

A.M. LARIONOV és Ju.P. SZELIVANOV /Elektronikus Számítástechnikai Kutatóközpont/ "A számítástechnikai eszközök fejlesztése az automatizált információs rendszer céljaira" c. előadásukban kifejtették, hogy a számítástechnika alkalmazásának alapvető irányzata jelenleg a különböző rendeltetésű adatfeldolgozó rendszerek kialakítása, amelyek hardware-ből, rendszer software-ből és alkalmazási programokból állnak.

Az Egységes Számítógép Rendszer kidolgozása hosszú távon biztosítja a fejlődésben lévő adatfeldolgozó rendszerek, ezen belül az automatizált információs rendszerek egységes hardware-rel történő ellátását.

Ju.A. SREIDER, a VINITI osztályvezetője "A természetes nyelvek mint az információs nyelvek alapjai" c. előadásában az információs nyelvek kidolgozásának elméleti problémáival foglalkozott. Megjegyezte, hogy az informatikai szempontból a szó szoros értelmében vett információs nyelveket az különbözteti meg, hogy az információval egy időben metainformációt - az információ kódolásáról és az adott szövegben vagy file-ban elfoglalt helyéről szóló információt - is képes leírni. A metainformáció fogalmával kapcsolatos az információs nyelveknek dokumentografikus és faktografikus nyelvekre történő felosztása. Az előadó ismertette a tájékoztatási gyakorlatnak azt a két irányát, amelyek az információs nyelveknek a természetes nyelvekhez való közelítéséhez vezetnek.

V.I. GORKOVA, a tájékoztatási szakemberek továbbképző intézetének rektora előadásában rávilágított a tájékoztatási szakemberek képzésének és továbbképzésének általános feladataira.

A konferencia egyetértett a gépi információkeresés, az információs rendszerek software-je, a tájékoztatási szolgáltatások hardware-je és a szakemberképzés területén meglévő hiányosságok felszámolására irányuló intézkedésekről szóló javaslattal.

/Naucsno-Tehnicoszskaja Informacija, Szer.l. 1974. 11.sz.
p.35-37./

Nemzetközi kollokvium számítógépek lexikográfiai alkalmazhatóságáról

A szocialista országok szótárkiadóinak 1970. évi moszkvai konferenciáján a résztvevők megvizsgálták az NDK szótárkiadóinak a műszaki és ágazati szótárak szóanyagának kölcsönös cseréjére és gyűjtésére vonatkozó tárgysorozati ponttal kapcsolatban előterjesztett javaslatát. Ebben felvetették, hogy meg kellene vizsgálni, mennyiben

volnának alkalmazhatók a szaknyelvi szótárszerkesztési munkákban a számítógépek. Kérték a konferenciát, hozzon határozatot a kérdéskomplexum megvizsgálására és számítógépeknek a lexikográfiai munkákba való bevonására. A Moszkvában elfogadott határozat szerint a gépi szótárszerkesztés kérdéseinek megvizsgálására, kellő előkészület után, nemzetközi kollokviumot kell összehívni.

A nemzetközi kollokviumot a drezdai Műszaki Egyetem készítette elő és rendezte meg. A kollokvium 1975. február 4-8. között ült össze, és azon az NDK több egyetemének, főiskolájának, kutatóintézetének és szótárkiadásának képviselőin kívül érkeztek küldöttek más szocialista országokból is /Bulgária, Csehszlovákia, Lengyelország, Magyarország, Szovjetunió/.

A nemzetközi kollokviumot a Műszaki Egyetem Alkalmazott nyelvészeti tanszéke az egyetem Számítástechnikai Központjával karöltve rendezte meg. A kollokviumon számos előadás hangzott el, a központjában a tanszék munkacsoportja által kidolgozott gépi szótárszerkesztő program ismertetése állt. A vita alá bocsátott nagy témakört /számítógép segítségével készíthető két- vagy több nyelvű szakszótár/ három nagy témára bontották, és pedig:

- a gépi szaknyelvi lexikográfia általános alapjai;
- szakszógyűjtemények gépi összeállítása;
- szakszógyűjtemények gépi kidolgozása.

E három téma nagyjából jól kiegyenlített nagyságrendű és mélységű előadássorozat keretében került megvizsgálásra, majd megvitatásra.

Az elhangzott előadások közül mint különösebb érdeklődésre méltót, ki kell emelni a következőket:

E. MAIER professzor /ilmenau Műszaki Főiskola/: Pszeudotövek képzése számítógépes feldolgozás céljaira; L. HOFFMANN professzor /Lipce/: Tudományos és szaknyelvi szótárak számítógépes szövegelemzéssel nyert gyakorisági szótárak; L. DROZD, docens /Tudományos Akadémia, Prága/: Rendezési rendszerek a terminológiában; A. LJUDSZKANOV, kandidátus /Szófia/: Néhány nyelvi, információelméleti és szervezési probléma terminológiai adatbankok készítése és felhasználása terén; NEUBERT, docens /Drezda, nyelvi tanszék/: A szaknyelvek elektronikus szótára - lexikográfiai strukturája és felhasználási lehetőségei; W. HORNUNG, docens /Humboldt Egyetem, Berlin/: A német-orsz szakszókincs ekvivalencia típusai lexikális-szemantikai sikon tezaurusz-készítésre tekintettel; E. MAYER, mérnök /Drezda/: A szakszavak elektronikus szótárának felépítéséből és használatából eredő problémák; A. TEJNOR /Prága, Tudományos Akadémia/: Szakszókincs lyukkártyás feldolgozásának problematikája; H. SPITZBARDT, professor /Jéna/: Az indonéz nyelv szótöveinek automatikus elemzése; W. BORSODORF /Drezda/: Közelítő módszerek a lexikográfiai szóanyagszervezés szintjének meghatározására; Ch. VOIGT /Drezda/: Számítógépi segédlet egy három nyelvű szakszótár kidolgozásához; W. REINHARDT, docens /Drezda/: A szaknyelvek elektronikus szótára lexikológiai kutatásokra való felhasználásának lehetőségeiről és problémáiról; W. KLENGEL /Drezda, Gra-

phischer Grossbetrieb Völkerfreundschaft/: Számítógépes vezérlésű fényszedés tapasztalatai két nyelvű szakszótárak előállításában.

A kollokvium résztvevőinek megállapításai és a vita eredményei a következők voltak:

Kétségtelen, hogy a gépi lexikográfia korszerű és gyors feldolgozást tesz lehetővé. Így nemcsak a feldolgozott irodalom volumenje nő meg rendkívüli mértékben és nemcsak a feldolgozás válhat sokkal szisztematikusabbá, hanem az ily módon gyűjtött anyag közzététele is sokkal gyorsabb, mint a hagyományos, kézi cédulázással készült szótáraké. A gépi lexikográfia segítségével lehetséges lenne egy nagyobb műszaki szótár 3-5 éves gyűjtési, szerkesztési és nyomdai átfutási idejét lényegesen lecsökkenteni.

A gépi feldolgozás előnye még az, hogy a gyűjtött anyagot nem kell ismételtelen leírni és rendezni kiindulási nyelv szerint, hanem a szavak kellő indexelésével a számítógép a két, illetve több nyelvű kombinációt tetszés szerint készíti el.

Kétségtelen azonban, hogy a számítógépes lexikográfia nyelvészeti problémái még koránt sem teljesen megoldottak. A gépi feldolgozás igen átfogó intellektuális emberi munkát feltételez: kísértő információk azonosító kulcsainak kidolgozását és kipróbálását, a szókincs leírását /ami egyet jelent a kísértő információk számának nagy megnövekedésével/. Modern számítógépek természetesen nagyobb mértékben képesek "intellektuális" műveleteket átvenni, de a programok és algoritmusok kidolgozása még bonyolult feladat - elsősorban számos elem és kapcsolat többértelműsége /poliszemiája/ stb. következtében. Megoldatlan még a kiszótárázó, elemző gépi program problémája, amelyvel természetes írott nyelvű szövegeket használható módon lehetne feldolgozni. A gépi lexikográfia nem képes a szövegekben előforduló összes jelsorozatot megragadni, így különösen hiányzanak pl. teljesen vagy részben a ragozott alakok, helyesírási variánsok, derivált és összetett alakok, neologizmusok, idegen szavak, rövidítések és betűszók, valamint a nevek.

Az ember, intellektuális adottságai folytán és különösen emlékező képessége, analógiai következtetései, általános társadalmi és különleges szakismeretei révén, képes szinte minden, fent említett jelsorozatot megérteni, még akkor is, ha az még szótárázva nem lett.

Az olyan nem intellektuális befogadó, mint az elektronikus adatfeldolgozó, a jelsorozatok azonosításához nem nélkülözhet sem egy szótárt, sem egy algoritmust /programot/ ahhoz, hogy a szövegben előforduló jelsorozatokat helyesen rendelje a szótárba. E problémák megvitatására ez év őszén Ilmenauban konferenciát fognak rendezni.

Kidolgozásra vár még egy több nyelvű terminológiai rendszer is.

Javaslat hangozott el arról is, hogy a további tudományos kutatások központja Drezda legyen, mivel itt megfelelő számú és már gyakorlattal bíró szakember van jelen /Alkalmazott Nyelvtudományi tanszék és Számítógépközpont/, valamint egy nagy kapacitású gép /BESZM/. Ezt

a javaslatot a résztvevők egyhangulag elfogadták, s felvetődött az is, hogy Drezdát e kérdések nemzetközi kutatása központjává kellene tenni.

Bár Magyarországon gépi szövegfeldolgozási munka folyik /MTA Nyelvtudományi Intézet, Kossuth Lajos Tudományegyetem Orosz filológiai tanszék, Budapesti Műszaki Egyetem Nyelvi Intézete/, számítógép két és több nyelvű lexikográfiai felhasználására még tudomásom szerint nem került sor. Kíváncos lenne, éppen a baráti országokban folyó ez irányú kutatásokra tekintettel, a számítógépes lexikográfiára irányuló kutatások számára Magyarországon is erre alkalmas intézményt kijelölni, miután előbb vagy utóbb e kérdés nálunk is aktuális lesz.

Végh Béla

TERMINOLÓGIA

Terminológiai adatbank és a fordítás

Az 1950-es évek végén a számítógépek alkalmazásával előreláthatóvá váltak a nem numerikus információk feldolgozásának nagy lehetőségei; a nyelvészeti felhasználás, különösen pedig a gépi fordítás tekintetében kissé korai lelkesedés volt észlelhető.

A gépi fordítással kapcsolatos nehézségek ellenére voltak próbálkozások a pozitív eredmények felhasználására. 1965-ben az Európai Községek Bizottságának Fordítói Csoportja kifejezés-szótár kidolgozását kezdte meg, mely a gépi fordítás módszereiből kölcsönzött morfológiai analízis sémájára alapult. Egy az NSZK-ben 1966-ban megjelent jelentés tanúsága szerint a Nyelvi Szövetségi Hivatal szóbank kialakítását tűzte ki célul. M.E.O. LIPPMANN, számítógép segítségével történő fordítási technikát dolgozott ki. LIPPMAN CAT /Computer Aided Translation/ rendszerében a fordítás mechanikus műveleteit a gép végzi, míg a fordító az analízis és szintézis műveleteinek irányítója marad. E rendszer a terminológia problémáját nyitott kérdésként meghagyja, és a fordító hatáskörébe utalja.

Eközben a terminológiai problémák egyre kényesebbé és megoldhatatlanabbá válnak, különösen a tudományos és műszaki fordítás területén.

1965-ben Kanadában kollokviumot tartottak a fordítás területén felmerülő terminológiai problémák tárgyában és javasolták egy gépesített terminológiai nyilvántartás elkészítését. 1970-ben létrejött a Montreali Egyetem terminológiai adatbankja.