanjima na tom području.

3.7. Još jedno zajedničko područje rada lingvista i informatičara svakako je i područje strojnog prevođenja. To područje, koje je imalo izražene uspone i padove, od početnog oduševljenja i prevelikih očekivanja početkom šezdesetih godina do potpunog razočaranja i prekida skoro svih istraživanja (Arthern 1979: 77-78), danas je opet veoma živa djelatnost, doduše s nešto umjerenijim i realističnijim ciljevima. I mada potpuno automatizirani prijevod ostaje krajnji dugoročni cilj na tom području istraživanja, uz naziv strojno prevođenje (machine translation) danas se sve više rabi i termin prevođenje pomoću računala (computer-aided translation). Tim se nazivom označava da u prijevodnom procesu ljudskom prevodiocu stoje na raspolaganju računalne terminološke baze podataka, kompjutorizirani jednojezični i višejezični rječnici te različiti elektronički uređaji koji po spješuju i racionaliziraju sve faze prijevodnog postupka. Za sada se koristi nekoliko sistema prevođenja uz pomoć računala (Tanke 1979:49) koji daju dosta dobre praktične rezultate, a usko su povezani s terminološkim bazama podataka. U tim

sistemima, osobito u planiranju i oblikovanju terminoloških baza podataka, rezultati lingvističkih proučavanja igraju važnu ulogu, ali je, isto tako, i korist takvih kompjutoriziranih sistema za lingvistiku izvanredno velika. U svakom slučaju, iako je potpuno automatizirano prevođenje još uvijek daleki cilj, jer ono, u krajnjoj liniji, pretpostavlja razumijevanje teksta, proučavanja na području umjetne inteligencije bacaju novo svjetlo na taj problem i otvaraju bolje mogućnosti njegova rješavanja (Wilks 1979:42). Istraživanja na području strojnog prevođenja danas se dobro uklapaju u suvremena stremljenja prema strojnom razumijevanju prirodnih jezika. Ta se djelatnost sada sagledava u širem sklopu istraživanja na području automatizacije lingvističkih procesa jer su problemi zajednički za različite domene, a metode i postupci razvijeni na jednom užem području mogu biti jednako dobro primjenjivi na drugom.

4.

4.1. Do sedamdesetih godina, bez obzira na prirodnu vezu između lingvistike i informatike, jezikoslovna istraživanja dosta su slabo utjecala na informatičko područje. Iako se ni u današnje vrijeme lingvistika i informatika ne prožimlju u onoj mjeri u kojoj bi se to moglo očekivati, noviji pravci izučavanja i na jednoj i na drugoj strani dovode do značajnijih međusobnih približavanja i korištenja postignutih rezultata. Lingvistika sve više širi svoj djelokrug i izvan proučavanja formalnog jezičnog sustava i jezične pojave sve više sagledava u širem sklopu čovjekova ponašanja i djelovanja. Informatika, s druge strane, unutar svojega proučavanja komuni kacijskih procesa uključuje i komuniciranje između čovjeka i stroja putem prirodnog jezika, što je neposredno upućuje da se u tim svojim nastojanjima koristi rezultatima lingvističkih istraživanja i da potiče izučavanje pojedinih aspekata funkcioniranja jezika sa stajališta svojih potreba.

4.2. Sinteza i prepoznavanje govora, umjetna inteligencija, strojno razumijevanje prirodnog jezika i strojno prevođenje široka su područja istraživanja i lingvistike i informatike, ali i niza drugih znanosti. Na ta se područja nadovezuju i tradicionalna područja indeksiranja i sažimanja dokumenata koja, unutar istraživanja na području umjetne inteligencije, otvaraju šire mogućnosti za automatizaciju tih procesa i njihovo obavljanje strojevima. Njima se, isto tako, pridružuju i terminološka izučavanja i, u vezi s tim, stvaranje računalnih terminoloških baza podataka, bilo da se radi o terminologiji same informatike ili bilo koje druge znanosti ili struke.

LITERATURA

- Arthern, Peter J. 1979. Machine Translation and Computerized Terminology Systems: A Translator's Viewpoint. In Barbara M. Snell, ed. Translating and the Computer. Amsterdam: North-Holland Publishing Company, 1979, str.77-108.
- Babić, Stjepan. 1986. Tvorba riječi u hrvatskom književnom jeziku. Zagreb: JAZU i Globus.
- Barr, Avron i Feigenbaum, Edward A., eds. 1981. The Handbook of Artificial Intelligence. Volume I. Los Altos: William Kaufmann, Inc.
- Bauer, Friedrich L. i Goos, Gerhard. 1973. Informatik: Eine einführende Übersicht. Erster Teil. Zweite Auflage. Berlin: Springer-Verlag.
- Brookes, B.C. 1972. Information Science (excluding IR). In H.A. Whately, ed. British Librarianship and Information Science 1966-1970. London: The Library Association, 1972, str. 137-149.

- Bugarski, Ranko. 1985. Lingvistika i njene discipline. Odjek, XXXVIII, 7, 1985, str. 3-4.
- Charniak, Eugene i Wilks, Yorick, eds. 1976. Computational Semantics: An Introduction to Artificial Intelligence and Natural Language Comprehension. Amsterdam: North-Holland Publishing Company.
- Crystal, David. 1985. Linguistics. 2nd ed. Harmondsworth: Penguin Books.
- Ferišak, Vilim, ur. 1981. Osnove informatike. Zagreb: Informator.
- Fromkina, M.J. 1978. Sovremennyj ruski jazyk: leksikologija. Moskva: Vysšaja škola.
- Gilchrist, Alan. 1971. The Thesaurus in Retrieval. London: Aslib.
- Gilchrist, Alan. 1972. Research in Information Science. In H.A. Whatley, ed. British Librarianship and Information Science 1966-1970. London: The Library Association, 1972, str. 320-341.
- Gluškov, V.M. 1987. Osnovy bezbuzmažnoj informatiki. Izd. 2-e. Moskva: Nauka.
- Han, Stjepan i Balaban, Nedo. 1983. Osnove informatike. 2. izd. Beograd: Savremena administracija.
- Harvey, John F. 1967. Professional Aspects of Information Science and Technology. ARIST, 2, 1967, str. 419-444.
- Katičić, Radoslav. 1967. Osnovni pojmovi suvremene lingvističke teorije. Zagreb: Sveučilište u Zagrebu.
- Kržak, Miroslav. 1975. O pojmu i terminu informatika. U knj. IX jugoslavenski simpozij "Automatska obrada podataka". Zagreb: VII Interbio, 1975, str. 9-19.
- Kunin, A.V. 1970. Anglijskaja frazeologija. (teoretički kurs). Moskva: Vysšaja škola.
- Leech, Geoffrey. 1974. Semantics. Harmondsworth: Penguin Books.
- Lyons, John, ed. 1970. New Horizons in Linguistics. Harmondsworth: Penguin Books.
- Mihajlov, A.I. i Giljarevskij, R.S. 1983. O perspektivah razvitija informatiki. NTI, serija 2, br. 5, 1983, str. 1-11.
- Mikhajlov, A.J. i Giljarevskij, R.S. 1971. An Introductory Course on Informatics/Documentation. Revised and enlarged edition. The Hague: International Federation for Documentation
- Palmer, F.R. 1976. Semantics: A New Outline. Cambridge: Cambridge University Press.
- Poulton, A.S. 1983. Microcomputer Speech Synthesis and Recognition. Wilnislow: Sigma Technical Books.
- Raben, Joseph i Burton, Sarah K. 1981. Information Systems and Services in the Arts and Humanities. ARIST, 16, 1981, str. 247-266.
- Ralston, Anthony i Meek, Chester L., eds. 1976. Encyclopedia of Computer Science. New York: Petrocelli/Charter.
- Robins, R.H. 1964. General Linguistics: An Introductory Survey. London: Longmans.

- Rumjancev, M.K. 1986. Estestvennaja i iskusstvennaja reč': jazykoznanie, kibernetika. Voprosy jazykoznanija, 5, 1986, str. 47-54.
- Rzevski, George, 1985. On Criteria for Assessing an Information Theory. The Computer Journal, 28, 3, 1985, str. 200-205.
- Semenjuk, E.P. 1984. Informatika: Kak ee ponimat'? (Analiz složivšejsja situacii). NTI, serija 2, br. 7, 1984, str. 1-9.
- Shera, Jesse H. i Cleveland, Donald B. 1977. History and Foundations of Information Science. ARIST, 12, 1977, str. 249-275.
- Simons, G.L. 1984. Introducing Artificial Intelligence. Manchester: The National Computing Centre Limited.
- Sparck Jones, Karen i Kay, Martin. 1977. Linguistics and Information Science: A Postscript. In Donald E. Walker, Hans Kalgren, Martin Kay, eds. Natural Language in Information Science: Perspectives and Directions for Research. Stockholm: Skriptor, 1977, str. 183-192.
- Stamper, R.K. 1985. Towards a Theory of Information. Information: Mystical Fluid or a Subject for Scientific Inquiry? The Computer Journal, 28, 3, 1985, str. 195-199.
- Tancig, P. i Vitas, D. 1985. O razvoju računarske lingvistike u Jugoslaviji. U knj. Zbornik III. znanstvenega srečanja Računalniška obdelava jezikovnih podatkov. Ljubljana: Institut "Jožef Štefan", Društvo za uporabno jezikoslovje Slovenije, 1985, str. 3-14.
- Tanke, Ebehard H. 1979. Implementing Machine Aids to Translation. In Barbara M. Snell, ed. Translating and the Computer. Amsterdam: North-Holland Publishing Company, 1979, str. 45-69.
- Taylor, Robert S. 1966. Professional Aspects of Information Science and Technology. ARIST, 1, 1966, str. 15-40.
- Težak, Božo. 1971. Informatika - ime za zburjujuči ili razjašnjavajuči koncept. Informatologija Jugoslavica, 3, 1-4, 1971, str. 1-13.
- Tosh, Wayne, 1969. Computer Linguistics. In Archibald A. Hill, ed. Linguistics Today. New York: Basic Books, Inc., 1969, str. 1989-206.
- Tudman, Miroslav. 1986. Teorija informacijske znanosti. Zagreb: Informator.
- Vickery, B.E. 1975. Classification and Indexing in Science. 3rd. ed. London: Butterworths.
- Wilks, Yorick. 1979. Machine Translation and Artificial Intelligence. In Barbara M. Snell, ed. Translating and the Computer. Amsterdam: North-Holland Publishing Company, 1979, str. 27-43.
- Winograd, Terry. 1983. Language as a Cognitive Process. Volume I. Syntax. Reading: Addison-Wesley Publishing Company.
- Willys, Ronald E. 1968. Extracting and Abstracting by Computer. In Harold Borko, ed. Automated Language Processing. New York: John Wiley and Sons, Inc., 1968, str. 127-179.
- Zunde, Pranas. 1981. Information Theory and Information Science. Information Processing & Management, 17, 6, 1981, str. 341-347.

Zunde, Pranas i Gehl, John. 1979. Empirical Foundations of Information Science. ARIST, 14, 1979, str. 67-92.

Primljeno: 1988-09-25

Vojvoda S. Linguistics and information science

S U M M A R Y

The paper deals with some points of contact between linguistics and information science. After a brief survey of the subject-matter and scope of each of these disciplines, some areas of research at the interface between linguistics and information science are discussed. Included here are such topics as abstracting and indexing, thesaurus construction, terminological research and terminological databases, artificial intelligence, automatic speech recognition and synthesis, natural language comprehension by machine and computer-aided translation.