

30. évf. 8–9. sz. 1983. aug.–szept.

## Tudományos és Műszaki Tájékoztatás

### A MAGYARORSZÁGON KÉSZÜLŐ TEZAUROSZOK A NEMZETKÖZI SZABVÁNYOSÍTÁS TÜKRÉBEN

*Ungváry Rudolf*  
OSZK Fejlesztési Csoport

#### Bevezető

A tezauszok nemzetközi szabványosítása immár több, mint egy évtizedes múltra tekint vissza. Az Unesco égisze alatt megjelenő első dokumentumok még 1970-ben láttak napvilágot [I és II], azóta szinte két-három évenként jelentek meg egyre bővülő, s még mindig nem végleges változataik [pl. III], a legutolsók 1980-ban és 1981-ben [IV és V]. A problémákat csak formálisan jelzi, hogy ezek az utolsó változatok már 64, illetve 85 oldalasak, s nem szabványnak, hanem módszertani irányelvnek készültek. Az NTMIK-ben folyó szabványosítás során (két KGST szabvány és két NTMIK normatív–műszaki előírás) külön dokumentumok készültek a tezausz szerkezetének, formájának szabványosítására [VII és VIII], és külön dokumentumok a készítésük egységesítésére [IX és X]. Időközben sorra jelentek meg az egyes nemzeti szabványok és irányelvek (pl. az NSZK-ban a XI és a XVI) és szabványjellegű dokumentumok (pl. az NDK-ban a különféle módszertani irányelvek, mint XII, XIII, XV). A magyar tezausz-szabvány nem az ISO, hanem a KGST-szabvány adaptációja, és 1976-ban jelent meg [XVII]. Ez a szabvány csak a tezausz szerkezetét, részeit és megjelenési formáját szabályozza; a készítésre és a részletekre vonatkozó magyar módszertani dokumentum megjelenése még várat magára.

A szabványosítás kérdése nyilvánvalóan azért nem jut – legalábbis időleges – nyugvópontra, mert a tezauszok a mai – elsősorban automatizált – információkereső rendszerek egyik meglehetősen változékony formában megjelenő s különböző irányokban fejlődő elemét képví-

selik. Különösen a tartalmi feldolgozás fokozódó automatizálása (a gépi tárolású tezauszok, a részben vagy egészben automatikusan végzett indexelés, az automatikus osztályozás) befolyásolja jelentősen a tezauszok formáját, de ez csak nagyon lassan változtatja meg a tezauszokról alkotott hagyományos elképzeléseket, melyek gyökerükben az osztályozó fogalomról alkotott „terjedelmi” felfogáson alapulnak. E felfogás szerint ugyanis csak az tekinthető szorosan vett osztályozó fogalomnak, melynek osztályába tárgyasítható jelenségek tartoznak, melyeknek terjedelme van, tehát valamilyen tárgykört képviselnek (pl. konkrét tárgyak, folyamatok, vagy meghatározható tudományos, elvont jelenségek, mint a SZÁM vagy a TÉTEL). Az automatikus szövegfeldolgozás és -elemzés ezzel szemben olyan elemeket is bevon a feldolgozásba, melyek első látásra csak nyelvek és nehezen állítható róluk, hogy terjedelmük van, hogy meghatározható dolgok osztályát képviselik (ilyenek például az olyan szavak, melyek pusztán jelzői szereppel rendelkeznek, mint KÖRKÖRÖS, MEREDEK, MIKRO-, vagy a számtartományok kifejezései, mint 1-TŐL 15 VOLTIG, de a tezauszokra vonatkozó hagyományos felfogás szerint ide tartozik minden olyan kifejezés, melyet az adott szakterületen csak módosító, „színező” értelemben használnak és ezért modifikátornak neveznek, mint IRÁNY, KÖVETKEZMÉNY, vagy akár az ANYAG, MOZGÁS, TULAJDONSÁG szavak is).

Az automatizálás hatására tehát egyre inkább olyan tezauszok születnek, melyekre a hagyományos teza-

urusz-felfogás csak korlátozottan érvényes. A szabványosításban pedig feltehetően ezért nem jutottak eddig közös nevezőre, mert a hagyományos felfogásból kiindulva próbáltak számot vetni a gépesítés okozta folytonos változásokkal.

Az első magyar nyelvű teaurusz 1970-ben jelent meg [1]. Azóta – tudomásom szerint – legalább 30 egynyelvű és 6 két- és többnyelvű teaurusz készült Magyarországon (beleszámítva a készüléfében levő, vagy félbemaradt változatokat is). Bár a korszerű információkereső rendszerek építésében korántsem tartunk ott, ahol számos nyugat-európai állam, a hazai teauruszok nagyjából mégis ugyanazokat a szerkezeti és készitési problémákat vetik fel. A teaurusz-szabványosítás nemzetközi dokumentumai a több, mint egy évtizedes beljük fektetett munka és viszonylag nagy terjedelmük következtében tartalmazzák mindazokat a kérdéseket, amelyekkel kapcsolatban a teauruszokkal – s így a hazai teauruszokkal is – kapcsolatos problémák felmerülnek. Az alábbi tanulmányban e szabványok – elsősorban az ISO legutóbbi szabványtervezetei [IV és V] – alapján megpróbálok áttekinteni azokat a gyakorlati problémákat, melyek a magyarországi teauruszkészítés nyomán felvetődnek.<sup>1</sup>

Az egyes szabványokat és szabvány-jellegű dokumentumokat érdemben nem bírálok; csak az utolsó fejezetben foglalom össze röviden a fontosabb következtetéseket. Annyit azonban megjegyzek, hogy az alkalmazás kérdésében az összes eddigi dokumentumban *szűken értelmezték a teaurusz fogalmát*, aminek következményei a legfontosabb gyakorlati kérdésben, a gépi felhasználású teauruszokban tapasztalhatók. E téren hiányzik a szabályozás. A gépi teauruszkészítés és használat kérdésével jelentősége miatt külön tanulmányban foglalkozom<sup>2</sup>, s ezért itt – a 6. fejezetben – csak röviden összegezem a gépesítéssel összefüggő hazai tapasztalatokat.

Az 1. táblázatban a vizsgált egynyelvű teauruszok legfontosabb jellemzőit foglaltam össze.

<sup>1</sup> Nem célok sem magának a teauruszkészítés gyakorlatának, sem maguknak a magyarországi teauruszoknak az ismertetése. Erre vonatkozóan utalok az alábbi dokumentumokra, melyekben a kérdésre vonatkozó további szakirodalom is megtalálható:

SZABÓ Julianna: A teauruszkérdés hazai irodalmának értékelése. Bp. 1979. (Szakdolgozat, ELTE), 80 p.

UNGVÁRY Rudolf: Teaurusz-technológia: Az információkereső teauruszok készítésének folyamata. Bp. OSZK KMK, 1979. 277 p.

UNGVÁRY Rudolf-ORBÁN Éva: A két- és többnyelvű teauruszok irodalma: Kritikai szemle és annotált bibliográfia. Bp. OSZK KMK-KG INFORMATIK, 1981. 124 p.

<sup>2</sup> *A teauruszok automatizálása. A gépi teauruszkészítés és használat néhány kérdése* címen a TMT következő számainak egyikében jelenik meg – *A szerk.*

Az egynyelvű teauruszok esetében az ISO 2788–74 nemzetközi szabvány [III] felülvizsgálatára készült ISO, DIS 2788 szabványtervezet 1981. március 2. kiadásának ideiglenes változatát [IV] – a továbbiakban: Tervezet – vettem alapul, a két- és többnyelvű teauruszok esetében pedig az UNISIST irányelvtervezetét [V], de figyelembe vettem a korábbi változatokat [II, VI], továbbá az NTMIR [VII, IX ill. VIII, X], valamint az NSZK-ban [XI ill. XVI] és az NDK-ban [XII, XV ill. XIII] megjelent dokumentumokat is.

### 1. A szabványok szerint készíthető teauruszok (a Tervezet 1.6 pontja)

A szabványosítás kezdettől fogva az ún. *információkereső teauruszokra* vonatkozott. A legkorábbi szabványban [III] még viszonylag általánosan határozták meg az alkalmazási területüket: „... alkalmazható az információ és dokumentáció területén”. Ezen azt értették, hogy „... felhasználható szóasszociációs jegyzékként, hogy segítse az indexelőt, vagy tekinthető az információkereső és a természetes nyelv közötti átalakítás eszközeinek”. A későbbi szabványtervezetben [II] erre vonatkozóan így fogalmaztak: „... a dokumentációs rendszerekben végzett tartalmi feltárára, osztályozásra és információkeresésre használható”

Ezeknek a meghatározásoknak még megfelelhettek azok a teauruszok is, melyeket nemcsak a szorosan vett bibliográfiai vagy katalógustételek, hanem másfajta információtételek tartalom szerinti keresésére is használhattak. A legújabb Tervezet viszont már sokkal szigorúbb megkötevéseket tartalmaz. 1.1 pontjában nemcsak „tárgymutatóról” és „indexelő nyelvről” beszélnek, hanem arról, hogy a teaurusz „... a dokumentumok azonosításának eszköze”. Az 1.6 (c) pont szerint a Tervezet csak azokra a teauruszokra vonatkozik, melyeket *emberi közreműködéssel végzett* (tehát nem automatikus) indexelésre használnak, az 1.6 (d) pont szerint pedig az ilyen teauruszt arra használják, hogy „... katalógusok vagy bibliográfiák formájában feltárt dokumentumok gyűjteményét indexeljék vele”. Csak mellékesen említik meg, hogy a jól szerkesztett teaurusz az automatikus indexeléshez és a szabadszöveges keresőrendszerekben is alkalmazható.

Ha meggondoljuk, hogy valójában éppen a gépesítés az oka a teauruszok rendkívüli elterjedésének, a fenti fogalmazás elég szerencsétlennek és szűknek tűnik. Ha ui. szó szerint értjük, akkor az automatizált bibliográfiai adatbázisokhoz használt teauruszok eleve nem jöhetnek szóba, hiszen a mágnesszalagon már nem valamiféle „katalógust” vagy „bibliográfiát” tárolnak, hanem elemi adattételek – többek között invertáltan – rendezett

Az ISO/DIS 2788 szabványtervezetben megadott és a Magyarországon elkészült tezauszok néhány jellemző adatának összehasonlítása

Vonatkozó szabványpont és jellemző megnevezés N /-szabványban nem szerepel	Egynyelvű tezauszok a bibliográfiai tételek sorrendjében [ 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 ]																													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
<b>Tezauszrészecskék:</b>																														
9.2 Betűrendes főrész	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
9.3 Szisztematikus rész	x																													
9.3.3 /b/ Mutató a szisztematikus részhez																														
9.4 Grafikus rész:																														
9.4.3 Pa-struktúra																														
9.4.4 Hálószerkezet																														
9.4.3/4 /b/ Mutató a grafikus részhez																														
N Permutált mutató	x	x	x																											
8.5.2 Szakterületek szerinti szervezethez	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
8.5.3 Faszették szerinti szervezethez																														
<b>Reldoidok:</b>																														
8.2 Nemdeszkriptor-deszkriptor kapcsolat	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
N ES-kapcsolat																														
N VAGY-kapcsolat																														
8.3.4 Nem-faj relidoid																														
szigorú hierarchia	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
laza hierarchia																														
8.3.5 Egész-rész relidoid	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
8.3.7 Polihierarchia	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
8.4 Rokonsági relidoid	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
N Irdnyultság-függőségi relidoid																														
N Egyéb rokonsági relidoid																														
<b>Készítés:</b>																														
10.1.1 /a/ deduktív																														
10.1.1 /b/ induktív	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
fordítás																														
10.6 gépi készítés	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
10.6.2 gépi tárolás																														

tömegét, amelyből a mindenkori felhasználó tetszés szerinti ún. célinformációkat választhat ki, melyeknek egyáltalán nem kell teljes bibliográfiai tételt képviselniük (lehetnek pl. csupán szerzőnevek, referátumok, sőt referátumok illetve eredeti szövegek részei is).

A gyakorlat kezdettől fogva ellentmondott a Tervezetben foglalt leszűkítésnek, s erről a hazai teauruszok is tanúskodnak. Számos olyan teaurusz készült, melynek nem elsődleges célja a hagyományos „katalógus” vagy „bibliográfia” előállítása, hanem tetszés szerinti adatok tartalom szerinti megállapítása a tárolóban [pl. 9, 10, 11, 13, 17, 18, 22, 24, 25, 26, 27, 29].

Különösen vonatkozik mindez az ún. terminológiai célból készült teauruszokra. A terminológiai és doku-

mentációs teauruszok nagyfokú hasonlóságát a szakirodalom már régen kiemeli.<sup>3</sup>

Magyarországon is megjelent olyan teaurusz [3], amely célját tekintve terminológiai, megjelenési formájában viszont megfelel az információkereső teaurusznak. A benne szereplő definíciók egyrészt tekinthetők a vezéreszkriptorhoz tartozó megjegyzéseknek, de felfoghatók úgy is, mint meghatározott információ- vagy adattétel (1. ábra). Ez utóbbi esetben az 1972-ben megjelent, strukturált informatikai fogalomgyűjteményünk [3] és a hozzá hasonló termékek valójában

<sup>3</sup> FELBER, H.: Some basic issues of terminology = The Incorporated Linguist, 21. köt. 1. sz. 1982. p. 12-23.

textuális összefüggés

SZ szövegösszefüggések

teaurusz

2.2

A szemantikai rendszerek → deszkriptorszótára. Tartalmazza az indexelő nyelvben elfogadott összes — deszkriptort, a deszkriptorok értelmi összefüggéseit; a természetes nyelv azon fogalmait, amelyek a felvett deszkriptorok → szinonimái vagy → kváziszinonimái / → tiltott szavai/, utalással arra, hogy a természetes nyelvi kifejezés helyett a szótár melyik kifejezést használja deszkriptorként. Jellemzői: ABC-je és szóképzési előírásai megegyeznek a természetes nyelvvel; deszkriptorállományát egy laza hierarchia keretén belül tárgyköri csoportokra osztja; eszközöket tartalmaz a deszkriptorok közötti kapcsolatok jelölésére. Felhasználása: osztályozásnál és → információkeresésnél.

- R információkereső nyelv szótára
- F ASTIA típusu teaurusz
- P EJC típusu teaurusz
- P EURATOM típusu teaurusz
- K deszkriptorszótár
- E szemantikai osztályozás, osztályosó rendszer
- K szóbokor
- K szótár
- K tiltott szó

Thermofax eljárás

"SZ" hőre másoló eljárások, termokópia

tiltott szó

2.1

A természetes nyelv azon szava, amely a → deszkriptor rokonértelmű szava. A szótár alapján deszkriptorként nem használható, hanem helyette a megfelelő deszkriptort kell használni.

- K teaurusz
- T nem deszkriptor

tizedes rendszer

2.2

Olyan osztályozó rendszer, amely felépítésében és jelzetelésében a tizedes számrendszert és a jeleket tizedes értelemben használja.

- K ETO/Egyetemes Tizedes Osztályozás/

a teaurusz vezérszavának vagy az invertált file kulcsának tekinthető elem

a „magyarázatnak” vagy az információ-tételnek tekinthető elem

a vezérszóhoz kapcsolódó vagy az invertált file kulcsaihoz mutatókkal /indexekkel/ kapcsolódó további kulcsok

1. ábra Részlet a terminológiai teauruszból [3]

A megjegyzéseket a jobb kiemelés kedvéért kereteztük be az ábrán; az eredeti teauruszban nincsenek bekeretezve

tezausz segítségével rendezett terminológiai adatokat (tétéleket) tartalmazó fájlnak vagy információkereső rendszernek tekinthetők, amelyben a tétéleket az egyes meghatározások képviselik.

Az ilyen „információkereső rendszerbe foglalt” tezausz kétségtelenül a Tervezet 1.1 pontjában említett „jól szerkesztett információkereső nyelvet” tartalmaz, de a Tervezet alapján nem valószínű, hogy tezauszknak tekinthető.

## 2. A megnevezések szintje

(a lexikai egységek és tartozékaik)

### 2.1 A kifejezések ellenőrzése és válogatása

A többnyelvű tezauszokra vonatkozó, Austin készítette 1980-as irányelvtervezetben [V] a tezauszt még „válogatott” kifejezések szótárának tekintik. A Tervezet 3.6 pontjában közölt meghatározásból, akárcsak az 1974-es szabványból [III] ez a jelző hiányzik. A Magyarországon elkészült, illetve készüléfelben levő 30 egynyelvű és 6 készüléfelben levő vagy félbemaradt két- és többnyelvű tezausz közül 16 esetében megállapítható, hogy legfeljebb a szó általános értelmében válogatták a kifejezéseket, amikor felvették őket a tezauszba. Ez a külföldön készült tezauszokban is gyakran megfigyelhető.<sup>4</sup> A felvett szavak ugyanis nem korlátozódtak kizárólag a szűken vett szakkifejezésekre, hanem „lingvisztikaibb” szempontok is szerepet játszottak a felvételben (pl. az olyan melléknevek felvételében, mint „fordító”, „átlag”, vagy az olyan elemek felvételében, mint „mikro-”, „-mentes” stb.). A [7, 9, 10, 11, 12, 13, 14, 17, 18, 22, 23, 24, 25, 26] valamint a [2T és 4T] tezausz tehát legfeljebb a Tervezetben meghatározott tezausz fogalmát elégíti ki maradéktalanul, a szabványban [III] meghatározott tezauszt nem egyértelműen. (Lásd még a 2.2.3 A szófajok című fejezetet!)

### 2.2 A szavak fajtái (a Tervezet 6.1 pontja)

A megvizsgált tezauszokban a szavak több fogalmi kategóriából kerültek ki, mint amennyit a tervezet 6.1.1 pontjában említene; másrészt a Tervezetben közölt kategória-típusizálás – mely a 2. táblázat első oszlopában látható – következtelen, ennek alapján a szavak nem kategorizálhatók egyértelműen.

A „földrajzi jelenségek” a Tervezetben ugyanabba a kategóriába kerültek, mint a „tárgyak” („things”), holott az utóbbiak sokkal inkább „helyek” (pl. alföld,

hegycsúcs, folyó), s mint ilyenek inkább a „tulajdonságok” vagy „jellemzők” kategóriájába tartoznak. Azokban a magyarországi tezauszokban, ahol kategorizálásukra sor került, a „földrajzi képződmény” önálló kategóriáját alkotják [9, 10, 11, 13, 17, 18, 19, 27] vagy az „egyéb” kategóriába kerültek [14].

Az ÉTK tezauszában [14] a tevékenységeket jelentő fogalmakat a „szellemi tevékenység”, illetve az „anyagi tevékenység” kategóriákba sorolták, a folyamatokat – a spontán tevékenységeket – pedig a „természeti tulajdonság” kategóriába. A [9, 10, 11, 13, 17, 18, 19, 27] tezauszokban „tevékenység” és „folyamat” kategóriákat különböztetnek meg. A Tervezet 6.1.1 (b) pontjával ellentétben, a folyamatokat egyik esetben sem tekintik elvont entitásoknak.

Ugyancsak különválasztják a hivatkozott tezauszokban a tulajdonságokat az elvont entitásoktól. Az ÉTK tezauszában [14] a „természeti tulajdonság” és „társadalmi tulajdonság” szerepel, a többiben „tulajdonság”, „idő” és „állapot”.

A Tervezet 6.1.1 pontja alapján nem lehet „mit kezdeni” az olyan fogalmakkal, melyek szellemi eszközöket (pl. fejlesztési koncepció), tudományágakat vagy ismeretterületeket (pl. matematika, vízügy), iparokat (pl. népgazdaság, bőripar), számtartományokat stb. jelentenek.

A 2. táblázatban a megvizsgált tezauszokban használt kategóriák összehasonlítása látható.

#### 2.2.1 Az intézménynevek (A Tervezet 6.6.8 pontja)

Az intézménynevek felvételét általában nem javasolják a Tervezetben. A tezauszokba mégis bekerülnek intézménynevekkel kapcsolatban ugyanakkor részletesen foglalkoznak az egyes névváltozatok (pl. ENSZ – UNO – EGYESÜLT NEMZETEK SZERVEZETE stb.) közötti nemdeszkriptor–deszkriptor kapcsolatok kialakításával.

A magyarországi gyakorlatban a szűkebb szakterületek tezauszaiiban gyakran szerepelnek intézménynevek is [2, 7, 13, 24, 25, 26], feltehetően azért, mert a kisebb szakterület következtében még áttekinthető a számuk. Mindig az eredeti nyelvű, hivatalos nevet vették fel. A Tervezetben ajánlott nemdeszkriptor–deszkriptor kapcsolatok kialakítására úgyszólván soha sem került sor, mivel a névváltozatokat – elsősorban az azonosíthatóság egyszerű volta következtében – nem vették fel nemdeszkriptorként.

Sokszor arra is sor került, hogy az intézménynevek között nem–faj hierarchián alapuló összefüggéseket jelöltek (pl. olyan fölérendelt fogalmak bevonásával, mint MAGYAR VÁLLALAT, NEMZETKÖZI VÁLLALAT, ALAPANYAGGYÁRTÓ, SZOLGÁLTATÓ INTÉZMÉNY, KÜLFÖLDI GÉPGYÁRTÓ VÁLLALAT stb.). Annak érdekében pedig, hogy a vállalatnevek, illetve a

<sup>4</sup> UNGVÁRY R.–ORBÁN É.: hivatkozott mű.

Az ISO/DIS 2788 szabványtervezetben közölt és a magyarországi teauruszokban használt fogalmi kategóriák és alkategóriák (fazetták) összehasonlítása

Tervezet	A [7, 9, 10, 11, 12, 13, 17, 18, 22, 23, 24, 25, 26] és a [2T, 4T] teauruszok kategóriái	A [12] és a [14] teauruszok kategóriái
Elvont entitások (i) Tevékenységek és események	(Mozgások) Folyamat Tevékenység	Anyagi tevékenység Szellemi tevékenység
Konkrét entitások (ii) Anyagok	(Körülhatárolt anyag)	Termék
(i) Tárgyak és fizikai részeik	(Körülhatárolatlan anyag) Tárgy Gép Épületem Létesítmény Földrajzi és csillogászati képződmény Élő anyag Szellemi termék	Anyagi eszköz  Termék Létesítmény  Szellemi termék
Elvont entitások (ii) Anyagok, tárgyak és tevékenységek tulajdonságai	(Jellemzők) Tulajdonság, képesség, állapot Paraméter, szám Hely, település Alakzat, felület Intézmény Idő	Természeti tulajdonság (Társadalmi tulajdonság)
	(Elvont és egyéb) Energiajelenség Kapcsolat Ismeretterület, ipar Gazdasági fogalom Tudományos fogalom Egyéb elvont fogalom	Tudományág Iparág Társadalmi tulajdonság
Egyedi entitások vagy tulajdonnevekkel jelölt „egyelemű osztályok”	—	—

termékek alapján kölcsönösen megállapíthatók a közöttük fennálló összefüggéseket, rokonsági relációval összekapcsolták a termékek, a hatóanyagok, a márkanevek deskriptorait a vállalatnevekkel [13, 24, 26]. Az ilyen teauruszok – ebből a szempontból legalábbis – már nemcsak az információkereső nyelvet tartalmazzák, hanem az adatbázisnak a tulajdonságaival is rendelkeznek, hiszen a vállalatnevek, illetve a termékek adattételeknek is felfoghatók, a közöttük jelölt kapcsolatok pedig ún. rekordok közötti kapcsolatoknak is minősíthetők. (Az

ilyen módon összekapcsolt tételek a teauruszban az adott szakterületre vonatkozó közvetlen, elsődleges információt képviselnek.)

Ezek az információk elsősorban az ad-hoc teauruszhasználó számára jelentősek; nem véletlen, hogy az intézménynevek és a közöttük fennálló kapcsolatok felvételére olyan teauruszokban kerül inkább sor, melyeket online körülmények között kívánnak használni, illetve használnak.

A vállalatnevek felvételekor különleges problémát jelentett a megnevezés végleges formája. Egyes tezauszokban a mindenkor teljes nevet vették fel (pl. SIEMENS AG, NIEDERBERGER+CO AG stb.), másokban – mivel az olyan elemek, mint AG., CO., Ltd. stb. a vállalatok fejlődésével változnak – csak az alapnevet (SIEMENS, NIEDERBERGER stb.). Erre vonatkozóan a Tervezet sem tartalmaz javaslatot.

## 2.2.2 A tartományszavak

Különleges lexikai egységeket képviselnek a számtartományokat jelentő kifejezések. Ezekre a speciális tezauszokban van szükség például a térképek méretarányok szerinti csoportosításakor, a különböző feszültségeken működő berendezések osztályozásakor stb. Ilyenkor számtartományokat állapítanak meg (pl. „1-től 45 voltig”, „46-től 250 voltig”, „251-től 500 voltig” stb.), és ezekről utalnak vagy egy-egy természetes nyelvű kifejezésre (pl. TÖRPEFESZÜLT-SÉG, KISFESZÜLT-SÉG, KÖZÉPFESZÜLT-SÉG stb.) vagy egy-egy számértékre (pl. az 5000-tól 75 000-ig terjedő méretaránytartományról az 1:75 000 méretarányértékre), mely a deszkriptor szerepét játssza, szemben a tartományértékekkel, melyek nemdeszkriptorok [27, 6T].

A Tervezetben a kérdéssel nem foglalkoznak. A 7.1 szakasz (c) pontjában közölt példa a tartományt megjegyzés formájában tünteti csak fel.

## 2.2.3 A szófajok

### 2.2.3.1 A főnevek (a Tervezet 6.2.1 és 3.3 pontja)

A Tervezet hivatkozott pontjai szerint a lexikai egységek lehetőleg főnevek vagy főnévi kifejezések legyenek. Egyes szaktezauszokban, a gépesített információkereső rendszerekben használt tezauszokban, továbbá a szláv nyelvekhez kapcsolódó két- és többnyelvű tezauszokban még ez a „lehetőleg” kifejezés is nehezen követhető, mivel nem-főnévi formákat – olykor elég nagy számban – is fel kell venni.

### 2.2.3.2 A melléknevek (a Tervezet 6.2.2.2 pontja)

Különösen sok melléknevet tartalmaztak a vizsgált tezauszok, holott a Tervezet deszkriptorként nem javasolja alkalmazásukat. Az 1.7 pont szerint a Tervezet előírásai „mind nyelvtől, mind kultúrkörötől függetlenek”. Egyes nyelvekben – pl. az oroszban – a melléknevek önálló szerepe a tezauszokon belül lényegesen nagyobb, mint amekkora figyelmet a kérdésnek a Tervezet 6.2.2.2 pontjában szentelnek [2T, 3T, 4T, 5T, 6T]. A gépesített információkereső rendszerekben alkalmazott tezauszokban pedig a feldolgozás hangsúlyozottan analitikus jellege következtében nő meg a mellékneveket képviselő deszkriptorok száma [9, 10, 11, 13, 17, 18, 24, 25, 26, 27].

Ilyen melléknevek például az OLDALIRÁNYÚ, TERELŐ, FORDÍTÓ stb. Gyakori, hogy közöttük nemdeszkriptor–deszkriptor kapcsolatokat határoznak meg (pl. az ÖSSZETETT képviseli a KOMBINÁLT, KAPCSOLT, KEVERT, VEGYES, BONYOLULT, KÉTÁLLÁSÚ, HÁROMÁLLÁSÚ, TÖBBRÉSZES stb. nemdeszkriptorokat).

### 2.2.3.3 Az igék (a Tervezet 6.2.4 pontja)

Jóllehet maguk az igék semmilyen formában nem kerülnek be közvetlenül a hagyományos információkereső tezauszokba, meg kell jegyezni, hogy automatizált indexelő rendszerekben az igék szótövei – akár csak a főnevek és egyéb szófajba tartozó szavak tövei – nemdeszkriptorként kapcsolódnak a tezausz deszkriptoraihoz. Mivel az ilyen tezauszokat általában gépi adathordozón tartják nyilván, a kinyomtatott változatban az ilyen – többnyire relatív töveket képviselő – nyelvi elemek nemdeszkriptorként jelennek meg a deszkriptorok alatt [6T]. A Tervezetben *elutasítják az igék felvételét*. Mivel a gyakorlat minden jel szerint mégis kikényszeríti a felvételüket, feltételezhető, hogy – szabványelőírások híján – a jövőben meglehetősen különböző megoldási formák alakulnak majd ki.

### 2.2.3.4 Az egyéb szófajok

A tartományszavak (lásd a 2.2.2 fejezetet) valójában számneveket képviselnek, s közöttük ugyanúgy állapítanak meg nemdeszkriptor–deszkriptor kapcsolatokat, akár a többi szófajhoz tartozó deszkriptorok között [27, 6T].

A 2.2.3.2 fejezetben említett okokból kerül sor az olyan nyelvi elemek felvételére is olykor, mint pl. MIKRO-, UTÁN-, SEGÉD-, SZERVO- stb. [24, 25, 2T, 4T, 6T]. Ezek között is előfordulhatnak nemdeszkriptor–deszkriptor kapcsolatok.

Pl.: NÉLKÜLI                      SEGÉD  
L MENTES                      L SZERVO

A Tervezet 6.2.3. pontjában az ilyen szavak felvételét elutasítják.

## 2.2.4 Az egyes- és többesszám kérdése (a Tervezet 6.4.2.1 pontja)

A megvizsgált tezauszok túlnyomó többségében minden lexikai egység – hacsak nem létezett kizárólag többesszámú alakban, vagy az egyes és többesszámú alak jelentése nem eltérő – következetesen egyszámúban szerepelt.

A Tervezet 6.4.2.1 pontja szerint a megszámlálható entitásokat lehetőleg többesszámúban kell felvenni, a többit pedig egyszámúban. Ugyanakkor számos anyag-név esetén is többesszámot javasolnak (pl. MŰANYAGOK). A gyakorlatban ez a szabály annyira bizonytalanul alkalmazható, hogy azokban a tezauszokban, ahol

megpróbálkoztak többszám alkalmazásával, teljes a következetlenség. Ha meggondoljuk, hogy a többszám-  
nak valójában semmiféle jelentésmódosító funkciója nem  
lehet az osztályozás szempontjából – kivéve azt az  
esetet, amikor a szó *csak* többszám-  
ban fordul elő vagy többszám-  
ban más jelentést, nem sok értelme van a  
többszámú alak megkülönböztetett kezelésének, és  
semmiféle egyértelmű és általánosan alkalmazható sza-  
bály nem állítható fel. Semmivel sem gazdagítja a  
teauruszt, ha a KÉMIAI REAKCIÓK többszám-  
ban szerepel, a BOMLÁS pedig – mint egyik alárendeltje –  
egyszám-  
ban, és a BOMLÁS további alárendeltjei (pl. a  
CSEREBOMLÁS) ugyancsak egyszám-  
ban.

Ilyen szabály minden jel szerint egyedül az állat- és  
növényrendszerekben állítható fel egyértelműen, mivel  
ezekben a rendszerekben az egyes hierarchiaszinteknek  
meghatározott *nevük* van, és a név alapján (pl. család,  
alcsalád, faj, alfaj, nemzetség stb.) meghatározható, hogy  
mely szint szavai szerepeljenek többszám-  
ban (pl. KUTYAFÉLÉK, melyek alárendeltje a FARKAS, a KUTYA  
stb.). Ilyen eljárást alkalmaztak például a növény-  
védelmi és agrokémiai teauruszban [13].

## 2.2.5 A homonímia (a Tervezet 6.5 pontja)

A Tervezet csak a valódi homonimákkal foglalkozik,  
holott – elsősorban a szűkebb tárgy körű teauruszokban  
– gyakori, hogy homonimaként kezelnek valójában  
közös főlérendelthez tartozó kifejezéseket /pl. HULLA-  
DÉK (BŐR), HULLADÉK (SZENNY)/, és ezeket nem  
rendelik semminek sem alá. Ennek az az oka, hogy az  
adott teauruszban pl. a HULLADÉK nagy része kizáró-  
lag bőrhulladék, kisebb hányaduk viszont mindenféle  
más, a gyártással összefüggő hulladék. Mivel a dokumen-  
tumban mindkét esetben többnyire hulladékról van szó,  
és csak a szövegből derül ki, hogy ezen mit értenek, az  
indexelő többnyire a HULLADÉK deszkriptort nézi meg  
először a teauruszban. Az alkalmazott megoldással az  
indexelő rákényszerül arra, hogy pontosabban megvizs-  
gálja, milyen hulladékról szól a dokumentum. Így jártak  
el pl. a bőripari teauruszban [25].

Ugyancsak homonimaként jelennek meg gyakran a  
szempont szerint felosztott fogalmak /pl. a SZÁMALK  
teauruszban a SZÁMÍTÓGÉP (MŰKÖDÉS SZERINT),  
SZÁMÍTÓGÉP (SZERKEZET SZERINT) stb./ (lásd a  
3.2.1.2 fejezetet).

## 2.2.6 A szinonímia

### 2.2.6.1 A szinonim szavak közötti választás (a Tervezet 6.6 pontja)

Annak, hogy a szinonim kifejezések közül melyiket  
válasszák deszkriptornak, a gépesített információkereső  
rendszerek teauruszaiban tulajdonítják a legkisebb je-  
lentőséget.

Egyes teauruszokban különleges problémákat okoz-  
nak a köznapi és a tudományos kifejezések közötti  
szinonim összefüggések. Mindaddig, míg a szóban forgó  
szinonima nem a speciális szakterület kifejezései közé  
tartozik, egyértelműen eldönthető, hogy melyik változat  
legyen a nemdeszkriptor és melyik a deszkriptor. Amint  
azonban fontos szakkifejezésnek számít, gyakran kide-  
rül, hogy a köznapi név nem egyértelmű; többnyire  
gyűjtőfogalma több tudományos névvel jelölt fogalom-  
nak (pl. a BÁLNA és a CET csak a köznyelvben  
szinonimák, az állatrendszerben az első alárendeltje a  
másodiknak) vagy különböző dolgokat kapcsol össze, és  
nem egyértelműen. Ilyenkor tehát nem állapítható meg  
egyértelműen a szinonímia. A körülményes homonimá-  
képzés /pl. CET (BÁLNA) L BÁLNA; CET (CETACEA)  
L CETACEA/ helyett sok teauruszban az egyszerűbb  
VAGY-kapcsolatot alkalmazzák [13, 22, 25, 2T].

Pl.: CET

LV BALANEA  
CETACEA

A szűkebb tárgy körű teauruszokban általában a  
tudományos megnevezést választják deszkriptornak, és a  
köznapi megnevezés a nemdeszkriptor (a Tervezet 8.2.2  
(b) pontja).

A Tervezet nemdeszkriptorok és deszkriptorok közöt-  
ti VAGY-kapcsolatokat egyáltalán nem említi: egyes  
hazai teauruszokban ezek is előfordulnak (lásd még a  
3.1.2 A VAGY-kapcsolat című fejezetet).

### 2.2.6.2 A kváziszinonimák

(a Tervezet 8.2.2 (b) és 8.2.4 pontja)

Kváziszinonim kapcsolatokat a magyarországi teza-  
uruszokban nemcsak a nem-faj reláció alapján jelölnek  
ki, ahogy ezt a Tervezet 8.2.3 pontjában javasolják,  
hanem nagyon gyakran egész-rész és egyéb rokonsági  
relációk között is (pl. TÖRÉSMUTATÓ L FÉNYTÖ-  
RÉS; PROGRAMHIBA L PROGRAMELLENŐRZÉS).  
(Lásd még a 3.1.1 Az ÉS-kapcsolat és a 3.1.2 A  
VAGY-kapcsolat című fejezeteket.)

### 2.2.6.3 A kölcsönszavak (a Tervezet 6.6.2 pontja)

Új szakterületeken viszonylag több a kölcsönszó. A  
számítástechnikai [24] és a formatervezési [26] teza-  
uruszban például szokatlanul nagy számban figyelhetők  
meg olyan kölcsönszavak, melyek magyar alternatívák  
híján eleve deszkriptorok, holott a Tervezetben azt  
javasolják, hogy ezek nemdeszkriptorok legyenek.

## 2.2.7 A megjegyzések (a Tervezet 6.7.1 pontja)

A lexikai egységekhez fűzött hosszabb, magyarázó  
megjegyzéseket csak kevés teauruszban alkalmazzák [3,  
15, 24, 25, 26]; ha alkalmazzák, akkor is anélkül, hogy a  
mezőt külön relációjel jelölné. A megjegyzés minden  
esetben a vezérszó alatt szerepel, a vezérszótól eltérő  
betűtípussal írva.



Számítógéppel kezelt teauruszokban olyan megoldás is előfordul, amelyben csak a vezérszó után szereplő jel utal arra, hogy a lexikai egységhez külön mellékletben hosszabb magyarázat tartozik. Ennek oka az, hogy hosszabb magyarázatok a tezurusznyomtató programokkal nehezen kezelhetők [24].

A lexikai egységek forrására, gyűjtésének idejére vonatkozó – a Tervezet 6.7.1 (a) és (b) pontjában említett – adatokat gépi teauruszkezeléskor a vizsgált rendszerekben a programrendszer a megnevezésrekordban speciális módon, külön visszakereshető formában tárolja és nyomtatja ki.

#### 2.2.8 A használatra vonatkozó jelölések (a Tervezet 10.5.2 pontja)

A Tervezet külön jel (pl. csillag) vagy megjegyzés – „Indexeléshez nem használható” – alkalmazását javasolja, ha olyan lexikai egységről van szó, melyet csak azért vettek föl a teauruszba, hogy a különben szétszórta speciális fogalmakat föléredelteként összefogja.

Ilyen gyűjtőfogalmakat különösen a gépi teauruszokban alkalmaznak, mégpedig annak érdekében, hogy az automatikus hierarchiaszint-változtatással járó előnyöket kihasználhassák. Többnyire nem kerül sor arra, hogy használatukat kizárják. Nyilván azért, mert átfogó dokumentumok indexelésékor szükség lehet rájuk. A Tervezetben javasolt különleges jelet felhasználják viszont arra, hogy jelöljék velük azokat a lexikai egységeket, melyeket mindig az általuk jelzőként jelzett szó után kell írni a deskriptorláncban (pl. az ALGEBRAI\* kifejezés esetén a használata jelzőként: /FÜGGVÉNY/ALGEBRAI/). Erre a megoldásra olyankor került sor, amikor az alkalmazott programrendszer – például az ISIS – képességeiből nem futotta a szintaktikai relációk használata [24].

#### 2.3 A szavak szerkezete (a Tervezet 7.3 pontja)

Feltűnő különbség figyelhető meg az összetett szavak kezelésében a géppel kezelhető és a géppel nem kezelhető, de egyébként intellektuálisan készült és emberi indexelő által használt teauruszok között. A Tervezet szemantikai (pl. MOSODA = MOSÁS+SZOLGÁLTATÁS+ÜZEM) és morfológiai (önkényesnek tűnő szóhasználatában: szintaktikai, pl. ACÉLTENGELY = ACÉL+TENGELY) összetevők felbontását különbözteti meg, és az első megoldást nem javasolja (a Tervezet 7.1.3 pontja), mivel zajt eredményezhet. A Tervezet készítői nyilván figyelmen kívül hagyták, hogy a zaj elsősorban akkor keletkezik, ha a deskriptorlánc összeállításakor nem alkalmazhatnak szintaktikai kapcsolatjelölőket.

Az automatizált információkereső rendszerekhez használt teauruszok esetén többnyire lehetséges az ilyen

eszközök használata. Ez is magyarázza, hogy az ilyen teauruszokban az összetett szavakat jóval gyakrabban fejezik ki összetevőikkel [9, 10, 11, 13, 17, 18, 22, 25].

Az olyan összetett szavakat, melyek összetevői egyszerű morfológiai elemek, és könnyen felismerhetők, sokszor fel sem veszik a teauruszokba (pl. ACÉLTENGELY, KECSKEBŐR, BÚTORTERVEZÉS stb.). Így a Tervezet 7.3.2.2 pontjában leírt megengedett felbontások nagyrészt teljesen érdektelenek, és sok teauruszban – s azon belül sok szócsoportban – elő sem fordulnak (pl. a Bőr-, szőrme-, műbőr- és cipőipari teauruszban [25] az összes állatbőr eleve hiányzik, és csak az egyes állatnevek és a BŐR szó szerepel; a teauruszhoz tartozó használati utasításban pedig általánosságban utalnak az ilyen kifejezések utólagos összekapcsolására). Azt is meg kell azonban jegyezni, hogy ez a megoldás mindig a gépi információkereső rendszerekkel összekapcsolt teauruszokra jellemző.

Ugyancsak jellemző, hogy az ilyen teauruszokban a jelzős összetételeket (a Tervezet 7.2.2 (c) és 7.3.2.1 (b) pontjai) is többnyire felbontják (pl. BILLENŐKAPCSOLÓ helyett külön van BILLENŐ és külön KAPCSOLÓ), és a mellékneveket nagyon gyakran szerepeltetik önálló deskriptorként (lásd a 2.2.3 fejezetben belül a mellékneveket és az egyéb szófajokat).

### 3. A relációk szintje

#### 3.1 Az ekvivalencia-relációk

A Tervezet gyakorlatilag csak egyetlen fajta nemdeskriptor–deskriptor kapcsolatot fogalmaz meg explicit módon, mégpedig azt, amely egyetlen nemdeskriptor és egyetlen deskriptor között áll fenn (4.4 pont és 8.2 pont). A szinonímia és a kváziszinonímia kérdésével már foglalkoztunk (lásd a 2.2.6 fejezetet). E helyen még a nemdeskriptor–deskriptor kapcsolatok további, a vizsgálat magyarországi teauruszokban előforduló fajtáira térünk ki.

##### 3.1.1 Az ÉS-kapcsolat

A Tervezet 4.5 és 8.2.2 (i) pontjában, valamint a Függelék 2. pontjában ugyan megemlíti az ÉS-kapcsolat lehetőségét, de a kérdést lényegében nem tárgyalja, és azt sem közli, hogy hogyan kell ennek az inverzét jelölni.

Az ISO 2788–1974 változatban ennek a kérdésnek még valamivel nagyobb jelentőséget tulajdonítottak. A magyar szabvány ezt a kapcsolatot pontosan szabályozta: az LL illetve a HL relációjeleket írja elő [XVII].

Az ÉS-kapcsolatot is tartalmazó teauruszokban [9, 10, 11, 13, 17, 18, 24, 25, 26, 27] nagyon gyakoriak voltak a szemantikai összetevőkre bontások:

Pl.:  
 KOMPENZÁLÁSI ÉRTÉK MICODIFOL  
 LL CSERÉLHETŐSÉG LL FOLPED  
 ILLESZTÉS ÖSSZETETT HA-  
 TÓANYAGÚ  
 NÖVÉNY-  
 VÉDŐ-  
 SZER

### 3.1.2 A VAGY-kapcsolat

Ezzel a kapcsolatfajttával a Tervezet egyáltalán nem foglalkozik (lásd még a 2.2.6.1 fejezetet). Ugyanazokban a tezauszokban használják, mint amelyekben az ÉS-kapcsolatot. Gazdaságosságának oka, hogy magyarázatot, illetve hátravetett értelmezőkkel többszörözött lexikai egységeket helyettesít, s ugyanakkor az indexelőt szinte rákényszeríti a gondosabb mérlegelésre.

A kapcsolatot a magyar szabvány szerint az LV relációjellel jelölik, fordítottját a HV relációjellel.

## 3.2 A deskriptorok közötti relációk

### 3.2.1 A nem–faj reláció

#### 3.2.1.1 A polihierarchia (a Tervezet 8.3.7 pontja)

A vizsgált tezauszok többségére a polihierarchia volt jellemző. Monohierarchiát elsősorban néhány korai [2, 4, 5, 6] és néhány speciális [19, 21] tezauszban használtak, s egy esetben az információkereső rendszerben használt programrendszer – a LOGEL – képességeiből nem futotta a polihierarchia [23].

A polihierarchikus kapcsolatok jelentős részét a rejtett rendeltetés–eszköz kapcsolatok alkották (pl. a KÉN egyik fölérendeltje a HALOGÉN ELEM, a másik fölérendeltje a FERTŐTLENÍTŐ SZER, amely azonban rendeltetés-relációban is kifejezhető volna). Ilyenkor – ez az automatikusan kezelt információkereső rendszerekben használt tezauszokra volt elsősorban jellemző – azért választották mégis a nem–faj fölérendelt kapcsolatot, hogy kihasználhassák az automatikus hierarchiaszint-változtatás adta stratégiai lehetőségeket. Más szóval: a relációk közötti értékkülönbség különösen az automatizált körülmények között vált érezhetőbbé.

#### 3.2.1.2 A szigorú és a laza hierarchia

(A Tervezet 8.3.4.2 pontja)

A laza (pl. a KUTYA fölérendeltje a HIÁZÓRZÉS) és a szigorú (pl. a KUTYA fölérendeltje a RAGADOZÓ) hierarchiát egyetlen magyar tezauszban sem jelölik eltérően, holott a Tervezetben ezt javasolják. Ennek egyik oka valószínűleg az, hogy ezt a megkülönböztetést maga a magyar tezausz-szabvány sem írja elő, mivel ezt a szabványt a KGST-szabvány alapján adaptálták, amely erre a kérdésre ugyancsak nem tér ki.

A nem–faj reláció laza kezelése különösen az [1, 5, 6, 15, 20, 21 és 23] tezauszokra jellemző. Ezekben – sajnos – a Tervezet 8.3.5.2 pontjában említett nem ajánlott megoldások is sűrűn előfordulnak a 2. ábrán látható formában.

A nem–faj reláció ilyen kezelése túlnyomó részben a prekoordinált, ún. hierarchikus osztályozási rendszerek, pontosabban az ETO befolyásának tulajdonítható Magyarországon.

#### 3.2.1.3 A tulajdonnév–fajta reláció

(a Tervezet 8.3.6.1 és 8.3.6.2 pontja)

Ez a kapcsolat elsősorban az intézményeket [13, 24, 26], illetve a földrajzi neveket [27] tartalmazó tezauszokban gyakori. Különösen jellemző, hogy az intézménynevek többszintű hierarchialánc legalsó tagjai, azaz a vállalatnév közvetlen fölérendeltje nem az általános fölérendeltet képviselő VÁLLALAT, hanem pl. a KÜLFÖLDI ALAPANYAGGYÁRTÓ VÁLLALAT, ez utóbbi pedig alárendeltje az ALAPANYAGGYÁRTÓ VÁLLALAT-nak és a KÜLFÖLDI VÁLLALAT-nak.

Az ilyen megoldásokat a keresési stratégia javítása érdekében alkalmazzák, és elsősorban online körülmények között van jelentősége (a kérdésbe bevonható deskriptorok hamarabb megtalálhatók, a hierarchiaszint emeléséből vagy süllyesztéséből adódó lehetőségek jobban kihasználhatók). (Lásd még a 2.2.1 *Az intézménynevek* című fejezetet is.)

#### 3.2.1.4 A fölérendelt fogalom szerepe a többi reláció jelölésében

Általánosan megfigyelhető – és a Tervezetben nem említett tény –, hogy azokat az egész–rész és rendeltetés–eszköz relációkat, melyek az alárendelt fogalmakhoz kapcsolódnak, s egyben a fölérendelthez is, mindig csak a fölérendelthez kapcsolódva jelölik. Ennek oka elsősorban a tezausz terjedelmének a csökkentése.

### 3.2.2 Az egész–rész reláció (a Tervezet 8.3.5 pontja)

A Tervezet ajánlásától eltérően, a magyarországi tezauszokban a tudományok és ismeretterületek között (pl. MECHANIKA–FIZIKA, NÖVÉNYTAN–BIOLÓGIA) nem egész–rész, hanem nem–faj relációt jelölnek. Ennek oka többnyire a nem–faj reláció automatikus kezelésének előnyeivel magyarázható.

### 3.2.3 Az irányultság–függőség reláció

(a Tervezet 8.4 pontja)

A Tervezet ezt a relációt nem különbözteti meg a rokonsági relációtól. Számos magyar tezauszban [9, 10, 11, 15, 17, 18, 19, 27] – a magyar tezausz-szabvány alapján [XVII] – külön jelölik. (Szerencsétlen módon a magyar szabványban rendeltetés–eszköznek

nevezik, holott ebbe tartozik minden olyan összefüggés, amely valamilyen értelemben következményt, irányultságot, célt, eszközt stb. fejez ki.)

3.2.4 Egyéb rokonsági relációk (a Tervezet 8.4 pontja)

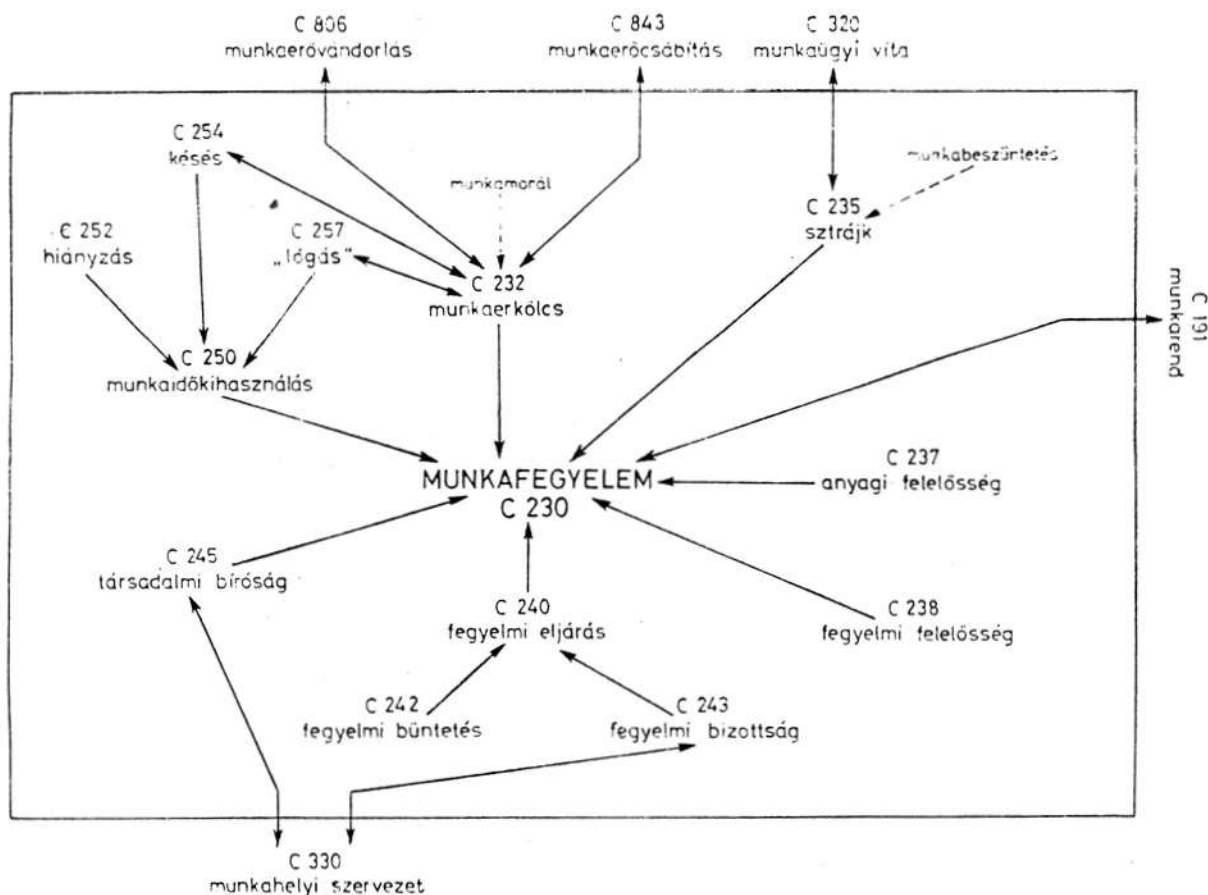
Speciális teauruszokban [13, 22] felmerül az igény, hogy bizonyos összefüggéseket a szokásosnál is pontosabban jelöljenek. Így pl. a növényvédelmi és agrokémiai teauruszban [13] megkülönböztetik a növény–károsítója, a növény–betegsége és a károsító–betegség relációkat. Ennek oka egyrészt az, hogy el akarják kerülni a különböző relációk jelentésének összemosódását, másrészt az a már említett törekvés, hogy a teaurusz egyben – a jobb gépi keresési stratégia kidolgozása érdekében – valamiféle közvetlen, elsődleges információt is tartalmazzon az említett összefüggésekre vonatkozóan. A magyar szabvány [XVII] a rokonsági reláció finomítását ajánlja; a nemzetközi dokumentumok ezzel a kérdéssel nem foglalkoznak. (Lásd még a 2.2.1 *Az intézménynevek* című fejezetet is.)

3.2.5 A többszintű teauruszcikk (a Tervezet 9.2.3 pontja)

A keresési stratégia támogatása érdekében – a Tervezet ajánlásával összhangban – előfordul, hogy a teauruszcikkben nemcsak a fölérendelt deskriptort, hanem az azt követő deskriptort is szerepeltetik [24]. Ez érvényes az egész–rész relációra is, sőt a két reláció együtt is előfordul ilyen összefüggésben.

Pl.: ZSÁKTÁROLÓ  
F1 TÁROLÓ  
F2 SZÁMÍTÓGÉP ARCHITEKTURA  
P1 SZÁMÍTÓGÉP (SZERKEZET SZERINT)

Az indexelés pontosságának növelése érdekében ugyanakkor előfordult, hogy több egymást követő alárendeltet is szerepeltettek [24]. A többszintű teauruszcikk a nyomtatott változatot túlságosan megnöveli, ezért elsősorban online körülmények között van jelentősége.



2. ábra Az ISO 2788 szabvány által nem ajánlott laza hierarchia példája.

(Részlet a munkaügyi teauruszból [5])

Pl.: A FEGYELMI BIZOTTSÁG nem jelent „eljárást”, tehát nem lehet fajtája a FEGYELMI ELJÁRÁS-nak.  
A FEGYELMI ELJÁRÁS nem jelent „fegyelmet”, s ezért nem lehet fajtája a MUNKAFEGYELEM-nek stb.

### 3.2.6 A csúcsheszkrítorcikk (a Tervezet 9.2.2 pontja)

A Tervezet 4.1 pontjában említett csúcsheszkrítor (top term) elsősorban a szisztematikusan is rendezett teauruszokban fordul elő (lásd a 4.1 fejezetet) [4, 14], de majdnem minden, Magyarországon gépi segítséggel készülő teauruszban jelölhető a mindenkor csúcsheszkrítor [pl. 13, 24]. A Tervezettel ellentétben nemcsak nem-faj, hanem egész-rész csúcshesz is előfordul, ami a reláció tranzitivitása alapján természetes.

## 4. A teauruszrészek szintje

A Tervezet a vezérszavak betűrendjében szereplő teauruszcikkeket tartalmazó főrész mellett még számos megjelenési formát javasol, melyek magyarországi változataival az alábbiakban foglalkozunk.

### 4.1 A szisztematikus rész

(a Tervezet 8.5.2 és 9.3 pontjai)

A szisztematikus szervezettségű részeket többnyire az olyan teauruszok tartalmazzák, melyekben a lexikai egységeket pragmatikus szempontok – pl. felhasználási területek – alapján csoportosították [1, 3, 23], és a hierarchikus kapcsolatokat nem a szigorú hierarchia szerint dolgozták ki. Az ilyen teauruszokban figyelhető meg elsősorban az a törekvés, hogy a teauruszcikkben a csúcsheszkrítor is megadják.

### 4.2 A fazettás rendezettségű rész

(a Tervezet 8.5.3 pontja)

Fazettás rendezettségű teauruszrészrel azok a teauruszok rendelkeznek, melyekben a lexikai egységeket a fogalmi kategóriák és alkategóriák alapján csoportosították. A gyakorlatban azonban az ilyen teauruszok nem tartalmazták önállóan a fazetták szerint csoportosított részt, mivel a fazetták kis száma miatt egy-egy fazettába rendkívül sok lexikai egység került volna, s az egyes részek áttekinthetetlenekké váltak volna. A fazettás rész a szakcsoportok szerint rendezett részbe olvadt be, mégpedig oly módon, hogy az egyes szakcsoportokon belül váltak szét egymástól a folyamatok, tevékenységek, anyagok, berendezések, tulajdonságok stb. [pl. 9, 10, 11, 13, 14, 17, 18, 26, 27]. Magának a fazettás beosztásnak a vizsgált teauruszokban elsősorban a teaurusz készítésében és karbantartásában volt szerepe: a fazettákon belül pontosabban és teljesebben lehetett megállapítani a lexikai egységek paradigmikus kapcsolatait.

Különleges esetet képvisel a *Növényvédőszer jegyzék* címet viselő teaurusz [22]; ez valójában a MÉM NAK növényvédelmi teauruszából készült, mégpedig oly módon, hogy kiválogatták belőle az ANYAG fazettába tartozó lexikai egységeket. Ez a teaurusz egyébként további példája a hagyományos információkereső teaurusz szélesebb körű felhasználásának. Közvetlen célja ugyanis nem a dokumentumtételek visszakeresése, hanem a terminológiai szerep: „... hogy eligazítsa a szakembereket a növényvédőszeres és velük kapcsolatos elnevezések rengetegében, tájékoztasson az egyes hatóanyagok kémiai szerkezetéről, a szerkombinációk minőségi összetételéről és a szerek növényvédelmi alkalmazásáról”. Közvetve ennek a teaurusznak is dokumentációs szerepe van, hiszen azt hivatott elősegíteni, hogy a felhasználók pontosabban, a MÉM NAK Növényvédelmi és Agrokémiai információs rendszerében használt teaurusznak megfelelően fogalmazzák meg a kérdéseiket.

### 4.3 A fa-struktúra (a Tervezet 9.4.3 pontja)

A fa-struktúra kizárólagos használatával magyarországi teauruszban csak az Építéstudományi Intézet teauruszában [12] találkozunk. Alkalmazásának oka feltehetően az volt, hogy magában a teauruszban csak nem-faj relációk állnak fenn a deszkriptorok között, rokonsági relációk nincsenek.

### 4.4 A hálószerkezet (a Tervezet 9.4.4 pontja)

A grafikus részt is tartalmazó teauruszokban – az előbb említett [12] teaurusz kivételével – a hálószerkezetet alkalmazták; a paradigmikus kapcsolatok ugyanis kikényszerítették a hálószerkezet alkalmazását. Ebben a rokonsági relációk gazdagságának döntő szerepe volt [pl. a 6, 7, 25 teauruszokban]; a 3. ábrán látható teauruszban a nem-faj kapcsolatok uralkodó szerepe tűnik ki.

A hagyományos nézetek szerint a teaurusz szerkezetének leglényegesebb vonása, hogy a nem-faj relációkat explicit formában kidolgozták benne. A hálószerkezet uralkodó jellege arra utal, hogy ez koránt sincs így: nemcsak a humán tudományok területén készülő teauruszokban, hanem a műszaki teauruszok számos szakcsoportjában sem lehet következetes és összefüggő nem-faj vagy egész-rész hierarchialáncokat kidolgozni; a rokonsági relációk hálójának kidolgozása viszont minden esetben lehetséges (4. ábra). A Tervezet erre a tényre nem utal.

A Tervezet 9.4.4 pontjában javasolt négyzetes beosztású háló – ha több ezer lexikai egységet tartalmazó teauruszról van szó – a gyakorlatban csak kompromisszumok árán valósítható meg. Ez az elrendezés feltételezi, hogy az egyes relációk szerint kapcsolódó







elemeknek van valami szabályos rendezettsége; másrészt pedig ezekben a négyzetekben vagy cellákban elég hely van minden elem számára, amit a gyakorlat nem igazol. Ha tehát következetesen akarják kidolgozni az ilyen hálót, rendkívül finom felbontást kell alkalmazni, azaz a grafikus rész terjedelme áttekinthetetlenül nagy lesz.

Négyzetes beosztású hálót két tezauszhoz dolgoztak ki. Az egyik a Nemzetközi Útügyi Dokumentációs Rendszer (International Road Research Documentation, IRRD) tezauszának fordítása következtében keletkezett (5. ábra) [28]. A magyar szerkesztők megjegyezték: „A grafikus rész – a cellák véges száma miatt – nem tartalmazza a deszkriptorok mindegyikét, a legspecifikusabb kifejezések elmaradnak”. A másik tezausz jelenleg kizárólag ilyen grafikus részből áll, s ebben a formában inkább az általa képviselt szakterület terminológiai rendezésének igényét képviseli, semmint a tartalmi feltárára és visszakereshető tárolásra alkalmas eszközt [30].

#### 4.5 A betűrendes mutató (a Tervezet 9.3.2 pontja)

A Tervezet csak a szisztematikus részhez javasol betűrendes mutatót (6. ábra). Ez a fajta mutató a vizsgált tezauszokból hiányzott. Feltehetően azért, mert csak akkor volna értelme az ilyen mutatónak, ha a szisztematikus rész valamiféle hierarchikus osztályozási rendszert képviselne – mint amilyen az ETO –, ha tehát a tezausz nem volna tezausz. Megtalálható volt viszont a tezauszokban olykor a lexikai egységek permutált mutatója [3, 4, 15, 24, 26], melyet a Tervezetben meg sem említettek.

Előfordultak még egyéb, speciális mutatók is. Így pl. az egyszerű betűrendes jegyzék (7. ábra), mely csak azt tünteti fel, hogy a lexikai egység deszkriptor vagy nemdeszkriptor [pl. 9, 10, 11 stb.]. Az ilyen jegyzékeknek elsősorban a tezausz karbantartásakor, új tezauszok szervezésekor, a szókészlet átnézésekor van jelentősége.

Különleges mutatót képvisel a *Növényvédőszer jegyzék*-ben [22] a növényvédőszer hatóanyagok szerkezeti képleteit és szabványos kémiai elnevezéseit tartalmazó mutató (8. ábra). Erre azért volt szükség, mert a tezauszba csak a hatóanyagok köznapi vagy kereskedelmi neve szerepelhetett (a szabványos kémiai megnevezés túlságosan hosszú és bonyolult), a kereséskor viszont sokszor nagyon pontos kémiai elnevezések alapján is tájékozódni kell. Ez a mutató valójában nemdeszkriptor–deszkriptor jegyzéknek is felfogható, mint amilyeneket a MEDLARS-ban például a „belépő szótárak” (entry vocabulary vagy lead-in-vocabulary) képviselnek. A Tervezetben az ilyen – s a gyakorlati életben egyáltalán nem rendhagyó – mutatókról vagy kiegészítő részekről nincs szó.

#### 4.6 Az előszó (a Tervezet 10.7.2 pontja)

A vizsgált tezauszok előszavai rendkívül eltérőek voltak. Előfordult az előszó teljes hiánya is [2, 19, 22, 23]. Az előszó megírását nem annyira a Tervezet javaslatai, mint a készítő tapasztalatai befolyásolják (így például az a tény, hogy a készítő előzőleg csak az olyan osztályozási rendszerekkel foglalkoztak, mint az

301	OPTICAL EQUIPMENT	35 mm CAMERAS	311
302	CAMERAS	CAMERAS	302
	RT: Photography	RT: Photography	824
303	MOVING PICTURE CAMERAS	CINE CAMERAS	304
	By medium	RT: Cinema	895
304	CINE CAMERAS	CINEMA	895
	RT: Cinema	RT: Cine cameras	304
305	UNDERWATER CINE CAMERAS	-----	
306	TELEVISION CAMERAS	MOVING PICTURE CAMERAS	303
	RT: Television	OPTICAL EQUIPMENT	301
307	STEREO CAMERAS	PHOTOGRAPHY	824
		RT: Cameras	302
		STEREO CAMERAS	308
		TELEVISION	897
		RT: Television cameras	306
		TELEVISION CAMERAS	306
		RT: Television	897
		UNDERWATER CINE CAMERAS	305; 317

6. ábra A Tervezetben javasolt mutató a tezauszok hierarchikusan rendezett részéhez



BETÜRENDES JEGYZÉK  
80.12.18.

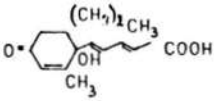
LAPSZ: 125

AZON	SZAK	FAZ	T	B	FOGALOMMEGNEVEZÉS
102669	314	D	7	N	GYOKERZET
123781	BA0	O	E		GYOKERZONA
123788	BFO	P	E	N	GYOKGUMOS TAKARMANY VAGASI ELLENALLASA
119574	I1	A	K		GYOKOS REAKCIO
031892	31G	D	4	O	GYOMBERGYOKER
031899	XB0	B	1	N	GYOMFELVETELEZES
123760	BGO	H	E	D	GYOMFESU
123767	BGO	B	E	D	GYCMIRTAS
125895	BGO	H	E	D	GYCMIRTOGEP
094080	1BA	L	6	D	GYCMIRTOSZER
102858	1BA	L	6	D	GYOMIRTOSZER HATASMECHANIZMUS SZERINT
031913	3BM	M	1	N	GYCMIRTOSZER OKOZTA KAROSODAS
031927	ZB	M	1	N	GYOMIRTOSZERERZEKENYSEG
031934	BAK	O	1	O	GYCMIRTOSZERES UGAR
031941	BB0	P	1	O	GYOMIRTOTT TERULET(HA)
031948	314	D	1	D	GYGMNAG
031955	XB0	B	1	N	GYOMMAGVIZSGALAT
035445	BB0	M	1	D	GYOMMENTESSEG
088165	310	D	1	O	GYOMNOVENY
031962	060	Y	1	N	GYOMNOVENYEK GAZDASAGI HATASA
031969	XB0	B	1	N	GYOMNOVENYSZAMLALAS
106974	OQ	V	7	O	GYOMNOVENY TAN
031976	1V1	L	6	D	GYCMORMEREG
031983	3AK	A	1	D	GYCMOSODAS
031990	3AK	Y	1	N	GYCMOSSAG KOVETKEZMENYE
031997	X08	B	1	N	GYCMPROGNOZIS
102676	BB0	B	7		GYOMTALANITAS
095452	BB0	M	1	N	GYGMTALANSAG
032004	31K	D	4	N	GYNGYAJAK
032011	31W	D	4	N	GYNGYIKE
032018	31K	D	4	N	GYNGYKOLES
032025	31W	D	4	O	GYNGYVIRAG
032032	31P	D	4	N	GYOPAR
032039	S7	O	1	O	GYOR
084532	C0	P	1	O	GYORSULAS
032046	31S	D	4	D	GYPSOPHYLA
032053	31S	D	4	N	GYPSOPHYLA AGRESTIS
032060	31S	D	4	D	GYPSOPHYLA MURALIS
102711	1B7	L	6	N	GYTTJA
123795	BFO	G	E	N	GYUJTOGEREBLYE
126728	LI2	I	F		GYUJTOMEDENCE
032067	31K	D	4	N	GYUJTOVANYFU
102452	ZB	B	1	D	GYULEKEZES(ALLAT)
032074	314	D	1	D	GYUMOLCS
097318	310	D	1	D	GYUMOLCSFA
124327	BFO	F	E	D	GYUMOLCSFA ALLVANY
032081	BBE	H	1	D	GYUMOLCSFA PERMETEZOGEP
032088	32D	D	5	N	GYUMOLCSFA SZUK
123802	BFO	H	E	N	GYUMOLCSFA VIBRATOR
125643	B10	H	E	D	GYUMOLCSFAAPOLO GEP
032102	1B5	L	6	N	GYUMOLCSFAQLAJ E

7. ábra Egyszerű betűrendes mutató a Tezauruszépítést Segítő Programrendszerrel (TSPR) kezelt növényvédelmi tezauruszból [13].

A mutató feltünteti a lexikai egység szakcsoportját (SZAK), fazettáját

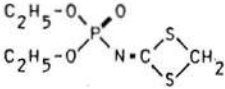
abszcizinsav



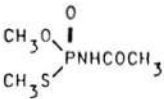
(6S)-5-(1-hidroxi-2,6,6-trimetil-4-oxo-2-ciklohexén-1-il)-3-metil-(Z,E)-2,4-pentadiénsav

AC 64475

O,O-dietil-(1,3-ditietán-2-illidén)-foszforamidát

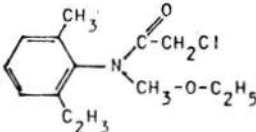


acefat



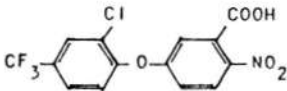
N-acetil-O,S-dimetil-foszforamidotioát

acetoklor



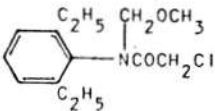
N-(2-etil-6-metil-fenil)-N-(2-etoxi-metil)-klór-acetamid

acifluorfen



5-[2-klór-4-(trifluor-metil)-fenoxi]-2-nitro-benzoésav

alaklor



N-(2,6-dietil-fenil)-N-(metoxi-metil)-klór-acetamid

8. ábra Részlet a Növényvédőszer teaurusz [22] hatóanyag-mutatójából

ETO). Statisztikai adatok többnyire csak olyankor szerepeltek, ha a teauruszt számítógépes segítséggel készítették. Ilyenkor azonban néha rendkívül részletes teaurusz-statisztikákkal is találkozni [9, 13, 24, 26].

Az igényesebb előszókhöz gyakran a szógyűjtéshez felhasznált források jegyzéke is kapcsolódott [9, 22, 26], s előfordult bennük részletes használati utasítás is.

4.7 A bejelentés (a Tervezet 10.8.3 pontja)

A vizsgált teauruszok közül 1982. végén 25 tekinthető teljesen késznek [1-től 24-ig, valamint 1T]. Ezek közül csak kettő készült el 1979. 12. 12. után. A

teauruszokat nyilvántartó nemzetközi gyűjtőhely, az Instytut Informacji Naukowej, Technicznej i Ekonomicznej (IINTE) nyilvántartása szerint ezzel szemben csak 10 magyarországi bejelentés szerepel [3, 5, 17, 8, 15, 14, 9, 10, 16], de közülük egyet [5] kétszer is bejelentettek.<sup>5</sup>

5. A készítés és a karbantartás szintje

5.1 A készítés (a Tervezet 10.1 pontja)

A készítés szervezési problémáival részletesen csak a Tervezetet megelőző 1974-es szabvány [III] korai Unesco változata [I], a DIN-szabvány [XI], valamint a külön a készítéssel foglalkozó NTMIR-dokumentumok [IX és X] foglalkoznak.

Az ISO-szabvány [III] 5.1 pontjában előírják, hogy egy-egy új teaurusz elkészítése előtt ellenőrzendő, nincs-e hasonló tárgy körben már kész teaurusz. A hazai teauruszok esetében azt tapasztaltuk, hogy az új teaurusz készítését nemcsak hogy nem befolyásolta a hasonló tárgy körű teaurusz léte, hanem az új teaurusz kidolgozásában sem használták fel szívesen a meglévő teauruszok adatait. Így például van két növényvédőszer teaurusz [19, 22] és három építésügyi [8, 12, 14]. A lexikai egységek összehasonlító vizsgálata azt mutatja, hogy a hasonló tárgy körű teauruszok szókincse inkább csak morfológiailag tér el egymástól, a rendelkezésre álló kifejezésekkel valójában majdnem mindig ugyanazok a fogalmak írhatók le. A feltűnő koordinációhiány egyik oka kétségtelenül a szakmai elzárkózás, aminek következtében meg sem ismerkednek alaposabban a szóba jöhető teauruszokkal. Mivel a gyakorlatban használt teauruszokban számos nyelvi megoldásnak az adott információkereső rendszer célkitűzéseivel összefüggő és a teauruszból magából első látásra nem felismerhető oka van, a felületes szemlélő hajlamos arra, hogy ezeket logikai-szemantikai hibának tekintse, ami jó ürügy arra, hogy tovább ne is mérlegelje az adaptáció kérdését. A másik ok azonban objektív: az eltérő fogalom szervezési szint (az összetett és az egyedi szavak alkalmazásának módja) nagyon függ a mindenkori adatbázis kezelési módjától. Más a fogalom szervezési szint fénylyukkártyák alkalmazása esetén [8] és más információkereső programrendszer használatkor [14]. De még programrendszerek között is alapvető különbségek vannak, attól függően, hogy az egyes rendszerekben hogyan kezelhetők az összetett kifejezések. Ezért más a fogalom szervezési szint az ISIS-szel, és más a GOLEM-mel kezelt teauruszokban [24, ill. 7].

<sup>5</sup> Bibliographic Bulletin of the Clearinghouse at IINTE. Warsaw. Institute for Scientific, Technical and Economic Information. Supplement.

Úgy tűnik, hogy az azonos vagy hasonló tárgykörben létező teauruszok kölcsönös felhasználhatóságának kérdése elsősorban a pre- és posztkombináció kérdésén múlik. A tapasztalatok szerint az eltérő tárgykörű, de azonos fogalomszervezési szinttel rendelkező teauruszok között könnyebben kerülhet sor lexikai egységek átvételére [pl. 9–13, 24, 25, 26], mint azonos tárgykörű, de eltérő fogalomszervezési szinttel rendelkező teauruszok között. Az is kétségtelen, hogy az átvétel kérdésében *döntő szerepe van* – automatizált információkereső rendszerek esetén – *a mindenkori programrendszernek*. Azonos programrendszer esetén még egyedi lexikai egységek átvétele is célszerű lehet. Éppen ezért a szabványosítás kérdésének kiemelkedő szerepe lehetne az információkereső programrendszerek megnevezés-rekordjainak egységes kialakításában és abban, hogy hogyan *kezeljék* az összetett kifejezéseket. Sajnos, ezekre a kérdésekre egyetlen szabály és irányelv sem tér ki, s a megnevezés-rekordok rögzítésének fontos kérdését is csak nagyon érintőlegesen és kizárólag a szellemi munkák szempontjából tárgyalják (a Tervezet 10.2 pontjában vagy az ISO-szabvány [III] 5.6 pontjában, vagy a DIN [XI] 7.6 pontjában).

#### 5.1.1 Az induktív és deduktív készítési módszer (a Tervezet 10.1.1 pontja)

A korábbi tervezetek – pl. [IV] – pragmatikus („analitikus”, induktív), illetve szisztematikus („Gestalt-”, deduktív) módszerről beszéltek.

A *pragmatikus módszer* esetén az összegyűjtött szavak alapján dolgozzák ki a teauruszt; ilyenkor tehát elsősorban a gyakorlatból vagy a tapasztalatból indulnak ki, és legfeljebb visszacsatolás formájában vonják be az elméleti rendezést.

A *szisztematikus módszer* esetén az elvileg elképzelhető és megalapozottnak tekintett fogalmi struktúrákból indulnak ki, és a gyakorlatban előforduló kifejezéseket mindenekelőtt ezekhez az alapstruktúrákhoz viszonyítva értékelik.

A két módszer többnyire összefonódik, s a gyakorlatban *átmeneti megoldások* alakulnak ki. Egyes teauruszokban azt tapasztaltuk, hogy a szisztematikus – és többnyelvű körülmények között deduktív – módszer alkalmazása mesterséges hierarchialáncokhoz, ETO-ra emlékeztető deskriptormegfogalmazásokhoz, elméleties és összetett kifejezésekhez vezetett [pl. 4, 21, 23].

## 5.2 A karbantartás

A karbantartás kérdésével a korábbi szabványtervezetek részletesebben foglalkoztak (pl. a nemzetközi szabvány [III] 6. pontja, a DIN-szabvány [XI] 8. pontja). Azt tapasztaltam, hogy a teauruszok készítésekor erre szinte

soha sem szoktak gondolni, mintha a készítéskor még az a meggyőződés uralkodnék, hogy az elkészült teaurusz további munkát már nem is igényel. Egyetlen teaurusz-csoporttal összefüggésben ismeretes csak részletes karbantartási szabályzat.<sup>6</sup> A Tervezet 10.5 pontja a karbantartás kérdésének csak egy részletét érinti, mégpedig elsősorban a szavak felvételét és törlését.

#### 5.2.1 A túl sokat, illetve túl keveset használt lexikai egységek (a Tervezet 10.5 pontja)

Csak egyetlen teauruszban tapasztaltuk, hogy az ilyen lexikai egységeket jelzéssel látták volna el [24] (pl. azzal, hogy „indexelésre nem használható”, illetve csak információkeresésre használható”). Az ilyen összefoglaló főlérendeltek ugyanis – automatizált rendszerekben – a hierarchiaszint emelését alkalmazó keresési stratégiában akkor is felhasználhatók, ha történetesen egyetlen tételt sem indexeltek velük. Mivel a szabványosítás a teauruszok használatával nem foglalkozik, erre a kérdésre sehol sem térnek ki. A tartalmi elemzéssel foglalkozó, az NDK-ban készült irányelvben [XV] sem térnek ki az ehhez hasonló kérdésekre. Előfordul, hogy az általánosabb jelentésű deskriptorhoz megjegyzést fűznek, hogy használatának gyakoriságát csökkentsék, mivel további alárendeltekkel nem rendelkezik, melyek elszívhatnák előle a dokumentumokat. Így például a bőripari teauruszban [25] a BŐRIPAR deskriptort – mivel nincsenek alárendeltjei – azzal a megjegyzéssel látták el, hogy „csak akkor használható, ha a dokumentum az egész iparágra vonatkozó szervezési kérdésekkel foglalkozik”. Ez is egy, a szabványokban nem említett módja a túl gyakori deskriptorhasználat csökkentésének.

## 6. A gépesítés szintje

A Tervezet az információkereső teaurusz egyik legfontosabb kritériumát abban látja, hogy nem automatizáltan, hanem *emberi közreműködéssel használják*. Eltekintve attól, hogy éppen ebben az egyre jelentősebb kérdésben kár a leszűkített értelmezés, az is sajnálatos, hogy a gépi – az információkereső programrendszerben integrált – teauruszokra vonatkozóan a tervezetben semmi sem szerepel, holott ezeket a teauruszokat – szemben az automatikus indexeléshez használtakkal – szintén emberi közreműködéssel használják indexelésre, csak éppen számos információkereső műveletet – például a szinonima-utalást, az asszociált fogalmak bevonását a keresésbe stb. – automatikusan lehet elvégeztetni. A hazai gépi teauruszok [7, 9, 10, 11, 13, 14, 17, 18, 19,

<sup>6</sup> A Kohó és gépipari teaurusz karbantartási irányelvei. (KG AIIR 8) Szerk. Ungváry R. Bp. KG INFORMATIK, 1978.

24, 29] tapasztalatai arról tanúskodnak, hogy célszerű volna például

- a valódi és a kváziszinonim összefüggéseken alapuló nemdeszkriptor–deszkriptor relációk megkülönböztetése, mivel az automatikus szinonima-utalásokhoz érdemben csak a kváziszinonímiát használják fel;
- figyelembe venni az összetett kifejezésekre vonatkozó előírásokban, hogy a gépi tezauszokban – ha a programrendszer a szintaktikai relációkat „ismeri” – egyáltalán nem okoz gondot az utólagos, indexeléskor végzett szóösszetétel (pl. az OLAJTRANSZFORMÁTOR illetve TRANSZFORMÁTOROLAJ esetén), mivel a szintaktikai relációval ezek a problémák egyértelműen megoldhatók (így például a GOLEM-ben [24]: TRANSZFORMÁTOR::OLAJ illetve OLAJ::TRANSZFORMÁTOR, ahol mindig a jelző következik a dupla kettőspont után);
- a hierarchialánc kialakításakor nagyobb mértékben figyelembe venni a vezérdeszkriptorhoz kapcsolódó többi deszkriptor jelölésének kérdését (és nem megjegyzéssel jelölni például az ún. elágazásokat, hanem a programrendszer által operátorként kezelhető reláció-jellel);
- figyelembe venni azt a tényt, hogy egyre több tezausz készül gépi segítséggel, s ezért célszerű volna szabályozni a megnevezés- és kapcsolattételek optimális szerkezetét, és az automatikusan ellenőrizhető hibák fajtáit.

## 7. A használat szintje

A tezauszok használata egyrészt szintaktikai kérdés, másrészt szorosan összefügg a gépesítéssel. Feltűnő, hogy a szakirodalomban elsősorban a készítéssel és a tezauszok szerkezetével foglalkoznak, a használatra alig szentelnek figyelmet. A nemzetközi szabványosítás se a gépesítéssel, se a szintaktikai kérdésekkel nem foglalkozott eddig érdemben. A tartalmi feltárással foglalkozó ISO/DIS 5963 szabványtervezet 1981 évi változatában [XIV] a szintaxis kérdését csak érintik: „... Egyes indexelő rendszerekben használnak szerepjelölőket, kapcsolatjelölőket, súlyozott deszkriptorokat stb. Az indexelőnek jól kell ismernie az összes ilyen, a rendszerben alkalmazható eszközt”. Ezenkívül utalnak az indexelés specifikusságának elvére: „... Ha tezauszt használunk, az adott fogalmat legspecifikusabban képviselő kifejezést válasszuk ki az indexelésre”. Az ennyire általános elveknek kevés befolyásuk van a gyakorlati munkára. A gyakorlattól való eltávolodás jele, hogy a szabványtervezetben egy sor olyan megállapítás is szerepel, amely a Tervezet megállapításainál is szűkebben értelmez bizonyos dolgokat. Az indexeléshez használható kifejezések-

ről például azt állítják, hogy ezek „... főnevek vagy főnévi kifejezések”. Ennek a magyarországi tezauszok többsége ellentmond (lásd a 2.2.3 fejezetet).

A nemzeti normatív dokumentumok közül csak az NDK-ban létezik olyan dokumentum, amely a tezausszal végzett indexeléssel foglalkozik [XV]. Ez kellő mértékben részletes, és figyelembe veszi a gyakorlati problémákat, Magyarországon azonban teljesen ismeretlen.

A vizsgált tezauszok készítését túlnyomó részben nem előzte meg a szabad tárgyszavakkal végzett, rendszeres indexelés, melynek során a jövődöbéli készítőik gyakorlati tapasztalatokat szerezhettek volna a természetes nyelven alapuló információkereső nyelvekkel végzett információfeltárással és kereséssel. A tezauszok többnyire alkalmazási tapasztalatok nélkül készültek, jórészt szakirodalmi forrásokra támaszkodva. A szintaktikai kérdések elhanyagolása valószínűleg ezzel is magyarázható. Nagyobb jelentőséget a felhasználásnak láthatóan mindig ott szenteltek, ahol a tezauszt automatizált információkereső rendszer számára készítették. A gépesítés igényei kikényszerítették az alkalmazási szempontok figyelembevételét, ami észrevehető a több szófajta tartalmazó, relációkban gazdagabb szerkezetben, és abban, hogy a gépi tezauszokkal összefüggésben tapasztalható utalások – mégha többnyire csak megjegyzésszerűen is. Egy – az ISIS-hez használt – gépi tezauszhoz társult részletesebb használati leírás [24].

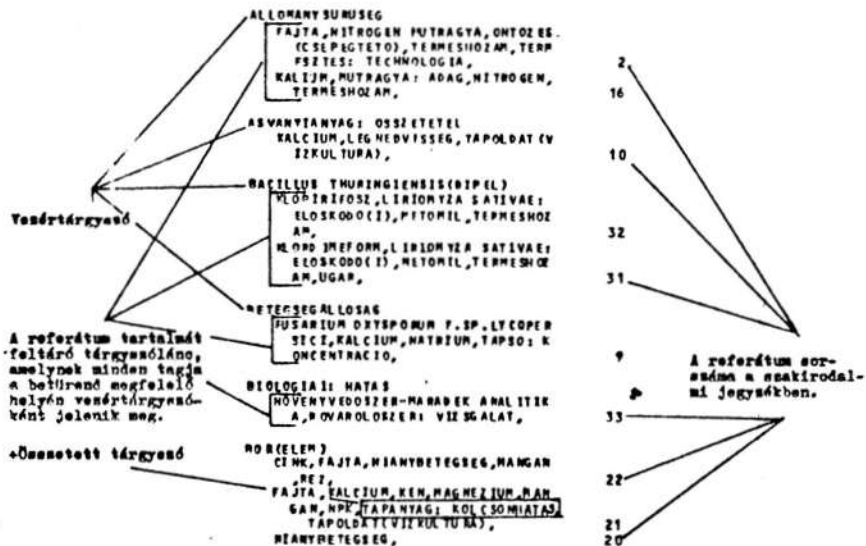
A szintaxisra vonatkozó megjegyzések gyakrabban fordulnak elő azokban a tájékoztató kiadványokban, melyekben a tezausszal végzett tartalmi feltárással eredményeként közreadják a bibliográfiai tételek jegyzékét. Ezekben a – többnyire automatizáltan készülő – kiadványokban szereplő megjegyzések áttételesen a tezausz használatára vonatkozóan is némi útbaigazítást adnak (9 ábra).

## 8. A két- és többnyelvű tezauszok

Késznek tekinthető két- és többnyelvű tezausz Magyarországon ma még csak egy van [1T]. A többi vagy torzóban maradt [2T, 4T], vagy jelenleg készül [3T, 5T, 6T]. Az alábbi következtetéseket ezért részben ezek alapján, részben pedig a nemzetközi tapasztalatok alapján vontam le.<sup>7</sup>

<sup>7</sup> UNGVÁRY R.–ORBÁN É.: hivatkozott mű.

**A PERMUTÁLT TÁRGYMUTATÓ HASZNÁLATA**



\* A tárgyszórendszert köztük jellege miatt Összetett tárgyszószavak is szerepelnek, köztük kettősponttal: pl. TÁPTRAGYKÖLCSÖNHATÁS. A KÖLCSÖNHATÁS túl általános fogalom, ezért önállóan nem, csak konkrétabb fogalomhoz kötve jelenik meg.

**TÁRGYMUTATÓ**

AGROTECHNIKAI VÉDEKEZÉS		ARARIS MOSAIC VIRUS	
HASTIDIOMICTES, TUNETAN,	33	GRAPEVINE FANLEAF VIRUS, PASPHERY RINGSFOT VIRUS, SZAFOPITOMYAG, SZEROFLOGIATÉLISA, VIRUSMINTESITÉS,	12
DIURON, GYOMNOVENY: TAPANYAGFELVÉTEL, PARAQUAT, SZIMAZIN, ANETTILLA OENOPHILA,	73 61		
AGROFARAB		ARISTOLOCHIA CLEMATITIS	
DUCP, FENAMIFOSZ, KARHOFURAN, MELU TOGYNE HAPLA, OXAMIL, PARATRICHODOPUS CHRISTIFI, XIPHINEMA INDEX,	42	AMINOTRIAZOL, CONVULVULUS ARVENSIS, CYNODON DACTYLON, DALAPON, DIKLORENIL, EQUISETUM, GLIFOSZAT, OXADIAZON, RUBUS, TERPHUMETON, ERBUTILLAZIN,	58
DIKLOPPROPAN, DIKLORFROPEN, SZULFOKARR, XIPHINEMA INDEX,	40		
ALUMINIUM ETILFOSZFÁT		APMILLARIFELLA MELLEAPOLYPORUS, STEREUM,	34
FOLPET, LS 74-783, PHOMOPSIS VITICOLA, PLASMOFARA VITICOLA, UNCI NULA NECATOR,	20	ATRAZIN	
FOLPET, PLASMOPARA VITICOLA,	21	AMINOTRIAZOL, LINURON, MONOLINURON, PARAQUAT, SZIMAZIN,	74
AMINOTRIAZOL		BACILLUS THURINGIENSIS	
ARISTOLOCHIA CLEMATITIS, CONVULVULUS ARVENSIS, CYNODON DACTYLON, DALAPON, DIKLORENIL, EQUISETUM, GLIFOSZAT, OXADIAZON, RUBUS, TERPHUMETON, ERBUTILLAZIN,	68	CLYSIA AMBIGUELLA, KEMIAI VEDEKEZÉS, LORFSIA POTRANA, CLYSIA AMBIGUELLA, LORFSIA BOTRANA, NOVENYVEDÉLMI FLOREJFLÉZESZEXFEROMON CSAPDA,	58 64
ATRAZIN, LINURON, MONOLINURON, PARAQUAT, SZIMAZIN,	74		
GLIFOSZAT, KLOKTIAMID, OXADIAZON, PARAQUAT, SZIMAZIN, TALAJMŰVELÉS NÉLKÜLI VÉDELIS, TERPHUMETON, TERBUTILLAZIN, 2, 4-D,	70	BARBAN	
		BOTRYTIS CINERIA, KARHOFURAN KLOPPROPAN, NOVENYVEDÉSERALLOSAG	37
ANTHIDIOMIKUM		BASIDIOMICTES	
MERCF BETEGSEG,	16	AGROTECHNIKAI VEDEKEZÉS, TUNETAN	35

9. ábra Részlet a Növényvédelmi és Agrokémiai Információ Visszakereső Rendszerben feldolgozott információtételek mutatójából.

A mutatóban szereplő tárgyszavak és tárgyszóláncok a Növényvédelmi és Agrokémiai Tezaurusz [13] deskriptoraiból állnak, melyeket szükség szerint szintaktikailag is összekapcsoltak.

### 8.1 A két- és többnyelvűség közötti különbség

Határozott különbség tapasztalható a két- és többnyelvű tezauszok között, amire a többnyelvű tezauszokkal foglalkozó szabványtervezetekben és irányelvekben [II, V, VIII, X, XIII] egyáltalán nem térnek ki. A vegyipari- és kőolajipari gépgyártási [2T] és a műszeripari [4T] többnyelvű tezauszok készítésekor például kiderült, hogy a bázis- vagy vezérnyelv szerepének sokkal nagyobb volt a jelentősége, amint át kellett térni a kétnyelvű környezetből a többnyelvű környezetbe. Míg kétnyelvű környezetben mindegy volt, hogy az adott nyelvek közül melyiket választják vezérnyelvnek, addig többnyelvű körülmények között rendkívül hátrányosan hatott, ha a vezérnyelv a résztvevő nyelvek közül nem a legfejlettebb volt, azaz nem rendelkezett a legnagyobb és az adott szakterületen a legbonyolultabb szókinccsel. A két- és többnyelvűség közötti különbségnek háttérben minden jel szerint az áll, hogy a többnyelvű tezauszoktól a közvetítőrendszer szerepe elválaszthatatlan, míg kétnyelvű tezausz esetén közvetítőnyelvre – legalábbis látszólag – nincs szükség.

### 8.2 A fordítókódok

Mivel a közvetítőrendszer szerepe ma még elméletileg tisztázatlan, a gyakorlatban a közvetítőrendszer leegyszerűsített helyettesítési formáit próbálják használni. Közöttük a legtöbb félreértést az ún. fordítókódok okozzák, melyeket egyes irányelvek – sajnos – ajánlanak is [VIII, X]. A Magyarországon készülő kétnyelvű tezauszok közül ezeket pl. az Agroiinform tezauszában használják [3T]. Ezek a fordítókódok valójában *csak* közvetítő elemek, s ezért nem automatizált körülmények között nincs értelmük; nem arra valók ugyanis, hogy velük indexeljenek (mint ahogy az univerzális grammatika mélystruktúrái vagy a rangnathani ideák sem arra valók, hogy közvetlenül használják őket). Alapvető tévedés tehát, ha a fordítókódok rendszerét tekintik az indexelés eszközének. Osztályozni az egyes nyelvek kifejezéseivel képviselt fogalmakkal kell, és az információkereséskor a nyelvközötti kompatibilitás mértékét kifejező relációjelek figyelembevételével kell a többnyelvű – tehát a *kevert* – keresőprofilot megszerkeszteni. Feltehetően éppen a két- és többnyelvű indexelés hosszabb gyakorlati tapasztalatainak a hiánya okozza, hogy a fordítókódok alkalmazásával majdnem minden rendszerben megpróbálkoznak.

Különösen abszurd következménye van a fordítókódok alkalmazásának, ha ezeket a kódokat ráadásul a hierarchikus osztályozási rendszerben szokásos módon kívánják megszerkeszteni, s ily módon ún. zárt tezauszokat alakítanak ki. Hogy az újabb lexikai egységeket később mégis felvehessék, ezeket a kódokat bizonyos

ugrásokkal alakítják ki, ami azonban a tezausz zártságán nem változtat, s az sem számítható ki előre, hogy a rendszer éppen hol fog túlszordulni. A szisztematikus felépítettségű fordítókódok [p. 3T] következtében éppen a tezausz-elv lényege – posztkoordinált jellegéből következő nyitottsága és fejleszthetősége – semmisül meg, mindaz tehát, ami a természetes nyelvi kiindulásból eleve következne, s aminek kihasználására a tezauszok megszülettek.

### 8.3 A nyelvek közötti relációk jelölése

A legvégzetesebb hatása annak a – normatív dokumentumokban is rögzített – törekvésnek van, hogy a nyelvek közötti ekvivalencia jelölését durván leegyszerűsítették, és arra törekszenek, hogy a pontatlan (átfedő) nyelvközötti ekvivalencia esetén is teljes ekvivalenciát tüntessenek fel. Ezáltal ugyanis hamis kép alakul ki a keresőben; a profil összeállításakor nem veheti a valóságnak megfelelő módon figyelembe a különféle nyelveken kifejezett fogalmi struktúrák közötti különbségeket.

Az Unesco legújabb irányelvtervezetében [V] ugyan felsorolják a nyelvek közötti ekvivalencia különféle lehetséges fajtáit (8.2.1 pont), de a relációk explicit jelölésekor már nem adnak megoldást arra vonatkozóan, hogy jelöljék őket egymástól megkülönböztetett formában. Ugyanakkor részletesen foglalkoznak az „egy a többhöz” reláció kérdésével, azaz azzal, hogy a kifejezésnek a másik nyelven együttesen több kifejezés felel meg (pl. SÜTÉS = BACKEN + BRATEN). Ez a reláció valójában a nyelvek közötti egész–rész relációk egy speciális esete. A magyarországi két- és többnyelvű tezauszok készítésekor éppen az volt a probléma, hogy nem lehetett magát a nyelvek közötti egész–rész relációt szabványosan jelölni, holott ezt az egész–rész összefüggést kell pontosan megkülönböztetni mind a teljes, mind pedig az átfedő vagy pontatlan ekvivalenciától. Különösen pedig a pontatlan ekvivalenciát nem volna szabad semmiképpen ugyanúgy jelölni, mint a részleges vagy a teljes ekvivalenciát.

### 8.4 A két- és többnyelvű tezausz készítése (lásd még a 8.2 fejezetet)

Feltűnő, hogy az irányelvtervezetben [V] a többnyelven párhuzamosan végzett tezauszkészítést szorgalmazták, a hazai gyakorlatban azonban a *fordítás dominál*. Ennek oka minden jel szerint az, hogy a párhuzamos készítéskor is fordítani kell, még ha ez a készítő előtt rejtve is marad, mivel öntudatlanul zajlik le, mintegy jelentéktelen szellemi részműveletként. Mikor a gyakorlati készítési munkára kerül sor, a fordítás rejtett ténye válik egyszerűen valósággá, és a munkák a fordítás irányába tolnak el.

Az a tény, hogy – a szabványtervezet [V] szerint – a párhuzamos eljárásakor kevesebb lexikai egység keletkezik, egyáltalában nem ennek az eljárásnak a szabvány készítői szerint vélt fölényét bizonyítja, hanem azt, hogy az eljárás szorgalmazói elsősorban a mesterséges nyelven alapuló, hierarchikus információkereső nyelvek hatása alatt állanak, és igyekeznek kiküszöbölni az igazán analitikus elemeket a feldolgozásból.

A fordítás uralkodó voltának mélyebb oka kétségtelenül a közvetítőrendszer szükségességében rejlik. Többnyelvű teauruszt nem lehet közvetítőrendszer nélkül készíteni. Mivel ennek a rendszernek igazi természetét még nem ismerjük eléggé, az egyszerűség kedvéért egy vezérnyelvet választanak a készítők. Amikor a teaurusz-készítési munka módszerének eldöntésére kerül sor, az illetékesek öntudatlanul is érzékelik a közvetítőrendszer szükségességét, s ösztönösen vezérnyelvet választanak, ami a módszer kérdését a fordítás javára többnyire eldönti.

### 8.5 A többnyelvű teaurusz megjelenítése

A többnyelvű teauruszokra vonatkozó ISO tervezetben [V] elsősorban a többhasábos megjelenítést javasolják. Ez a forma eléggé elterjedt [1T, 3T, 4T], bár vannak egyhasábos formák is, amelyen belül az idegennyelvű megfelelőt külön relációjellel jelölik [2T, 6T]. A tervezet [V] sajnos semmiféle javaslatot nem ad arra vonatkozóan, hogy a nyelvek közötti ekvivalenciák különböző fajtáit milyen relációjelekkel jelöljék. A többhasábos megjelenítés ennek a kérdésnek jelentőségét eleve elmosza, és azt a képzetet kelti, mintha minden a legnagyobb rendben volna. Valójában azonban mind az egyhasábos, kevert, mind a többhasábos megjelenítési formában meghatározott relációjelekkel kellene jelölni a nyelvek közötti ekvivalenciák fajtáját.

## 9. Következtetések

Az egyes nemzetközi és számos nemzeti szabványban és szabvány jellegű dokumentumban sok elvi kérdést megoldottak, s ezek alapján a teauruszok egységes megjelenése biztosítottnak látszik. Ugyanakkor ezek a dokumentumok – magyarországi teauruszok tapasztalatai alapján – nem vetnek eléggé számot a gyakorlattal, különösen, ami a gépesítés kérdését illeti. Annál érthetlenebb ez, ha meggondoljuk, hogy a teauruszok valójában az információkereső rendszerek gépesítésének köszönhetik elterjedésüket.

A hazai teauruszokban rugalmasabban és tágabban értelmezik az osztályozáshoz felhasználható nyelvi ele-

mekeket; valójában éppen a teauruszok lényegében analitikus jellegű felhasználását tartják szem előtt, amikor így járnak el. Korántsem korlátozzák például olyan mértékben az egyes szófajok alkalmazhatóságát, mint ahogy ezt a nemzetközi szabvány teszi; így például nem lehet magyar teauruszban lemondani arról, hogy a mellékneveket viszonylag gyakran önállóan, deskriptorként használják. A lexikai egységek között lényegesen több relációtípust különböztetnek meg, mint amennyit a nemzetközi és a hazai előírások megkövetelnek. Különösen feltűnő a nemdeskriptorok és a deskriptorok közötti ÉS- és VAGY-kapcsolatoknak a szerepe: ezt a két relációt nem ismerik a nemzetközi szabvány-dokumentumok. Ugyancsak több relációtípust különböztetnek meg az ún. rokonsági relációkon belül. Mivel a nemzetközi és hazai szabványokban és szabvány jellegű dokumentumokban lényegében nem vetnek számot azazal, hogy amint a teauruszokat automatizált információkereső rendszerekben használják, az alkalmazott programrendszer lehetőségei döntően befolyásolják a teaurusz szerkezetét, a legnagyobb – és elkerülhetetlen – eltérések a szabványoktól éppen az ilyen, gépi teauruszokban tapasztalhatók. Ezek a befolyások azonban nem olyan anarchikusak, hogy ne volnának megfelelő formában rendszerezhetőek, és ennek alapján ne lehetne rájuk tekintettel a gépesítésnek jobban megfelelő teauruszirányelveket, illetve szabványokat kidolgozni. A magyar teaurusz-szabvány kidolgozásakor – különösen pedig, ha a készítésre vonatkozó NTMIR normatív–műszaki előírások adaptációjára is sor kerül – fokozott mértékben figyelembe kellene venni a hazai teauruszkészítésben tapasztalt problémákat és igényeket.

## Bibliográfia

### 1. rész: A Magyarországon elkészült teauruszok\*

#### Egynyelvű teauruszok

- [1] Általános műszaki fogalmak teaurusza. I., II. és III. köt. Bp. Országos Műszaki Könyvtár és Dokumentációs Központ, 1970. 1113 p.
- [2] Földtani tárgyszógyűjtemény. Bp. Magyar Állami Földtani Intézet, 1971. 40 p.
- [3] Fogalomgyűjtemény az informatika válogatott területeiről. Összeállította: Várady É. Bp. Országos Műszaki Könyvtár és Dokumentációs Központ, 1972. 192 p.

\* A bibliográfia nem tart igényt a teljességre. Szerepel benne kész, illetve készülőfélben levő vagy félbemaradt teaurusz, továbbá olyan dokumentum, melyet kiadói teauruszként jelentettek be. Lezárva 1982. decemberében. A sorrendjük az időrendet követi.

- [4] Számítástechnikai fogalmak tezausz. Bp. Országos Műszaki Könyvtár és Dokumentációs Központ, 1974. 396 p.
- [5] „Munkaiügy. Munkatudomány” fogalomkör tezausz-modellje. = Molnár I.: Az információkezelés fogalomrendszere. Az információs szaktezausz fogalma, rendszere és felépítése. Bp. Közgazdasági és Jogi Könyvkiadó, 1973. p. 147–202.
- [6] A bányászati és a rokon tudományok fogalmi. Segéd-könyv az indexeléshez, szakrendi osztályozáshoz, továbbá a szakirodalom kutatásához. Szerk. Uhlmann A.–Uhlmann N. Bp. 1974. 223 p. (Bányászati Kutató Intézet Dokumentációs Osztályának kiadványa 1.) (A Nehézipari Műszaki Egyetem Központi Könyvtárának kiadványa 16.)
- [7] Számítástechnikai tezausz. Budapest, Számítástechnikai Koordinációs Intézet, 1974. (mágneses adathordozón)
- [8] Deszkriptorszótár a fénylukkartyás építéstervezési információkereső rendszerhez. I., II. és III. köt. Bp. Országos Műszaki Könyvtár és Dokumentációs Központ, 1974. 596 p.
- [9] Gépgyártástechnológiai szaktezausz. Szerk. Ungváry R. Bp. Kohó- és Gépipari Tudományos Műszaki Tájékoztató Intézet, 1975. 735 p. (KG AIIR 1. I.)
- [10] Vaskohászati szaktezausz. Szerk. Ungváry R. Bp. Kohó- és Gépipari Tudományos Műszaki Tájékoztató Intézet, 1975. 608 p. (KG AIIR 1. II.)
- [11] Erősáramú elektrotechnikai szaktezausz. Szerk. Ungváry R. Bp. Kohó- és Gépipari Tudományos Műszaki Tájékoztató Intézet, 1975. 470 p. (KG AIIR 1. III.)
- [12] Az Építéstudományi Intézet tematikai információs rendszerének tezausz. Bp. Építéstudományi Intézet, 1975. 300 p.
- [13] Növényvédelmi deszkriptorszótár. Szerk. Somi Kovács Mária–Ungváry R. Bp. MÉM Növényvédelmi és Agrokémiai Központ, 1975. (mágneses adathordozón) (NÖMTIR III. 1.)
- [14] Építésügyi tezausz. I., II. és III. köt. Bp. Építésügyi Tájékoztatói Központ, 1977. (lapsz. n.)
- [15] Könyvtári és tájékoztatói tárgyszójegyzék. Tezausz tervezet. Bp. OSZK Könyvtártudományi és Módszertani Központ, 1976. 187 p.
- [16] Fogalomgyűjtemény az anyagmozgatás, csomagolás, raktározás és a disztribúció témaköreiből. Bp. Anyagmozgatói és Csomagolási Intézet, 1977. (előkészületben)
- [17] Gépipari szaktezausz. A kohó- és gépipar „Gépek és gépi berendezések” szakágazatának tezausz. Szerk. Ungváry R. Bp. Kohó- és Gépipari Tudományos Műszaki Tájékoztató Intézet, 1977. (mágneses adathordozón) (KG AIIR 1. V.)
- [18] Járműipari szaktezausz. Szerk. Ungváry R. Bp. Kohó- és Gépipari Tudományos Műszaki Tájékoztató Intézet, 1977. (mágneses adathordozón) (KG AIIR 1. VI.)
- [19] Növényvédőszer deszkriptorszótár. Veszprém, Nehézipari Kutató Intézet Növényvédelmi Főosztály, 1977. 27 p.
- [20] Az Unesco ideiglenes deszkriptorjegyzékének magyar fordítása. Szerk. A Népművelési Intézet Szakkönyvtára Tezauszbizottsága. OSZK Könyvtártudományi és Módszertani Központ, Bp. 1977. 109 p.
- [21] Muzeológiai szakrend és szójegyzék. Bp. Múzeumi Restaurátor és Módszertani Központ, 1979. 183 p.
- [22] Növényvédőszer jegyzék. Összeállította Vadász Á. Bp. MÉM Növényvédelmi és Agrokémiai Központ, 1979. 214 p.
- [23] K+F irányítási információs tezausz. I. és II. köt. Bp. Könyvüipari Szervezési Intézet, 1980. 119 p. (Kutató-fejlesztő tevékenység és „MÜFA” gazdálkodás irányításának központi számítógépre szervezett információs rendszere).
- [24] Számalk-tezausz. Szerk. Ungváry R.–Raffalszky K. Bp. Számítástechnika-alkalmazási Vállalat, 1982. 872 p.
- [25] Bőr-, szőrme-, műbőr- és cipőipari tezausz. Szerk. Ungváry R. Bp. Bőr-, Műbőr- és Szőrmeipari Kutató Intézet, 1982–1983. (mágneses adathordozón) (előkészületben)
- [26] Formatervezési tezausz. Szerk. Ungváry R. Bp. Formatervezési Központ, Országos Műszaki Információs Központ, 1982–1983. (mágneses adathordozón) (előkészületben)
- [27] Kartográfiai tezausz. Bp. Kartográfiai Vállalat, 1982–1984. (előkészületben)
- [28] A KPM K+F célprogramokat alátámasztó információ keresőrendszerének fogalomgyűjteménye (tezausz). Szerk. Horváth T. [et al.] Bp. Közlekedéstudományi Intézet, 1982.
- [29] Hadtudományi tezausz. Szerk. Bakos Klára–Tóth Bálintné. Bp. Zrínyi Katonai Akadémia, 1980–1983. (mágneses adathordozón) (előkészületben)
- [30] Magyar külgazdasági tezausz. I. és II. rész. Szerk. Kónya B. Bp. Magyar Kereskedelmi Kamara (Okmányegységszabványügyi és Adatfeldolgozási Koordinációs tagozata – HUNGPRO –), 1982. 588 p.

*Két- és többnyelvű tezauszok*

- [1T] Nemzetközi Építésügyi Tezausz. KGST ÉÁB, 1976. (kézirat)
- [2T] A vegyipari és kőolajipari gépgyártás orosz–magyar tezausz. Bp. KG INFORMATIK, 1980. (mágneses adathordozón)
- [3T] AGROINFORM-tezausz. Bp. Agroinform, 1981. (kézirat)
- [4T] Műszeripari orosz–magyar tezausz. Bp. KG INFORMATIK, 1980. (mágneses adathordozón)
- [5T] Kartográfiai magyar–orosz tezausz. Bp. Kartográfiai Vállalat, 1982. (kézirat)
- [6T] INFORMELEKTRO-tezausz. Orosz–magyar fordítás. Bp. KG INFORMATIK, 1980. (kézirat)

**2. rész: Az összehasonlításba bevont nemzetközi és nemzeti szabványok és szabvány jellegű dokumentumok**

- [I] Guidelines for the establishment and development of monolingual scientific and technical thesauri for information retrieval. Paris, Unesco, 1970. 14 p. (SC/MD/20)



- [II] Guidelines for the establishment and development of multilingual scientific and technical thesauri for information retrieval. Paris, Unesco, 1971. 20 p. (SC/WS/501)
- [III] ISO 2788-1974 Documentation. Guidelines for the establishment and development of monolingual thesauri. First edition.
- [IV] ISO/DIS 2788 Guidelines for the establishment and development of monolingual thesauri. 2nd edition. Provisional version of second edition. 1981. (TC 46/SC 5 No. 42 E)
- [V] UNISIST Guidelines for the establishment and development of multilingual thesauri. Revised Text. By Austin, D.-Waters, J. Paris, May 1980, Unesco. (ISO/TC 46/SC 5)
- [VI] ISO/DIS 5127/3 b Information and documentation. Vocabulary. Section 3 b: Documentary languages. 1980. (TC 46/SC 3)
- [VII] ST SEV 174-75 Egynyelvű információkereső tezausz. Szerkezet, összetétel és megjelenési forma. Az NTMIR normatív-műszaki dokumentumai, 2. I. rész. p. 45-50.
- [VIII] NTP MCNTI 3-75 Többnyelvű információkereső tezausz. Tartalom, szerkezet és megjelenési forma. Az NTMIR normatív-műszaki dokumentumai, 2. I. rész. p. 41-44.
- [IX] NTP MCNTI 9-75 Az egynyelvű információkereső tezausz kidolgozásának szabályai. Az NTMIR normatív-műszaki dokumentumai, 2. I. rész. p. 72-83.
- [X] NTP MCNTI 15-76 A többnyelvű információkereső tezauszok kidolgozásának szabályai. Az NTMIR normatív-műszaki dokumentumai, 6. II. rész. p.
- [XI] DIN 1463 Richtlinien für die Erstellung und Weiterentwicklung von Thesauri. März 1976.
- [XII] Methodische Rahmenregelung zur Erarbeitung von Thesauren. Berlin, Zentralinstitut für Information und Dokumentation, 1973. 58 p.
- [XIII] Methodische Empfehlungen für die Erarbeitung von russisch-deutschsprachigen Informationsrecherchesauren. Berlin, Zentralinstitut für Information und Dokumentation, 1978. 92 p.
- [XIV] ISO/DIS 5963 Documentation. Methods for examining documents, determining their subjects, and selecting indexing terms. 1981. (TC 46/WG 5)
- [XV] Methodische Rahmenregelung zum Indexieren mit Hilfe eines Thesaurus. Berlin, Zentralinstitut für Information und Dokumentation, 1973. 32 p.
- [XVI] Methodische Hinweise für die Übersetzung fremdsprachig vorliegender Thesauri in die deutsche Sprache. Zusammenge stellt von Beling, G.-Schuck, J. H.-Wersig, G. Frankfurt M. Deutsche Gesellschaft für Dokumentation e. V., 1972. 83 p. /DGD-Schrift (KTF-2/KTS-1) 9/72/
- [XVII] MSZ 3418-76 Magyar nyelvű információkereső tezauszok szerkezete, részei, formája. Bp. Magyar Szabványügyi Hivatal, 1977. 7 p.

*UNGVÁRY Rudolf: A Magyarországon készülő tezauszok a nemzetközi szabványosítás tükrében*

Az elkészült és készüléfélben levő egy és többnyelvű tezauszok tulajdonságait összehasonlítottuk a nemzetközi szabványok és irányelvek követelményeivel. Külső megjelenésükben a Magyarországon készülő tezauszok többnyire nem térnek el a nemzetközi előírásoktól, tartalmilag azonban néhol jelentősek e különbségek. Szófajok tekintetében tapasztalhatók a legkisebb megköttöttségek: a magyar tezauszokban többnyire nem mondanak le arról, hogy a mellékneveket viszonylag gyakran használják önállóan, deskriptorként. A lexikai egységek között több relációtípust különböztetnek meg a szabványban ajánlottaknál. Különösen feltűnő a nendeszkriptorok és deskriptorok közötti ÉS- és VAGY-kapcsolatok szerepe; ez a két reláció teljesen hiányzik a nemzetközi dokumentumokból. A legnagyobb eltérések a szabványoktól és irányelvektől a gépesített bibliográfiai adatbázisokhoz használt tezauszokban tapasztalhatók. Az alkalmazott programrendszer lehetőségei döntően befolyásolják a tezausz szerkezetét. A szabványosítás mindeddig nem volt tekintettel erre a tényre. A gépesítés befolyása azonban korántsem olyan anarchikus, hogy a szabványosítás ne vehetné ezt a hatást tekintetbe.

\* \* \*

*UNGVÁRY, R.: The compliance of Hungarian thesauri with international standards*

The characteristics of mono- and multilingual thesauri including those under preparation have been compared to the requirements of international standards and guidelines. Their formats comply more or less with the international rules, however, in their contents considerable differences have been observed. The constraints concerning the parts of speech are minimum: adjectives are quite often used as individual descriptors. More relations between lexical units are differentiated than recommended by the standards. Particularly striking is the role of AND- and OR relations between non-descriptors and descriptors; the international documents do not refer to these relations at all. The most significant deviations from the standards and guidelines exist in thesauri used for machine-readable bibliographic databases: the computer programs affect considerably the structure of the thesauri. Standardization activities have neglected this fact, though the influence of automation should have a more pronounced role in standardization than the present practices.

\* \* \*

УНГВАРИ, Р.: Разработка тезаурусов ВНР в свете международной стандартизации

Проводилось сравнение особенностей разработанных и разрабатываемых одноязычных и многоязычных тезаурусов с требованиями международных стандартов. Формально — по внешнему оформлению — тезаурусы, разработанные в ВНР, не отличаются от международных требований, однако по содержанию в отдельных местах есть значительные отличия. С точки зрения частей речи требования менее строгие: в венгерских тезаурусах довольно часто встречается то, что имена прилагательные применяются самостоятельно, в качестве дескрипторов. Среди лексических единиц различают больше типов реляций, чем это предложено по стандарту. Особенно заметна роль соединений недескрипторов и дескрипторов при помощи отношений „И” и „ИЛИ”, этих двух реляций вообще нет в международных стандартах. Наибольшие отклонения от стандартов существуют в тезаурусах, применяемых в автоматизированных библиографических базах данных. Возможности применяемой системы программ оказывают воздействие на структуру тезауруса. При стандартизации до сих пор на это не обращалось внимание, несмотря на то, что автоматизация не является настолько анархичной, чтобы этот факт не учитывался.

UNGVÁRY, R.: Die in Ungarn erstellten Thesauri im Spiegel der internationalen Normen

Die vorhandenen bzw. in der Entwicklung befindlichen ein- und mehrsprachigen Thesauri wurden mit den Erfordernissen der internationalen Normen und Richtlinien verglichen. In ihrer äusseren Erscheinung entsprechen die untersuchten Thesauri den internationalen Vorschriften, inhaltlich sind jedoch bedeutende Abweichungen vorhanden. Die Eigenschaftswörter werden verhältnismässig oft als Deskriptoren verwendet und auch andere — durch die geltende Normen nicht akzeptierte — Wortarten kommen vor. Zwischen den lexikalischen Einheiten werden mehr Relationstypen definiert, wie dies in den Normen üblich ist. Besonders auffallend ist das Hervorkommen der Relation UND und ODER zwischen den Nichtdeskriptoren und Deskriptoren; sie fehlen völlig unter den Empfehlungen der internationalen Normen. Die grössten Abweichungen von den Normvorschriften sind unter den Thesauri zu bemerken, die in automatisierten Informationsrecherchesystemen verwendet werden. Die Möglichkeiten der gegebenen Programmsysteme beeinflussen entscheidend die Struktur der Thesauri. Die internationale Standardisierung widmet dieser Tatsache zurzeit wenig Augenmerk, obwohl der Einfluss der Automatisierung immer wie wichtiger wird. Dieser Einfluss ist genügend definierbar, um bei der Standardisierung regulieren zu können.

## Külföldi hírek

Az amerikai *Advanced Technology/Libraries* c. híradó által a lap előfizetői között végzett felmérés szerint a *mikroszámítógépek könyvtári alkalmazásában a közművelődési könyvtárak vezetnek*, megelőzve a felsőoktatási és szakkönyvtárakat. A megkérdezett közművelődési könyvtárak csaknem kétharmada rendelkezik legalább egy mikroszámítógéppel, míg ugyanezt a felsőoktatási könyvtáraknak 47%-a, a szakkönyvtáraknak pedig csak 28%-a mondhatja el magáról. A könyvtár típusok szerinti különbségek ellenére a mikroszámítógépek jövője megalapozottnak látszik: az összes megkérdezett könyvtár kétharmada vagy már rendelkezik mikroszámítógéppel, vagy rövid időn belül tervezi ezen (az egyik könyvtáros szerint) „elkerülhetetlen” berendezés beszerzését.

A *British Library Tanácsa* (British Library Board) úgy döntött, hogy *nem bocsátják az OCLC rendelkezésére* a British Library MARC adatbázisának mágnesszalagjait. A Tanács döntését azzal indokolta, hogy a rekordokhoz az OCLC által biztosított nemzetközi hozzáférés meghiusítaná az Egyesült Királyságban létrehozandó online könyvtári adatbázis-rendszer tervét, amely szerint az 1984 végéig üzembe helyezendő rendszer központi adatbázisa kezdetben a British Library és a résztvevő brit könyvtárak több, mint 4 millió rekordját fogja tartalmazni.

(*Advanced Technology/Libraries*, 1983 március)