

## A tartalom formája és a változás logikája

### A tartalmi feltárás néhány elméleti kérdése

*A könyvtári feldolgozás automatizálása olyan lehetőségeket nyitott meg a formai és tartalmi feltárásban, amelyek hagyományos keretek között megvalósíthatatlanok. Ezeknek a lehetőségeknek a kiaknázásához azonban sokkal következetesebben kell elemezni többek között a tartalmi feltárás műveleteit. Az elemzés vezérfonala az a felismerés, hogy ezek a műveletek egymással párhuzamosan különböző (fizikai, felhasználói, nyelvi, fogalmi) szinten játszódnak le, és e szintek között kölcsönhatások állnak fenn. A tartalmi feltárás problémáit szintenként vesszük szemügyre.*

Az évszázad utolsó évében a legnagyobb egyetemi, köz- és szakkönyvtárakban, de zömmel a városi közkönyvtárakban is különböző típusú és színvonalú integrált könyvtári rendszerek automatizálják azokat a munkafolyamatokat, amelyeket korábban gépek nélkül, vagy sokkal egyszerűbb programokkal végeztek el.

A szakmai zsargonban csak „gépesítésnek” nevezett folyamat olyan ütemben robbant a könyvtárak egyébként inkább „konzervatív üzemű” életébe, amellyel nem minden munkafolyamat volt képes lépést tartani. Bizonyos területeken kellően előkészültek a gépi feldolgozásra – pl. a formai feltárás szabványok általi szabályozása –, más területeken kevésbé. A tartalmi feltárás ilyen előkészítetlen terület maradt. Már a 80-as években látni lehetett, hogy az ETO – mint 100 éves osztályozási rendszer – sokat emlegetett hibái miatt már nehezen képes tükrözni az időközben a tudományok szerkezetében bekövetkezett és azzal párhuzamosan jelentkező fogalmi változásokat. Az ETO alapvető problémája, hogy mesterséges nyelven alapul, a dokumentum, illetve a keresőkérdés természetes nyelven kifejezett tartalmát tehát erre a mesterséges nyelvre kell lefordítani, ami elkerülhetetlenül nehezíti az osztályozási rendszer laikus használójának dolgát. Ez a nehézség az automatizált rendszerekben viszonylag könnyen feloldható azáltal, hogy a felhasználó és az ETO közé természetes nyelvű „felületet” iktatnak be. A természetes nyelv információkereső nyelvi használata viszont más, újabb nehézségeket vet föl.

Milyen elméleti problémákkal kell szembenéznünk azoknak, akik a tartalmi feltáráshoz a természetes nyelven alapuló, deskriptoros/tárgyszavas információkereső nyelvet választják? Ezeknek a kérdéseknek az elemzéséhez abból kell kiindulni,

hogy a problémák különböző szinteken jelennek meg. Természetes nyelven alapuló információkereső nyelvek esetében is lejátszódik fordítás, mégpedig a beszélt nyelv és az információkereső nyelv között. A fordítás tényét azonban elrejt, hogy mindkét nyelv (a beszélt és az információkereső nyelv) esetében természetes nyelvekről van szó. Ha a kérdést a nyelvi rendszerek különféle szintjein vizsgáljuk, ez a fordítási folyamat jobban megragadható. A továbbiakban az egyszerűség kedvéért a természetes nyelven alapuló kötött szótáras információkereső nyelvek használatát indexelésnek fogjuk nevezni.

#### Fizikai szint

A legelső, *fizikai szinten* a dokumentumban testet öltött mű(alkotás) az indexelés kiindulási pontja. Ha a formai feltárás analógiájához hűek akarunk maradni a tartalmi feltárásban (miért is ne tennénk, hiszen lényegileg itt is másodlagos azonosítók segítségével végrehajtott leírás készül), akkor a „tükrözés” módszerét alkalmazhatjuk. Ez annyit jelent, hogy a kézbe vett dokumentum tárgyát megpróbáljuk tárgyszavakba sűrítve vissza-tükrözni. A formai és tartalmi feltárás módszerének nemcsak azért kell összhangban lennie, mert azonosítási műveleteik egy töről fakadnak, hanem azért is, mert több ponton egymásra támaszkodnak és támaszkodhatnak. A formai feltárásban általában legalább annyi tartalomra utaló jelzés lehet, mint fordítva. Gondoljunk az alcím körülíró funkciójának tartalmi elemeire, a sokatmondó műfaji utalásokra, amelyek jelzik a tartalom formáját.

A tartalmi feltárásban gondolhatunk az ETO esetén a formai alosztásokra, tárgyszavazásban a

dokumentum típusát megnevező formai tárgyszavakra. Ezek az elemek valójában átmenetet képeznek a formai és tartalmi feltárás között. A gépi feldolgozásra való áttéréssel még felismerhetőbbé vált az a sajátosság, hogy a formai és tartalmi feltárás ugyanúgy nem határolható el egymástól, mint ahogyan általában a tartalom és a forma nem választható szét teljesen, sőt minden eddigénél inkább konvergálnak egymáshoz.

A hagyományos kézi feltárás során a formai feltárásban léteztek, de háttérbe szorultak a tartalomra utaló elemek (pl. az alcímek tartalmi körülíró funkciói), ugyanígy a tartalmi feltárásban sem voltak olyan rugalmasan visszakereshetőek a tartalomra utaló formai elemek. Az ETO erői és hibái között egyaránt említhetjük azt az átfedést, ami bizonyos számok főtáblázati és önálló, általánosan közös alosztásokkal való párhuzamos használatát lehetővé teszi [pl. a szótárak esetében a  $801.316.4:33=20=945.11$  és  $33(038)=20=945.11$ , vagy a  $030/050/06$  főtáblázati számok formai alosztásként való használati lehetőségei]. Az általánosan közös alosztások kiemelhetősége viszont [pl. (031) vagy (086.7)] éppen a formai visszakeresésre adott esélyt, bár nem jellemző általában, hogy a közkönyvtárak gyakorta éltek volna ezekkel a lehetőségekkel. Az általános gyakorlat inkább azt tükrözte, hogy a dokumentum típusára utaló alosztások a keresés szempontjából másodlagos vagy harmadlagos elemeknek számítottak. Ez a gépi feltárással teljességgel megváltozott, hiszen minden jelzetelem (főtáblázati szám és általánosan közös alosztás) ugyanolyan könnyen felhasználható az automatizált kereséskor.

A címek ugyan a formai feltárás tárgyai, de a címszavakra végzett keresés tartalmi feltárásnak tekinthető. A tartalmi feltárás elemeinek egy része viszont formai jellegű (pl. a dokumentum típusa, vagy a személy neve, akiről a dokumentum szól). Ez a sajátosság (konvergencia) a kézi feltárásban elsikkadt, jelentőség nélkül maradt.

A tartalmi feltáráskor a könyvtárosnak természetesen nincs arra lehetősége, hogy a dokumentumokat sorról sorra elolvassa, de számos önfeltáró segédeszköz áll rendelkezésére a dokumentumban munkája könnyítéséhez (tartalomjegyzék, mutatók, előszó, utószó, fűszöveg, sőt, megint visszakapcsolódva a formai feltáráshoz: kiadó, megjegyzések, sorozat-adatelemek és -csoportok előfordulási értékei).

Ezen a fizikai szinten, amikor a dokumentum (maradva csupán a könyvnél) tárgyról beszélünk, a szakirodalom többször visszatért arra a művelleti ellentmondásra, amely szerint egyrészt a könyv tárgyának azonosítása fizikai egyedhez kapcsolódik (szelektív azonosítás), másrészt az adott tárgy azonosítása esetében ugyanaz a könyv már egy

csoport tagja (átfogó áttekintés). Eredetileg Cutter fogalmazta meg ebben a kettősségben a tárgyszó-katalógus célját, amelyben Dunkin ellentmondást látott [1]. Az ellentmondás látszólagosságát könnyen beláthatjuk, ha pontosabban megvizsgáljuk, mi a helyzet a fogalmi szinten. A könyv tárgyát az első esetben valóban a tárgyszó, illetve deskriptor azonosítja. Viszont az adott tárgy irodalmának nem a könyv, hanem a tárgyszó az azonosítója, a könyv pusztán annak előfordulási értéke, és ilyen értelemben valóban egy csoport tagja.

## Felhasználói szint

A fizikai szint mögött, háttérként dominál a *felhasználás szintje*. Amikor a tartalmi feltáró könyvtáros kézbe veszi a feltárandó művet, elsődlegesen magából a dokumentumból és annak feltárási szintjéből indul ki. Tudatában azonban akaratlanul is jelen van annak a felhasználónak a személye, akivel munkája során közvetett kapcsolatba kerül, akinek dolgozik. Bizonyos könyvtár típusoknál a felhasználó személye konkrét rétegekhez kötődik, ami specifikusan orientálja a könyvtáros feltáró munkáját. Nehezebb helyzetben vannak ebből a szempontból a nagy közkönyvtárak, ahol a gyűjtemény általános körű, az olvasók pedig sokféle réteget képviselnek. A feltárónak abból a jóhiszemű feltételezésből kell kiindulnia, hogy a megfelelő feltárási szinteket a hasonló mélységben érdeklődő felhasználók megtalálják. Az az olvasó, aki nem ismeri a tribológia szakkifejezést, nyilván nem is fogja keresni, aki viszont ismeri, feltehetően így keresi. Más szóval: a tartalmi feltárás függ a felhasználói körtől is.

## Nyelvi szint

A felhasználási szint után átlépünk egy bonyolult, sokrétű problémakörbe, amely a tartalmi feltárás elméleti kérdéseit a legmélyebben érinti, ez pedig a *nyelvi szint*.

## Terminológia

A problémakör egyik területe a *terminológia* kérdésköre. A hazai könyvtári szakirodalomban ez már a múlt század végén (Kudoránál) és a század első felében komoly problémát jelentett, amit Kálmány Géza a következőképpen fogalmazott meg 1943-ban: „... a szófogalmak [ma szakszerűen úgy mondjuk: az általános nevek] s a szavak közértelmű jelentései az idők folyamán, mint tapasztaljuk, elég gyorsan változnak, egyes szók

## Szinkronitás és diakronitás

A tárgyszavak szinkronitási és diakronitási problematikája is a terminológia ellentmondásai közé tartozik. Általános gyűjtőkörű nagy könyvtárakban a kettő párhuzamosan előfordulhat.

A természetes nyelvi változásokban az avulás folyamata állandóan jelen lévő probléma. Történeti tárgyú dokumentumok esetén nem kerülhető el az elavult szakkifejezések használata (pl. a lyukkártyákról szóló tanulmány esetén a már sehol sem használatos régi adathordozók neve). Az iskolai értesítők mint dokumentumtípusok is elavultak, illetve most nem ez a nevük, de a róluk készült bibliográfiát legpontosabban mégiscsak ezzel a történetivé vált kifejezéssel jellemezhetjük tartalmilag. A retrospektív tartalmi feldolgozások ezek a problémák halmozottan fordulnak elő.

Az avulás együtt jár a keletkezéssel. Ha valami egészen új dolog, arra egyszerre több kifejezés lehet egyaránt érvényes (pl. noteszgép, notebook, laptop, thinkpad). A szinkronitás és történetiség kérdése a tartalmi feltáráskor szorosan összefügg a szemantikai kérdésekkel, melyekről később lesz szó.

A terminológia szintjén a legkomolyabb problémát az írásbeliség és szóbeliség közti szakadék okozza, amelyre napjainkban hívta fel a nyelvelmélet és filozófia a figyelmet.

Amikor indexelünk, valójában írott anyagot az információkereső nyelv szavaira fordítunk, és írásban rögzítünk. A felhasználó pedig, amikor írott anyagot keres, először a saját, majd az információkereső nyelv szavaival próbálja megközelíteni az írásban rögzített tárgyszavakat. Ezért az indexelés művelete olyan, mintha szakadékon át folyton-folyvást oda-vissza ugrálnánk (osztályozó és felhasználó egyidejűleg).

Ez a fordítási, transzformációs folyamat a fogalmi és a nyelvi, majd a nyelveken belüli szintek között játszódik le.

## Szemantika

A nyelvi szint egyik alszintje a szemantikai szint. Ezen a szinten az indexelést meghatározóan nehezítik a következő elemek: szinonimák, homonimák, az átmenetek, melyeket nevezhetünk kettős alakzatoknak, valamint azok a deformálható alakzatok, amelyekre valójában még kifejezés sem létezik.

A szinonímia kérdésköre a lexikográfia egyik ugyancsak problematikus ága. A 20. századra nem a szinonimák létezésének kérdése a probléma tárgya, hanem „mint Quine és Carnap vitája tanúsítja – a szinonima fogalom meghatározhatósága” [4].

avultakká válnak (ki beszél ma élenyről, légenyről, légszeszről? stb.)” [2].

Századunk egyik legdinamikusabb nyelvfejlődési tendenciája a szakszókincs – minden szakterületet érintő, de bizonyos területeken robbanásszerűen bekövetkező – folyamatos gyarapodása. „A szakszókincs mennyiségi növekedése és differenciálódása éppen a legújabb idők, napjaink nyelvi fejlődésének egyik legfontosabb, legjellemzőbb sajátossága” [3]. Tudományelméleti szempontból ez a tendencia egyrészt az azonos terminust más jelentésben használó elméletek összemérhetetlenségét, másrészt tágabb értelemben a tudományágak terminológiai globalizálódását eredményezi. Utóbbi annyit jelent, hogy egy-egy paradigmán belül jellemző az általános tudományos terminológia horizontális, egyre több tudományágba hatoló szétterjedése. Egyértelmű, hogy ez a folyamat egy adott paradigmán belül a gondolkodásmód hasonlóságának függvénye.

Egy szakterület alaposabb terminológiai vizsgálatai bizonyíthatják, hogy a terminológiai változásokon belül hosszabb és rövidebb távú tendenciák egyaránt jellemzők. A könyvtartudományban ilyen hosszabb távú (kb. fél évszázados) változás volt a dokumentáció és információ szakkifejezések terminológiai váltása. Számtalan közép- és rövid távú (maximum 10–20 év, minimum 1–2 év alatt lezajló) terminológiai változást a technikai fejlődés eredményezett (pl. a lyukkártyák különböző típusú kifejezéseinek pályafutása).

A terminológia megválasztásában tartózkodnunk kell a tudományos vagy népszerű terminológia közötti ellentmondás csapdájától. A művet a feltárási szintje határozza meg elsősorban. Tudományos dokumentumnak nem adhatunk népszerűsítő tárgyszót, mint ahogy fordítva sem. Nem szabad elvonatkoztatnunk, vagyis egy fogalmi szinttel feljebb lépni, amire Cutter legismertebb szabálya – miszerint mindig a legspecifikusabb tárgyszóval kell osztályozni – is fölhívja a figyelmet. Nem tévesztheti meg a tartalmi feltárót az egyedi tárgyszóalkotásban az sem, hogy az általános is lehet egyedi (átfogó téma esetében, ha például a dokumentum tárgya maga a Tudomány vagy az Élet).

Át kell hidalnunk terminológiai szempontból azt az ellentmondást is, amely a szövegelemzés szempontjai és a terminológia között áll: vagyis azt az elvet, hogy az újdonságot a többnyire egyed- és nyelvspecifikus szakkifejezések hordozzák. A tudományos szaknyelvből számtalan példa hozható arra, hogy az eredetileg egyedi szakkifejezés hosszú távon fennmarad, többnyire alkotója nevéhez kötődően a nyelvben (pl. a „fekete lyuk” és John Wheeler, a „másodlagos szóbeliség” és Walter J. Ong).

Nyelvészeti szempontból még a tökéletesnek látszó szinonimák is rejtenek magukban olyan ellentmondást, amely megkérdőjelezi fogalmi azonosságukat. A szinonimák sajátossága a dominanciára való törekvés. Ebből a szempontból a szinonimák egymás kiszorítására irányuló versenytársaknak tekinthetők. Fejlődésük során éppen ezért válhatnak belőlük időnként antonimák (*ellentétpárok*, pl. áld – átkoz). Ha a kiszorítójáték döntetlenül végződik, a szinonimák *rokon értelmű* szavak maradnak. Igen ám, de ez nem jelenti feltétlenül azt, hogy nyelvészetileg bármilyen esetben helyettesíthetik egymást – ami logikailag a tökéletes szinonimitás kritériuma lenne. Klasszikus példája lehet ennek a kutya – eb szinonimapár. A köznyelvben elterjedt kutya alak nem helyettesítheti a hivatalos kifejezésekben (pl. testületi nevekben) és frazeológiákban megőrződött *eb*-et (nem cserélhetünk kutyán gubát), mint ahogy az *eb* sem helyettesítheti a kutyát bármilyen esetben. Az ehhez hasonló nyelvészeti problémákat utalással lehet feloldani.

Pszudoszinonimáknak vagy „kváziszinonimáknak” nevezhetők osztályozási szempontból a jelentésárnyalaton, túl specifikus nem-faj vagy egyéb, főként laza logikai kapcsolaton alapuló összefüggések [5].

Osztályozási szempontból megkülönböztetünk alaki és jelentéstani szinonimát. Az alaki szinonimák kérdésköréhez kapcsolódik részben az idegen vagy magyar nyelvű alak használata (pl. dyslexia és diszlexia), amelyeknél a kevésbé elterjedt alakról utalunk az elterjedtebbre, valamint a mozaikszavak kérdése (mivel a mozaikszó homonimia is lehet, ezért a rövid alakról utalunk a teljesre, pl. MNB lehet Magyar Nemzeti Bibliográfia és Magyar Nemzeti Bank is).

A jelentéstani szinonimáknál is szerepel az idegen vagy magyar nyelvű (tudományos és köznyelvi) alak alternatív használati lehetősége (pl. lélektan vagy pszichológia, a különböző összetételek ellentmondásaival együtt, pl. gyermeklélektan, de szociálpszichológia), amelyet ugyancsak a gyakoribb előfordulás és ugyanakkor az indexelő nyelvi ízlése határoz meg. Ilyen következetes nyelvi ízlés lehet az összetett szavak etimológiai homogenitása, vagyis az arra való törekvés, hogy összetett szavak esetén azonos eredetű, csak idegen nyelvű vagy csak magyar nyelvű elemeket válasszunk, követve a nyelvi előfordulás gyakorisági szabályait (pl. a nehézkes gyermekpszichológia helyett a gyermeklélektan, ám a kevésbé gyakori társadalomlélektan helyett az elterjedtebb szociálpszichológiát válasszunk, és engedjünk a nyelvi kompromisszumnak, ha szükséges). Ehhez persze nem csupán nyelvi ízlés, hanem érzék is szükséges.

Azokat az idegen szavakat, amelyek jelölete a magyar kultúrában – egyelőre – ismeretlen, magyarisítás helyett eredeti alakjukban logikusabb átvenni (origami, aikido stb.). Az egyéb jelentéstani szinonimák kérdését egyértelműen következetes utalással egységesíthetjük.

A terminológiai újdonságokból eredő, „halmozódó” szinonimákat (pl. noteszgép, notebook, laptop vagy honlap, otthallap, weblap) egy, lehetőleg a gyakoriság alapján választott alakban egységesítjük.

A *homonimákat* értelmező teszi egyértelművé. Az egyszerű jelentéstani (szótári) homonimák kérdése (azonos alak, más jelentés) nem okoz gondot a tárgyszavazásban. Nagyobb problémát jelentenek a homonimák akkor, ha ugyanaz a kifejezés számos tudományterületen előfordul: ilyen pl. az analízis. Ha a dokumentumot ilyen összefoglaló kifejezéssel indexelik, az nyilván teljes átfogó feltárási szintet jelent. Ha viszont a dokumentumot részterületi finomsággal kell feltárni, az indexelőnek el kell döntenie, hogyan specifikálja a tartalmat: analízis (matematikai) vagy matematikai analízis. Az értelmezős homonimával szemben természetesebb a jelzős szerkezet használata, de nem mindig célravezető (pl. mondhatjuk, hogy villanykörte, de azt nem, hogy gyümölcskörte. Ezért ilyenkor elkerülhetetlen a hátravetett értelmező: körte [gyümölcs]). Ezeket a kétféleképpen feloldható homonimákat tekinthetjük kettős alakzatoknak. A nyelvtani homonímia kérdéskörének (egy lexéma és egy toldalékolt alak, vagy két toldalékolt alak közötti azonosalakúság) a számítógépes nyelvészetben van megkülönböztető jelentősége.

Még nehezebb feladattal találkozunk az indexelő abban az esetben, ha nincs olyan elfogadott kifejezés, amellyel jelölhetné a dokumentum tartalmát. (Az ilyen esetek jobb híján nevezhetők deformábilis alakzatoknak.) Ez a magyar nyelvben olyankor fordulhat elő, amikor a gyűjtőfogalom hiányzik. (Pl. az erdei és hegyi sportoknak nincs olyan átfogó gyűjtőfogalmuk, amely mindkét terület valamennyi sportágát [földi és légi] magába foglalná, vagyis a nyelvi rendszerben itt hiányzik egy láncszem, amelyet fogalmilag mégis megragadunk.) „Az egyes nyelvek eltérnek a fölérendelt kifejezéseik és az ilyen kifejezések alá sorolt hiponimák tekintetében. A klasszikus görögben például az »ács«, az »orvos« és a »fuvalás« lexémák és egyéb más foglalkozások mind a *demiourgos* hiponimái; az angolban és a magyarban azonban nincs egy ennek megfelelő fölérendelt kifejezés” [6]. A nyelvi szint tehát a szemantikai szinten keresztül különösképpen kapcsolódik a fogalmi szinthez.

## Morfológia

A nyelvi szint nem kevésbé fontos alszintje a *morfológiai*, amely formálisan tartja kordában az indexelőt és az információkereső nyelvet.

Alapvető elemei a szófajhatárok kijelölése (gondoljunk a főnevesítés szabályára), az esetek tiltása, a többes szám szabályozása, az egybe- és különírás használata. Ellentmondásai elsősorban a szóösszetételek esetén jelentkezhetnek. A magyar nyelvben a 80-as évek végéig megfigyelhető volt az általános szóösszevonási tendencia, amelynek a helyesírási szabályai ma is érvényesek. A 90-es években azonban a különírás tendenciája került előtérbe, ami a köznyelvi használatban szóösszetétel helyett a szókapcsolatos szerkezeteknek kedvez inkább.

A tárgyszavak helyesírásában nem függetleníthetjük magunkat a magyar helyesírás szabályaitól, és az érvényes könyvtári szabványoktól. Ismét vissza kell utalni arra a bizonyos konvergenciára, amely a gépi kezelés következtében idősebbé vált a formai és tartalmi feltárásban. Mivel azt az elvet követjük, hogy a kettő egy töről fakadó művelet (dokumentumazonosítás két szempontból), ezért szinkronban kell lenniük. Ha a formai feltárás során a „róla szóló” tárgyi melléktételekben a szabványos személynév, testületi név, földrajzi név vagy cím alakokat használjuk, akkor úgy logikus, ha a tartalmi feltárás hasonló eseteiben ugyanígy járunk el, vagyis a köznyelvi tulajdonnév-változatokról utalunk az egységesített alakokra.

Itt egy ponton van komoly ellentmondás, ez pedig a folyóiratok, napilapok címének írása, melyben a két „szabvány” (A magyar helyesírás szabályai, 196–197. és MSZ 3424/2) eltér egymástól. Egy másik hiátus a márkanevek írása, amiben csak a magyar helyesírás szabályaira támaszkodhatunk.

A tulajdonnevek helyesírására vonatkozó nagybetűs szabály az említett kivétel nélkül mindenhol érvényesül. A könyvtári szabványok által egységesített tulajdonneveket a köznyelvi névváltozatokról egységesítő vagy magyarázatos utalók közlésével kell feloldani.

Ha külön-külön indexekben egységesen kezeljük a formai és tartalmi feltárás során alkalmazott tulajdonneveket, valamint a tárgyszavakat, akkor a köznévi tárgyszavak egységes állományt alkothatnak.

A tárgyszavak helyesírása követi a magyar helyesírási szabályzatot, tehát a köznévi tárgyszavak kisbetűvel kezdődnek; ez a gyakorlat ma mégsem általánosan jellemző Magyarországon.

Összegezve megállapítható, hogy a nyelvi struktúra valamennyi szintje meghatározó kapcsolatban áll az indexeléssel.

## Fogalmi szint

A nyelvi szint nem tárgyalható a *fogalmi szint* nélkül. A modern nyelvelméletben és filozófiában az egyik kulcskérdés, hogy „... vajon a valóság fogalmi tagolását nyelvünk határozza-e meg, vagy éppen fordítva, a nyelvi rendszer megszületését megelőző gondolati fejlődés determinálja-e az alapvető szemantikai és szintaktikai struktúrák kialakulását?” [7]

Több tudományterület átfogó kutatásai igazolják, hogy „a megismerés határozza meg a nyelvet” [8], vagyis a nyelvi „címkezés” folyamat feltételez előzetes gondolkodási, fogalomalkotási, kategorizálási folyamatokat. Ez nem kisebbíti Chomsky nyelvi kompetencia fogalmának jelentőségét, sőt tágabb értelemben kiterjeszthetővé teszi azt a nyelv evolúciós eredetének gondolatára [9]. A kategorizálás műveletének alacsony szintű képességét állatkísérletekkel is igazolták. Vagyis kategorizálási művelet nyelvi közvetítés nélkül is létezhet. A bonyolultabb fogalomalkotás és nyelvi kompetencia azonban csak az ember sajátossága. A fogalomalkotás képessége kölcsönhatásban áll a nyelvi képességgel. A fogalmi szint azonban nem azonosítható a nyelvi szint szemantika síkjával. A nyelvi szinten ugyanis a különböző síkok is kölcsönhatásban állnak. Ezt jelzi az az egyszerű tény is, hogy a szemantika és a szintaxis kölcsönösen hat egymásra. A jelentés kontextusfüggő, és ugyanakkor, egyidejűleg a szöveg a jelentés függvényében válik valamilyen szintaxis-szerkezetté.

Az indexelés fogalmi szintje több szinten kapcsolódik a nyelvihez. Utaltunk a szemantikai síkon a nyelvi rendszer lehetséges betöltetlen üres helyeire (melyek egyes nyelveken be lehetnek töltve). Szóba került Cutter specifikus tárgyszavazási szabálya, amely a nem fogalom és faj fogalom pontos meghatározására hívja fel a figyelmet.

A gondolkodó ember általában hajlamos az elvonatoztatásra, a dolgok több szempontú megközelítésére, az összevetésre. Ez nehezítheti az indexelés műveletében a megfelelő faj fogalom vagy nem fogalom kiválasztását. A hiponimák példája igazolja, hogy maga a nyelvi rendszer is osztályoz akkor, amikor a *tulipán*, *árvácska*, *ibolya* lexémákról azt állítja, hogy a *virág* hiponimái. Ugyanakkor, mint említettük, a nem fogalom és a faj fogalom a dokumentum tárgya felől egyaránt lehet specifikus tárgyszó (az általános is lehet egyedi), az információkereső nyelv szempontjából pedig kváziszinonimák is lehetnek. Vagyis a fogalmi szint és a szemantikai szint ugyanúgy egyidejűleg meghatározó, ahogyan a nyelvi szinten belül a szemantika és a szintaxis egyidejű kölcsönhatásban áll.

A fogalmi és a nyelvi szint kölcsönhatásait sokkal nehezebb szétválasztani, mint amennyire könnyen össze lehet keverni. De azok a bizonyos hiányzó láncszemek (nyelvi szinten betöltetlen lexémahelyek) és ugyanakkor fogalmilag megragadható csomópontok tanúsítják, hogy a szintek között hasonlóságok és különbségek létezhetnek.

Ha a felsorolt (fizikai, felhasználói, nyelvi, fogalmi) szintekről az indexelés végén visszatérünk a fizikai szintre, a tárgyszó végső formájára, megállapítható, hogy az szigorúan függ mindattól, amit az előzőekben vázoltunk.

Mégis ezen a bizonyos fizikai szinten, melyet a 20. sz. végén a számítógép használata segít, számos olyan technikai könnyítés áll a keresési folyamat rendelkezésére, amely növeli az információkereső nyelv hatékonyságát. Ezek a gépi keresés klasszikus felhasználóbarát lehetőségei, a különböző típusú csonkolási lehetőségek, a tárgyszavakkal/deszriptorokkal végezhető logikai műveletek, a keresőelemek listájából (a mutatókból) való választás stb. Nem árt hangsúlyozni, hogy gépi rendszerben a cím és/vagy a kivonat szavaira való kereshetőség is szolgálhat segédeszközzel.

Mivel a gépi keresés lehetőségei elméletileg korlátlanok, ha ragaszkodunk osztályozási rendszerünkhöz (pl. az ETO-hoz) is, akkor a két rendszer elemeinek egymáshoz rendelése, konkordanciája (pl. az ETO és a természetes nyelvű mutatója) sem kivitelezhetetlen.

Az indexelés bonyolult művelete számos meghatározó szint kölcsönhatására épül. A művelet lehetséges modelljeit, mintázatait a különböző

szintek teljes mélységű feltárása, és a kapcsolatok térképezése körvonalazhatja. A tanulmány csupán vázlatosan érintette a fontosabb szinteket és néhány tipikus problémát. A természetes nyelvet használó információkereső rendszerek joggal tekinthetők dinamikus rendszereknek.

## Irodalom

- [1] SZILÁGYI Tibor (szerk.): A tárgyszó és a tárgyszó-katalógus. Budapest, OSZK KMK, 1968. p. 150.
- [2] KÁPLÁNY Géza: Könyvtárak korszerű rendezése és fejlesztése. Budapest, Technológiai Könyvtárak Barátainak Egyesülete, 1943. p. 80.
- [3] BÁRCZI-BENKŐ-BERRÁR: A magyar nyelv története. Budapest, Tankönyvkiadó, 1967. p. 371.
- [4] KELEMEN János: A nyelvfilozófia kérdései. Budapest, Kossuth Kiadó-Akadémiai Kiadó, 1977. p. 177. ISBN 963-09-0834-4 (Kossuth), ISBN 963-05-1302-1 (Akadémiai)
- [5] UNGVÁRY Rudolf: Tezaurusz-technológia. Budapest, NPI, 1979. p. 118-121. ISBN 963-201-062-0
- [6] CRYSTAL, David: A nyelv enciklopédiája. Budapest, Osiris, 1998. p. 138. ISBN 963-379-211-8
- [7] SZÉCSI Gábor: Tudat, nyelv, kommunikáció. Budapest, Áron, 1998. p. 69. ISBN 963-85864-5-1
- [8] NEUMER Katalin (szerk.): Nyelv, gondolkodás, relativizmus. Budapest, Osiris, 1999. p. 48. ISBN 963-379-585-0
- [9] SMITH, J. M.-SZATHMÁRY Eörs: Az evolúció nagy lépései. Budapest, Scientia, 1997. p. 351. ISBN 963-8326-11-5

Beérkezett: 2000. III. 9-én.

## ENSZ-webhely a harmadik világ szolgálatára

Az Egyesült Nemzetek Fejlesztési Programja (United Nations Development Programme = UNDP) nagyszabású projektet indított *NetAid* címmel, hogy az internettechnika és -használat terjesztésével is segítse a harmadik világ egyes országainak társadalmi fejlődését. A *NetAid* projekt célja az internet-hozzáférés, a televízió- és rádióműsorok, valamint híres előadóművészek műsorainak integrálása. A szeptemberben munkába állt webhely keretében három koncertet tartottak New Yorkban, Londonban és Genfben, amelyet élőben közvetítettek televízión és a weben.

A Cisco Systems által fejlesztett hálózatot több mint 1500 szerverre építették ki, amelyek több mint 90 adatközpontban vannak a világ különböző helyein.

/Information Retrieval and Library Automation, 35. köt. 4. sz. 1999. p. 4-5 /

## Online pizzarendelés

A Domino's Pizza az Egyesült Királyság első gyorséttermi hálózata, ahol elektronikusan lehet ételeket rendelni az interneten, így pizzát is; ezért is nevezik a szolgáltatást *E-pizzának*. Ezt a szolgáltatást kísérletképpen a hálózathoz tartozó öt webhelyről lehet megrendelni: Cambridge, Maida Vale, Milton Kynes, Reading-East és Reading-West üzleteiben (valamennyi helység Nagy-Britannia kiemelt tudományos vagy csúcstechnikai központja). A megrendelőknek csak első alkalommal kell regisztrálniuk magukat, ezután azonosítójukkal tudnak elektronikus úton pizzát rendelni.

A hálózat webcíme: [www.dominos.co.uk](http://www.dominos.co.uk)

/ASLIB Online and CD Notes, 12. köt. 9. sz. 1999. p. 9 /

(R. P.)