

# Műhely

Az Országos Tudományos Diákköri Konferencia  
díjnyertes pályaművei (2. rész)



*Folytatjuk az előző számban elkezdett díjnyertes OTDK-pályaművek bemutatását. Ezúttal két második helyezést elért pályamunka, Dula Marina és Girhiny Eszter dolgozatának szerkesztett változatát mutatjuk be. (A szerk.)*

## A többnyelvű teauruszok kialakítási lehetőségei

DULA Marina

A 21. században a többnyelvű eszközök egyre nagyobb teret nyernek és egyre szélesebb körű felhasználói csoporthoz jutnak el, ezen belül olyanokhoz, akik eltérő kulturális és nyelvi háttérrel rendelkeznek, és az ismeretek hasonlóan változatos területei iránt érdeklődnek. Tény, hogy a teauruszalapú keresőrendszerek rendelkeznek a legjobb pontossági aránnyal, jelenleg azonban nincs a többnyelvű teauruszok kialakítására vonatkozó magyar szabványunk. Így alapvető hiányt pótolunk azzal, ha vizsgáljuk ezen eszközök létrehozásának lehetőségeit, a klasszikus és modern módszerek előnyeit és hátrányait.

A fenti indíttatás tükrében a szakirodalom tanulmányozásával két kérdésre kerestem a választ:

- 1) milyen tulajdonságok jellemeznék egy jól kivitelezett többnyelvű teauruszt,
- 2) hogyan lehet kialakítani egy ilyen rendszer hatékony és felhasználóbarát verzióját.<sup>1</sup>

### 1. A többnyelvű teauruszok kialakításának kezdeti lépései

A többnyelvű teauruszok kialakításának kezdeti lépései az egynyelvű változatokéhoz nagyban hasonlóak. Az előkészítő munka során meg kell határozni a célokat, a teaurusz fő témaköreit és rögzíteni kell a továbbfejleszthetőség lehetőségeit. A következő lépés az induló szóanyag gyűjtése, majd a szókészlet elemzésének során a szakemberek kiválasztják a legfontosabb deskriptorokat, illetve elkészítik a betűrendes mutatót.<sup>2</sup> A többnyelvű szótárak esetében a munkálatok kiegészülnek a célstruktúra megválasztásának feladatával.

#### 1.1. A nem-szimmetrikus célstruktúra

Az egyik alternatíva egy nem azonos, nem-szimmetrikus teauruszszerkezet kialakítása, ahol a

deszkriptorok száma és a kapcsolódás típusai nem szükségszerűen egyeznek meg az egyes nyelveken. Ebben a szerkezetben forrásnyelvi deszkriptor állhat önállóan egyenértékű célnyelvi kifejezés nélkül. *Michael Hudon*<sup>3</sup> szerint helyesebb, ha elfogadjuk ezt a nem-szimmetrikus struktúrát és a deszkriptorok számát változtathatóvá tesszük az egyes nyelvi változatokban. Indoklásának alapja az, hogy adott kultúrában létező fogalomnak, amely egy bizonyos nyelven képviselteti magát, egyáltalán nem biztos, hogy létezik ekvivalense egy másik kultúrában és az sem, hogy a megfelelő verbális alak áll majd rendelkezésünkre. A gyakorlatban, tehát az indexelő szótár kialakítása során nem lehet megegyező szerkezetet kialakítani, mivel gyakran előfordul, hogy két csúcshoz deszkriptor vagy főrendelt kifejezés, pontatlan vagy részleges nyelvi megfelelést mutat – vagyis, bizonyos mértékben eltérő jelentésűek –, aminek következtében eltérő lesz az alárendelt kifejezés is. Erre példa az angol „*education*” kifejezés, amelynek jelentése tágabb, mint a francia „*éducation*” kifejezése, vagy másképpen kifejezve az „*education*” és a „*teaching*” fogalmak világosabban különülnek el az angolban, mint az „*éducation*” és az „*enseignement*” francia kifejezések (1. táblázat). Hudon<sup>4</sup> véleménye szerint azok a többnyelvű tezauszok, amelyekben a szerkezet nem egyforma, nagyobb valószínűség szerint tükrözik hitelesen a világegyetem fogalmainak képviselését minden kultúrában és nyelven.

Angol (forrásnyelv)	Francia (célnyelv)
education	éducation enseignement formation
teaching	enseignement
instruction	enseignement instruction

1. táblázat

*A relációs hálózat bonyolultsága (Hudon, 2001):*

## 1.2. A szimmetrikus célstruktúra

A leggyakoribb nézet azonban az, hogy a többnyelvű tezausz minden egyes nyelvi változatának azonosnak kell lennie. Amennyiben adott tezauszban a deszkriptoroknak minden nyelven kizárólag egy egyenértékű megfelelője van, valamint az egyes deszkriptorok kapcsolódási típusai megegyeznek, szimmetrikus tezauszról beszélünk. Ez a struktúra annyiban rugalmas, hogy a deszkriptorok közötti

kapcsolat alapulhat teljes vagy részleges egyenértékűségen, illetve a nem-deszkriptorok száma eltérő lehet. A szimmetrikus tezausz kialakítóinak célja a tökéletes egyezés létrehozása a deszkriptorok és a nem-deszkriptorok között, valamint az egyes nyelvi ekvivalensek között. Ennek következményeként azonban gyakran előfordul, hogy a nyelvek közötti egyenértékűséget kényszeríteni kell, még ott is, ahol a valóságban nincsenek meg ezek az egyezések, így megkérdőjelezhető kapcsolati szerkezetek alakulhatnak ki.

A szimmetrikus tezauszban a kifejezések jelentését gyakran szűkíteni, vagy tágítani kell ahhoz képest, amilyen az adott kifejezés a természetes nyelvi használatban, de még így sem lehet elkerülni a mesterségesen „megalkotott” kifejezéseket, neologizmusokat. Ennél a szerkezetnél elfogadhatatlan, ha nincs ekvivalens a célnyelven, így ebben az esetben a tezausz alkotói általában eltávolítják a rendszerből a „gazdátlan” forrásnyelvi kifejezéseket.

## 2. Többnyelvű tezausz kialakítása alulról-felfelé építkezve

Az IFLA irányelvében<sup>5</sup> a többnyelvű tezauszok három kialakítási módszerét határozzák meg. Az első lehetőség, hogy alulról-felfelé építkezve alakítsunk ki egy tezauszt. Ezen belül további két variáció közül választhatunk: kezdetjük a munkát egy nyelvvvel, majd hozzá adhatunk egy másik nyelvet vagy nyelveket, illetve a másik opció, ha egyszerre több nyelvvvel kezdjük meg a struktúra kialakítását.

A szakirodalom a hetvenes évek közepétől azt hirdeti, hogy ez a párhuzamos (ab initio) tezauszkészítési módszer a legjobb. Ezek a tezauszok nemzetközi intézményeknél és világcégeknél készülnek, mivel nehéz és költséges feladat hosszú időn át olyan szakértői csoportot összefogni és együtt tartani, amelyek tagjai értenek mind az adott szakterülethez, mind az adott (idegen és anya)nyelvhöz, valamint a dokumentáció szakterületéhez is. Erre a módszerre jellemző a legtakarékosabb szógyűjtés, aminek következtében kevesebb homonima és pontatlan kifejezés is keletkezik, ami nagy pozitívum. Különböző nyelvi változatok együttes fejlesztése egyenértékűséget biztosít minden nyelven, azonban a célnyelvek és a forrásnyelvek váltakozásával, gyakran mesterkélt kifejezések keletkeznek. Ennek orvoslására a kulturális sajátosságokból adódó különbözőségeket az alkotók úgy kezelik, hogy a kifejezéseket összehangolják, amit a *terminológia harmonizációjának* neveznek.<sup>6</sup>

## 2.1. A tulajdonnevek megjelenítése a párhuzamos szerkezetben

Minden egyes deszkriptor, ami a tezauszban szerepel, egyedi fogalmat jelenít meg. A deszkriptor lehet egyszavas kifejezés vagy többszavas összetett kifejezés is, a lényeg az, hogy természetes nyelvi formájú legyen.<sup>7</sup> Abban az esetben, ha adott deszkriptorunk tulajdonnév a tazuruszban, a következő három alkalmazási lehetőség közül választhatunk:

- 1) kiemelhetjük a tulajdonnév egyik nyelvi változatát, és azt használhatjuk minden nyelven;
- 2) használhatunk egy tulajdonnevet az adott nyelveknek megfelelő alakban (abban a nyelvi környezetben viszont, ahol a kifejezés hiányzik, ott az eredeti nyelvű tulajdonnevet alkalmazzuk);
- 3) alkalmazhatunk fordításokat, amennyiben ez észszerű és lehetséges.

Minden esetben a kiválasztott forma megfelelő nem-deszkriptorokkal egészül ki, amelyekhez ajánlatos az egyes nyelvek engedélyezett neveinek listáját használni (ld. 2. táblázat). Az angol és a francia deszkriptorok az adott nyelvnek megfelelő hivatalos formában szerepelnek, a harmadik oszlopban azonban a holland deszkriptor kiegészül az angol és a francia kifejezésekkel, amelyek nem-deszkriptorként szerepelnek.

## 2.2. A homonimák kérdése

A homonimák (azonos alakú szavak) nemcsak az egyes nyelveken belül, hanem a nyelvek között is előfordulhatnak. Így ha ilyen kifejezéseket kell alkalmaznunk a többnyelvű környezetben, célszerű magyarázattal kiegészíteni őket, vagy a jelentések értelmét tisztázni egy zárójeles minősítő hozzáadásával.

Példa a minősítők alkalmazására a „Cranes” angol kifejezés, amely a magyar nyelvben is teljesen egyező homonimapárt alkot, mivel a „Cranes” (birds) vagy a „Cranes” (*lifting equipment*) magyarul „Daruk” (madarak) és a „Daruk” (emelő berendezések) teljes egyenértékű egymásnak. Azonban a holland nyelvben az itt megfigyeltetett kifejezéspár már nem mutat ilyen alaki hasonlóságot, ezért szükséges a magyarázat hozzákapcsolása (3. táblázat).

A tezausz megjelenítésének vannak olyan lehetőségei, amikor nem tüntetik fel a megjegyzéseket, ahogy ez jellemző az alfabetikusan rendezett soroknál. Azonban a szerkezet átláthatóságát tekintve jóval előnyösebb, ha az egyes nyelvek kifejezései külön listában kapnak helyet. Amennyiben mégis az előbbi opciót használjuk, a kifejezés mellett zárójelben elhelyezett nyelvmegjelölés alkalmazása ajánlott. A 4. táblázat felső részében a nyelvek közötti ekviva-

Angol	Francia	Holland
International Committee of the Blue Shield (ICBS)	Comité international du Bouclier bleu (CIBB)	Internationaal Comitee van het Blauwe Schild UF International Committee of the Blue Shield (ICBS) UF Comité international du Bouclier bleu (CIBB)

2. táblázat

Tulajdonnevek a többnyelvű tezauszban (IFLA, 2009)

Brit angol	Amerikai angol	Holland	Francia	Magyar
Cranes (birds)	Cranes (birds)	Kraanvogels	Grue (oiseau)	Daruk (madarak)
Cranes (lifting equipment)	Cranes (lifting equipment)	Hijskranen SN voor andere typen kranen, zie aldaar	Grue (appareil de levage)	Daruk (emelő berendezések)

3. táblázat

Homonimák a többnyelvű tezauszban, megjegyzéssel kiegészítve (IFLA, 2009)

lencia esetét és párhuzamait, míg az alsó cellában a kifejezések betűrendes megjelenítését láthatjuk.

Chance (angol) = Aasard (francia) / jelentésük: esély
Fortune (angol) = Chance (francia) / jelentésük: véletlen
Alfabetikus sorrend:
Chance (eng)
Chance (fre)
Fortune (eng)
Hasard (fre)

#### 4. táblázat

*Homonimák az alfabetikus többnyelvű megjelenítésben (s.k.)*

### 2.3. Az összetett kifejezések kérdése

Az összetett kifejezés több mint egy szóból vagy összetett szóból áll. Az összetett szavak olyan szavak, amelyek több mint egy szóból (általában főnévből) állnak, illetve egy vagy több betűből álló kötőszó kapcsolhatja össze őket. Az összetett kifejezések általában prekoordináltak, vagyis előzetesen egyesítenek két vagy több egyszerű fogalmat egy összetett fogalommal. Ezen kifejezések használata a tezauszban és a tárgyszólistákban növeli a specifikusság szintjét. A specifikussággal párhuzamban nő a pontosság is, ami az adott fogalom precíz képviselését eredményezi. Ezzel egyetemben pedig nő az eszköz hatékonysága is, mivel a rendszer a nem kívánatos, tárgykörbe nem tartozó dokumentumokat ki tudja zárni. Éppen ez a hiányosság jelentett problémát *Mortimer Taube*, *Unitermjénél* is, mivel a rendszer nem volt képes összetett kifejezésekkel jellemezni az adott tartalmat, így bizonyos speciális jelentések kiszorultak vagy torzultak a struktúrában belül.

Az összetett kifejezések egyes elemeit a következő elvek szerint lehet megkülönböztetni. A középpont, vagyis a főnév az az összetevő, amely azonosítja a fogalmak általános osztályát, amelyre a kifejezés egésze vonatkozik. A „*Printed indexes*” kifejezés főnévi komponense az „*indexes*”. A „*Hospitals for children*” kifejezésben a „*hospitals*” ugyanezt a szerepet tölti be, míg a „*Broadband*” esetében a főnévi

rész a „*band*”. Az összetett kifejezések másik része, a minősítő, amely az az egy vagy több komponens, ami arra szolgál, hogy szűkítse a kiterjeszhetőség középpontját, és így megadja az alosztályokat. Ilyen például a „*Printed*” melléknév a „*Printed indexes*” kifejezésben. Az előjáró-kötőszó-főnév összetételben, vagyis a „*Hospitals for children*” kifejezésben a „*for children*”, míg a „*Broadband*” kifejezésben a „*Broad*” tölti be ezt a szerepet. Általában, ha egy összetett kifejezés két vagy több kifejezésből áll, valószínűleg a vele egyenértékű idegen nyelvű kifejezés is ilyen formában lesz fellelhető.

A különböző nyelvű deskriptorok formája a gyakorlatban azonban nem mindig egyezik, mert pl. az angol „*Rail safety*” kifejezés holland megfelelője, az előjárószavas „*Veiligheid bij spoorwegen*” összetétel. Egy nem-szimmetrikus tezauszban lehetőség van a USE indikátor, mögötte pedig az összetevő deskriptorok plusz jeles kapcsolására (5. táblázat). Ennek az eshetőségnek viszont komoly következményei vannak a tezausz szerkezetére.

Angol	Holland
Rail safety	Veiligheid bij spoorwegen USE Veiligheid + Spoorwegen

#### 5. táblázat

*Megoldás az összetett kifejezések alaki eltérésére a nem-szimmetrikus struktúrában (IFLA, 2009)*

Az összetett kifejezések előfordulásuknak két formája van, a jelzős és az előjárószavas kifejezések. Az utóbbiak jelenlétét lehetőleg kerülni kell a tezauszban. Éppen ezért lehetőség szerint az angol „*Carbohydrate metabolism*” (szénhidrát anyagcsere) kifejezést kell előnyben részesítenünk, a megegyező jelentésű „*Metabolism of carbohydrates*” kifejezéssel szemben. A jelzős kifejezések felépítése az angol és a magyar nyelvben azonos, vagyis a jelző a főnév előtt helyezkedik el. A szavak elhelyezkedése a hollandban is ezzel megegyező, azonban a két szó egybeolvad. Az olasz nyelvben a jelző a főnév mögött kap helyet, közéjük pedig egy kötőszó ékelődik be (6. táblázat).

Angol	Holland	Olasz	Magyar
Library catalogues	Bibliotheekcatalogi	Cataloghi di biblioteca	Könyvtári katalógusok

#### 6. táblázat

*A jelzős kifejezések formai különbségei az egyes nyelveken (IFLA, 2009)*

Az is előfordulhat, hogy az előljárószavas kifejezésnek még sincs közvetlen egyenértékűje a másik nyelven. Ebben az esetben eltérő indikátorhasználat lesz jellemző a szimmetrikus és a nem-szimmetrikus tezaurusokban (7. táblázat).

Megoldás a szimmetrikus tezaurusban	
Angol	Holland:
Accessories UF Accessories after the fact	Medeplichtigen
Megoldás a nemszimmetrikus tezaurusban	
Angol	Holland:
Accessories NT Accessories after the fact	Medeplichtigen

7. táblázat

Pontatlan ekvivalencia a szimmetrikus és a nemszimmetrikus tezaurusban (s. k.)

### 2.3. Az ekvivalenciák

A nyelvek közötti ekvivalenciának három megközelítése van: szemantikai, kulturális és strukturális. Az előbbi két szempont a kifejezések jelentésére és arra vonatkozik, ahogyan ezeket az egyenértékű alakokat egy adott nyelven vagy kultúrában használják. Így ezek a megközelítések csak a deskriptorokkal kapcsolatban lesznek relevánsak, a legtöbb többnyelvű tezaurusban ugyanis nem tesznek kísérletet arra, hogy egyenértékű kifejezéseket rendeljenek minden nem-deszkriptorhoz. Ez a gyakorlat a neologizmusok elkerülése miatt használatos.

Német	Angol	Francia	Holland	Magyar
Schiennetz	Rail network	Résau ferroviaire	Spoorwegnet	vasúthálózat
UF Eisenbahnlinie Eisenbahnstrecke	UF Railway line Railway track	UF Ligne ferroviaire Voie ferrée	UF Spoorlijn Spoorweg	UF vasútszakasz vasútvonal vasútvonalszakasz

8. táblázat

Pontos egyenértékűség az EuroVoc-ban (s. k.)

Angol	Spanyol	Francia
Historic settlements	= Asentamientos históricos	≈ Site de peuplement

9. táblázat

Pontatlan egyenértékűség (HEREIN Tezaurusz, 2003)

A deskriptorok közötti interlingvális ekvivalenciának öt különböző fokát lehet megkülönböztetni az alulról-felfelé kialakított tezaurusokban: a pontos, a pontatlan, a részleges, az egy a többhöz egyenértékűséget, illetve azt az esetet, amikor nem létezik ekvivalens a másik nyelven.

- 1) A pontos egyenértékűség, más néven nyelvek közötti szinonima, abban az esetben áll fenn, amikor X és Y nyelven a deskriptorok szemantikai és nyelvi egyenértékűséget egyaránt mutatnak (8. táblázat).
- 2) A következő típus a pontatlan vagy közeli egyenértékűség, amely a nyelvek közötti kvázi-szinonimák, vagyis adott szakterület azonosnak tekinthető kifejezéseinek nézőpontbeli különbségeinél merülnek fel. Ebben az értelemben, a deskriptorok X és Y nyelven ugyanazt az általános fogalmat fejezik ki, de a kifejezések jelentése X és Y nyelven nem teljesen azonosak. Ezek a differenciák gyakran inkább kulturális jellegűek, mint szemantikaiak, azaz jelentésárnyalatbeli vagy felfogásbeli különbségről beszélhetünk. Így például: A „*Strong-willed*” és a „*Pig-headed*” angol kifejezések jelentése egyaránt „*Stubborn*” vagyis magyarul makacs, de míg az első kifejezés pozitív jelentéssel bír, addig a második negatívval. Az IFLA irányelvei alapján a pontatlan egyenértékűség esetében a kifejezéseket úgy kell kezelni, mintha pontos egyenértékűség állna fenn a kifejezések között (9. táblázat).
- 3) Az egyenértékűség következő megkülönböztetési fokán a részleges ekvivalenciák helyezkednek el, amelyek a nyelvek közötti kvázi-szinonimák között specifikusságbeli differenciát mutatnak. Ebben az esetben a deskriptor, a nyelvek egyi-

kén egy kissé tágabb vagy szűkebb értelmű, mint ugyanaz a deskriptor egy másik nyelven. A részleges ekvivalencia esetében az IFLA irányelv három lehetőséget kínál. Az egyik a már fentebb is használt megoldás, vagyis az, hogy kezeljük úgy a kifejezéseket, mintha egymás egyenértékű lennének. A második az, hogy minden egyes nyelvi változatot tüntessünk fel a tezauszban, és rendezzük ezeket hierarchikusan, vagyis a tágabb értelmű kifejezést fölérendeltként, míg a szűkebb értelmű kifejezést alárendeltként adjuk meg (10. táblázat). Kezelhetjük úgy is az esetet, mintha egy a többhöz egyenértékűség lenne.

Német	Angol
Wissenschaft	Wissenschaft SN Loan term adopted from German NT Science
Science D Lehnwort aus dem Englischen OB Wissenschaft	Science  BT Wissenschaft

**10. táblázat**

*Részleges ekvivalencia (ISO 5964,1985)*

4) Hudon szerint az egy a többhöz ekvivalencia, abban az esetben áll fenn, ha túl sok vagy túl kevés egyenértékű kifejezés áll a rendelkezésünkre. Ennek az ekvivalenciának tehát két altípusa van, az egyik, amikor a forrásnyelvi kifejezésnek több mint egy egyenértékű kifejezése van a célnyelven. A másik pedig ennek az ellentéte, vagyis ha nincs elég célnyelvi kifejezés. Túl sok célnyelvi kifejezés esetén három lehetséges megoldás létezik:

a) Létrehozhatunk egy előre összevont célnyelvi deskriptort (11. táblázat).

Angol	Francia
Listed building	Édifice inscrit + Édifice classé EP Édifice inscrit EP Édifice classé

**11. táblázat**

*Egy a többhöz ekvivalencia, összevont célnyelvi deskriptor létrehozása. (HEREIN tezausz)*

b) Módosíthatjuk vagy specifikálhatjuk a forrásnyelvi kifejezést például egy minősítő hozzáadásával (12. táblázat).

Angol	Német
Festival (feast) RT Holiday	Fest RT Feiertag
Festival (series of performances) RT Event	Festival  RT Event

**12. táblázat**

*Egy a többhöz ekvivalencia, specifikáció minősítővel (Schott, 2006)*

c) Létrehozhatunk egy vagy több célnyelvi nemdeskriptort a deskriptorhoz kapcsolódóan (13. táblázat). Az utóbbi megoldás a leginkább elfogadható a nyelvi ekvivalencia szempontjából, mivel így nem keletkeznek mesterségesen létrehozott kifejezések.

Angol	Német
Festival	Festival BF Fest VB Holiday
	Fest BS Festival

**13. táblázat**

*Egy a többhöz ekvivalencia, célnyelvi nemdeskriptorral (Schott, 2006)*

Az egy a többhöz ekvivalencia másik altípusa abban az esetben áll fenn, ha a célnyelvi kifejezés csak forráskifejezések kombinációja révén képviselheti a fogalmat, vagyis nincs megfelelő célnyelvi kifejezés. Erre a problémára a következő négy megoldást alkalmazhatjuk:

a) Létrehozunk egy összevont célnyelvi kifejezést (14. táblázat).

Német	Angol
Berufspraktikum	Vocational education + Internship

**14. táblázat**

*Egy a többhöz ekvivalencia, összevont célnyelvi kifejezésekkel (Schott, 2006)*

b) Produkálhatunk egy általunk megalkotott kifejezést, ún. neologizmust (15. táblázat).

Német	Angol
Berufs- praktikum	Vocational training SN Use this descriptor for a combination of vocational education and internship

**15. táblázat**

*Egy a többhöz ekvivalencia, neologizmus  
(Schott, 2006)*

- c) Használhatunk fölérendelt célnyelvi kifejezést, minősítővel együtt (16. táblázat).

Német	Angol
Gymnasium	Secondary school (gymnasium)

**16. táblázat**

*Egy a többhöz ekvivalencia fölérendelt kifejezés  
minősítővel (Schott, 2006)*

- d) Kialakíthatunk egy vagy több forrásnyelvi nem-deszkriptort, amely(ek) kapcsolatban áll(nak) a deszkriptorral (17. táblázat).

Angol	Német
Town redevelopment	BS Stadt und Sanierung
Town	Stadt
Redevelopment	Sanierung BF Stadtsanierung

**17. táblázat**

*Egy a többhöz ekvivalencia forrásnyelvi  
nemdeszkriptorokkal (Schott, 2006)*

- 5) A nyelvek közötti ekvivalencia utolsó típusa abban a helyzetben áll fenn, ha nem létezik a forrásnyelvével egyenértékű kifejezés, amely azonos jelentéssel bírna a célnyelven. Hudon szerint (1997) ebben az esetben a szimmetrikus struktúra érdekében, a legegyszerűbb megoldás az, ha eltávolítjuk a „gazdátlan” forrásnyelvi kifejezést, különösen akkor, ha nagyon speciális fogalmat képvisel. Ha azonban az eltávolítás kivitelezhetetlen, további három lehetőség áll rendelkezésünkre:

- a) Átalakíthatjuk a „gazdátlan” kifejezést nem-deszkriptorrá és ezt összekapcsolhatjuk egy olyan deszkriptorral amellyel számos tulajdonosságban osztozik (18. táblázat). Ez a megoldás nincs különösebb hatással a szerkezetre, így ez lehet az ideális választás, feltéve, ha a „gazdát-

lan” kifejezés valójában közeli rokona az adott deszkriptornak.

Angol	Francia
Classroom environment USE Learning environment	
Learning environment UF Classroom environment UF Educational environment	Milieu éducatif EP Conditions d’apprentissage EP Milieu pédagogique EP Milieu d’apprentissage

**18. táblázat**

*Ha nincs ekvivalencia, átalakítjuk a kifejezést nem-deszkriptorrá (Canadian literacy thesaurus, 1996)*

- b) Importálhatjuk a forráskifejezést a célnyelvre (19. táblázat). Ennek a lehetőségnek a hátulütője, hogy a felhasználók így olyan információhoz juthatnak, amelyek csak egy másik társadalomban léteznek, a sajátjukra nem érvényesek.

Német	Angol
Diakonisches Werk	Diakonisches Werk SN Form of social work done by protestant churches in Germany

**19. táblázat**

*Ha nincs ekvivalencia, importálhatjuk a forrás  
kifejezést a célnyelvre. (Schott, 2006)*

- c) Létrehozhatunk egy egyenértékű új szót vagy kifejezést (20. táblázat). A neologizmusok alkalmazása azonban soha nem a legjobb megoldás. A teaurusz nem egy terminológiai kifejezésbank, a teaurusz önmagában nem idéz elő lényeges változásokat a nyelvben, hanem inkább tükröznie kell a nyelvek specializált használatát a társadalom bizonyos szegmenseiben.

Német	Angol
Hochrechnung	Projection (statistical)

**20. táblázat**

*Ha nincs ekvivalencia, neologizmust alkalmazhatunk  
(Schott, 2006)*

### 3. Már kialakított tezauszok összeolvasztása, összekapcsolása

Arra is lehetőség van, hogy ötvözzünk két vagy több már kialakított tezauszot egyetlen új többnyelvű változatba, illetve már kialakított tezauszok és tárgyszójegyzékek összekapcsolásával is létrehozhatunk többnyelvű struktúrát. A cél az, hogy felhasználjuk a meglévő szótárak szókincsét, amennyire csak lehetséges. Az ötlet mögött az az elképzelés áll, hogy a használók továbbra is azzal az ismert szókincssel dolgozhatnak, amivel már korábban is találkoztak, így egy ismerős szókészlethez egy új funkciót rendelhetünk.

Az összekapcsolás különböző formákban lehetséges. Az első esetben két vagy több megegyező nyelvű szótárat kapcsolhatunk össze, erre példát találunk az orosz nyelvterület gyakorlatában, vagy különböző nyelvű, egynyelvű szótárak összekapcsolását is elvégezhetjük, például német és francia szótárak kombinálását. Egynyelvű és többnyelvű szótár összevonásával is megvalósítható a többnyelvű tezausz kialakítása. Az összekapcsolás általában a tezauszok kifejezéseinek és a tárgyszójegyzékek szavainak feltérképezésével kezdődik, innen ered a „mapping” kifejezés. A találatokat sokféleképpen jeleníthetik meg, például egy táblázatban, amely annyi oszlopot tartalmaz, amennyi a felhasznált szótárak száma. A táblázat minden egyes sorában a különböző szótárak adott egyenértékű kifejezései helyezkednek el, melyben egy-egy cella nulla, vagy annál több kifejezést is tartalmazhat.

Hudon az alábbiakat írja az összekapcsolásról. Bár már jóval elfogadhatóbb ez a módszer, mint a fordítás, de nagyon komoly gyakorlati problémákkal jár együtt. Ezt az eljárást a legnehezebb intellektuálisan kezelni, mivel minden egyes tezausz eltér egymástól bizonyos mértékben, lefedettség mélységben, a prekordináció fokában, a specifikusság szintjén stb. Emellett valószínű, hogy a legnagyobb és legfejlettebb tezausz nyelve és struktúrája lesz a domináns, és az is, hogy a többi tezausz struktúráját fogják igazítani/egyeztetni ehhez az egyetlen dominánshoz.

#### 3.1. Az ekvivalenciák

Az összekapcsolással kialakított többnyelvű tezauszokban három ekvivalencia típust különböztetünk meg egymástól: a teljes és a hiányos ekvivalenciát, valamint azt az esetet, amikor nincs ekvivalense a ke-

resett kifejezésnek. Az utóbbi helyzetben tájékoztatni kell róla a felhasználót, hogy az általa keresett kifejezésnek nincs ekvivalense a másik nyelv(ek)en.

1) A teljes ekvivalenciánál, minden szókészletben adott kifejezésre az ekvivalensek keresése során egy és csak egy deskriptor létezik (21. táblázat). Az összekapcsolás azonban lehet hozzávetőleges. Az egyenértékűség nem azt jelenti, hogy itt a kifejezések azonos nyelvi jelentésűek, hanem azt, hogy várhatóan az eltérő szókészletekből származó kifejezések megegyező témájú forrásokhoz vezetnek.

Angol	Francia	Német
Decathlete	Décathloniens	Zehnkämpfer
Decathlon	Décathlon	Zehnkampf
Discus throwing	Lancer du disqu	Diskuswurf
Divers	Plongeurs	Kunstspringer
Diving	Plongeon	Wasserspringen
Hammer throwing	Lancer du marteau	Hammerwurf
Hurdle racing	Course de haies	Hürdenlauf

21. táblázat

*Teljes ekvivalencia, az összekapcsolás eljárásánál (MACS, 2004)*

2) Hiányos ekvivalencia esetében legalább az egyik szókészletben két vagy több kifejezést kell kombinálnunk, amelyek egymás egyenértékű kifejezéseiként használhatóak. Ez a kombináció alapulhat a Boole-algebra ÉS, illetve VAGY operátorain. A VAGY összekapcsolás akkor szükséges, ha a forrásnyelvi deskriptor a másik nyelven két kifejezést jelöl (22. táblázat).

Agol	Német	Francia
Jumping	Hochsprung OR Sprung	Saut en hauteur OR Sauts (athlétisme)
Long jump	Weitsprung	Saut en longeur

22. táblázat

*Hiányos ekvivalencia, VAGY operátorral (MACS, 2004)*

Az ÉS operátorra akkor van szükség, amikor az egyik szókészlet kifejezése hierarchikusan alacsonyabb szinten áll, mint a releváns kifejezés legalább egy

felhasznált másik szókészletben (23. táblázat).

Angol	Német	Francia
Cycling	Radsport	Cyclisme
Cycling– Training	Radsport AND Training	Cyclisme AND Entraînement

23. táblázat

*Hiányos ekvivalencia, ÉS operátorral (MACS, 2004)*

## 4. A tezauszok fordítása

### 4.1. A fordítás elmélete

A fordítási eljárás igen vitatott, éppen ezért ezt a megközelítést a szabványok és irányelvek nem is tartalmazzák, noha a meglévő tezausz egy vagy több új nyelvre fordítása igen népszerű, főleg gazdasági okokból. Ez a módszer nem teszi lehetővé az összes érintett nyelv közötti egyenlő kezelést, így mindig a forrásnyelv lesz a domináns. Az így előállított eszköz nem tükrözi megfelelően a célkultúrát, kulturálisan mindig elfogult lesz, mivel az egyenes fordítás egyfajta kulturális imperializmushoz vezet. A szakirodalomban vannak, akik azt állítják, hogy lehetetlen fordítás útján tezauszt készíteni, azonban a gyakorlatban a két- és többnyelvű tezauszok többnyire úgy keletkeznek, hogy egy forrástezauszt a célnyelvre fordítanak. Fordítás esetében is ugyanolyan szakembergárdára van szükség, mint a párhuzamos eljárásnál, legfeljebb a specialisták térben és időben el vannak választva egymástól. Emellett a fordítási eljárás számos nehézség forrása lehet. Ennél a módszernél alapvetően nagyobb kifejezéshalmaz keletkezik, ezzel egyetemben pedig rendkívül megnő a homonimák és a pontatlan szakkifejezések száma. Így több egyeztetésre van szükség az egyes munkafolyamatok között, megnő az ellenőrzés szerepe, ami még több munkát generál.

Hudon a következő három problematikus elemre hívja fel a figyelmünket, amelyek hagyományosan kapcsolódnak a többnyelvű tezauszok fordításához. Az első probléma az, hogy a kifejezések szó szerinti fordítása a forrásnyelvről semmitmondó kifejezéseket eredményezhet a célnyelven. A második kérdéses pont az, ha egy adott nyelvet tágitunk annak érdekében, hogy illeszkedjen egy idegen fogalmi struktúra azonos pontjaihoz. Ez azt eredményezheti, hogy a nyelv alig felismerhetővé válik beszélői számára.

Az utolsó intés a fordítással kapcsolatban az, hogy mielőtt áthelyeznénk egy teljes fogalmi struktúrát az egyik kultúrából a másikba, jól gondoljuk meg egyáltalán ez mennyire célszerű.

*Martin Doerr*<sup>8</sup> szerint a terminológiai források szemantikai különbözőségére gyakran hivatkoznak a szakirodalomban alapvető problémaként. Ezt azonban a szerkezetek szisztematikus elemzésével nem lehet kiküszöbölni. Véleménye szerint a fordított tezausz is tezausz, ahol minden fogalmat optimálisan értelmeztek több nyelven, ezek a fordítások azonban nem minden esetben megalapozottak.

Bár úgy tűnhet, hogy általános megegyezés van abban, hogy nincs értelme a tükör-fordításnak, a fordítást nem tekinthetjük egyszerű műveletnek. *Eugene Nida*<sup>9</sup> szemlélete a fordítót a saját kultúrája „foglyának” tartja, mivel a fordító általában a saját nyelvi-kulturális kontextusában érti meg az üzenetet, majd megpróbálja megoldani a problémát idegen eredetű kifejezések használatával. Ez a folyamat független attól, hogy egy forrásnyelvi fogalom szemantikailag nem mindig egyenértékű a célnyelv megegyező fogalmával. Ezért egy másik kultúrába ágyazásnál a magyarázó kifejezések használata sokkal jobb, mint ha egy lefordított idegen nyelvű kifejezést használunk. Ez az elv tükröződik a párhuzamosan szerkesztett tezauszkészítési szabványokban és irányelvekben is, mivel a jövevényszavak, kölcsönzött kifejezések, neologizmusok használata sehol sem ajánlott. Az egyenértékűség típusainak felosztása is ebből a felismerésből fakad. A funkcionális, kommunikatív fordítási elméletek közé tartozik többek között a dinamikus ekvivalencia elmélete, a funkcionális ekvivalencia elmélete és a SKOPOS elmélet.<sup>10</sup>

A dinamikus ekvivalencia elméletét Nida fejlesztette ki 1969-ben, ez volt az első kommunikatív fordítási elmélet. A dinamikus szó arra az elképzelésre utal, hogy a fordításnak nem kell ugyanazt a hatást eredményeznie, mint a forrásszövegnek az eredeti kontextusban. Nida fordítási meghatározása szerint a fordítás újraalkotja a célnyelvben a forrásnyelv legközelebbi természetes ekvivalensét, először a kifejezések értelmében, másodsor a kifejezések stílusában.

A rendeltetési vagy funkcionális ekvivalencia teóriájának az az elképzelés az alapja, hogy a forrásszöveg rendeltetése megegyezik, vagy hasonló a célszöveg rendeltetésével. Más szavakkal a fordítás rendeltetése hozzáigazodik a forrásszöveg rendeltetéséhez.<sup>11</sup>

A SKOPOS elmélet a fordítási műveletet hangsúlyozza, tehát a fordítás cselekvési folyamatát. *Vermeer*<sup>12</sup> szerint a fordítási művelet, olyan, mint egy bizo-

nyos viselkedéstípus, ezt a viselkedést a cselekmény szempontjából cselekedetnek nevezhetjük, mivel az azt végző személynek kell potenciálisan képesnek lennie arra, hogy megmagyarázza, miért járt el úgy ahogy, bár más módon is eljárhatott volna. Továbbá a cselekedet valódi okai mindig kifejezhetőek a célok szándékainak vagy állításainak segítségével a kifejezésekben, mint egy jó okkal elkövetett cselekvés.<sup>13</sup> *Vehmas-Lehto* gondos meghatározást adott a fordítás meghatározására: *fordításkor valaki kifejez valamit a célnyelv eszközeivel, amit korábban már kifejeztek a forrásnyelv eszközeivel*. A definíció jelzi, hogy leginkább a fordítás tartalma a fontos, a jelentés megtartása és nem maga a nyelv.

#### 4.5. A fordítást kiegészítő statisztikai eljárások

*Susanna Keränen* a fordítási eljárást statisztikai módszerekkel is támogatná, hogy a megfelelő szókészlet feltérképezésén túl, az adott nyelv kifejezései közötti kapcsolatok is átláthatóbbá váljanak a fordító számára. Az általa javasolt eljárások: a társszó analízis, a szó társítási módszer és a szemantikus alkotórész analízis.

##### A társszó elemzés

A társszó elemzés nagyon hasonlít az együtt idézés elemzésére, a kifejezések együttes előfordulását elemzi a dokumentumokban. A társszó analízist úgy lehet jellemezni, mint egy kétdimenziós térképet. A térképen a relatív kör mérete összefügg az előfordulások számával, ez alatt azt érthetjük, hogy minél nagyobb a deskriptorok száma, értelemszerűen annál nagyobb lesz a körünk, viszont ha kisebb a kör, akkor több kapcsolat lesz az egymással relációban álló deskriptorok között. A társszó analízis eredménye kétdimenziós térképen is bemutatható.

##### A szótársítási módszer

A szótársítási módszert már korábban is alkalmazták a tezauszokészítés keretein belül. *Marianne Nielsen*<sup>14</sup> szerint a szótársítási módszer egy használható és működőképes tezauszot eredményezhet, amely jobban teljesít, mint a hagyományos tezauszokészítési módszerekkel kialakított társaik. A szótársítási módszer előnye az, hogy gazdagíthatja a tezauszépítést az aktuális, releváns és területspecifikus információk biztosításával.

##### A szemantikai komponens analízis

A szemantikai komponens analízist gyakran említik

jelentés analízisként is. Az eljárás a szó jelentésváltozatait osztja kisebb elemekre, szemantikai jellemzőkre. Az összetevők elemzése különösen hasznos a poliszémiák esetében.

#### 5. A többnyelvűség kezelésének nehézségei a modern igényeknek megfelelően

A többnyelvű információkeresés biztosítására nincs egyszerű és kézenfekvő megoldás. *Kerstin Jorna*<sup>15</sup> úgy véli, hogy a többnyelvű tezauszok leggyakoribb formái jelenleg nem igazán megfelelőek erre a szerepre. Véleménye szerint a többnyelvű eszközök ezen típusának új formájára van szükség. E célból kezdte el kialakítani az InfoDEFT nevű kísérleti tezauszát, amit az online tezauszok lehetséges új modelljének tart. Az InfoDEFT egy szemantikailag strukturált, enciklopédikus jellegű és többnyelvű tezausz, amely felhasználható a mesterséges tanulást segítő programok esetében is, ezzel csökkentve az eszköz manuális előállításának költséges mivoltát.

A többnyelvű tezauszok összetett kialakítási problémáira *Hahn* és *Angelova*<sup>16</sup> szerint ideális opció lehet, egy független nyelvi tudásbázis bevezetése, amely közös fogalmi struktúrát alkothatna minden egyes nyelven. *Jorna* szerint egy ilyen tudásbázis kiépítése azonban rendkívül költséges lenne, és nagy erejű nemzetközi összefogást kívánna meg. Éppen ezért a problémára inkább a háromnyelvű kísérleti tezauszában az InfoDEFT-ben látja a választ, mely rendkívül hasznosan alkalmazható az információ visszakeresésénél és a felhasználók számára is a tények, kifejezések megértéséhez. Erre a modern értelmezésre *Schmitz-Esser*<sup>17</sup> EXPO 2000 tezauszja is rávilágított, azonban utóbbi esetében nem hívják fel a használó figyelmét az egyes nyelvek közötti kulturális különbségekre. *Jorna* verziójában azonban az egyes kifejezések szoros ellenőrzése nagy figyelmet kap. Erre az InfoDEFT-et kialakító csoport facetták alkalmazását találta a legideálisabbnak, vagyis egy olyan szerkezet létrehozását, amely eltér az egyes nyelvektől, ezzel is felhívva a figyelmet a különbségekre, nem csupán a szókészletben, hanem a fogalmi kontextusban is.

*Jorna* három fő sarokpontot határoz meg, amelyek egy jó tezausz alapkövei kell, hogy legyenek. Először is a tezauszoknak többnyelvűnek kell lennie, hogy támogassa az információhoz való hozzáférést az egyre inkább globálissá váló információs társadalomban. Másodsor ezeknek a rendszereknek az alfabetikus forma helyett szemantikailag strukturáltaknak

kell lenniük, ez a szempont részben lehetővé teszi a felhasznált nyelvek egyenlő képviselését. Harmadszor segíteniük kell a felhasználók számára a megértést, a felidézést és a precíz információ-visszakeresést a definíciók és a megjegyzések használatával.

Míg a modern információs és kommunikációs technológia megkönnyíti a globális adatok transzferét, ugyanez nem segít a globális kommunikáció megvalósulásában. Jelenleg mind az internet és a nemzetközi vállalkozások is nagyrészt az angol nyelvet használják a nemzetközi kommunikációban, amely problémákat vet fel a kultúrák közötti kommunikációban az angol anyanyelvűek és a más ajkúak között. A legújabb kutatások szerint az angol nyelv hamarosan megszűnik uralni a globális kommunikációt, helyét Ázsia és Latin-Amerika egyes nyelvcsoportjai fogják átvenni.<sup>18</sup>

Az említett előrejelzések tükrében új többnyelvű teauruszokra lesz szükség, hogy támogassák a későbbi uralkodói státuszt nyerő nyelvcsoportok közötti kommunikációt, illetve a nyugat-európai nyelvekhez kapcsolódó diskurzust. Manapság sok úgynevezett többnyelvű eszközhöz férhetünk hozzá az interneten vagy a könyvtárak OPAC-jában, valójában azonban ezek egynyelvű teauruszok, amelyek két- vagy többnyelvű szótárakkal együtt működnek. A többnyelvű információ visszakereső rendszerek ezen formái gyakran csődöt mondanak, amit igazolni lehet a LISA (Library and Information Science Abstracts) részletes vizsgálatával. A LISA teaurusza a címkéző kifejezések legnagyobb csoportját angol nyelven tartalmazza, annak ellenére, hogy több mint 60 ország szakterméséből tartalmaz absztraktokat és teljes szövegű hozzáférést. A LISA-nál minden absztraktot angol nyelvre fordítanak, ami súlyosan ellentmondásos keresési eredményekhez vezet abban az esetben, ha nem adottak a pontos ekvivalensek a nyelvek között. A francia „*documentalistes*” a német „*Dokumentar*” és az angol „*documentalists*” egymás egyenértékű kifejezéseiként vannak feltüntetve a teauruszban, de különböző keresési eredményekhez vezetnek, a különböző nyelvi változatokra keresve 55, 4, vagy 110 rekordot kapunk. Amennyiben az „*information professionals*” kifejezéssel keresünk, akkor pedig 3372 rekordot kapunk, de ha a kifejezés nem-deszkriptorával a „*documentalists*” fogalommal ugyanezt végrehajtjuk csak 590 rekordot kapunk. Az ilyen következetlenségek legyőzhetőek abban az esetben, ha egyenlő hozzáférést nyújtunk minden nyelven ahelyett, hogy az angol nyelvet részesítsenénk előnyben. Ez nem csak a kultúrák és így az emberek közötti

egyenlő bánásmódot jelenti, hanem lehetővé teszi a teaurusz tárgyainak elfogulatlan kezelését is.

Sajnos nem könnyű feladat minden nyelv egyenlő kezelése a rendszeren belül. Valaha még voltak fizikai korlátai a többnyelvű megjelenítésnek, de mára ezeket le lehet küzdeni az elektronika segítségével. A technológia azonban nem tud győzedelmeskedni a szemantikai egyenértékűség intellektuális problémái fölött. A természetes nyelvek komplexitása, amely minden terminológia része, ritkán engedi pontos ekvivalenciák kialakítását a nyelvek között, ami igen sok gondot okoz a szemantikailag strukturált teauruszokban. Ugyanez a hiba kevésbé okoz fejtörést az alfabetikus teauruszok esetében, de a szakemberek szerint ez az elrendezés nem is lehet alapja egy jó teaurusznak, csak egyfajta kiindulópontként kell szolgálnia.<sup>19</sup>

### 5.1. Az EXPO 2000 és az InfoDEFT újszerű megoldásai

Újszerű megoldás, ha facetták vagy más szemantikai struktúrák segítik a felhasználót a kifejezések megértésében, és annak az átlátásában, hogyan viszonyulnak egymáshoz a mögöttes fogalmi kapcsolatok. A Schmitz-Esser és csapata által Hannoverben kialakított EXPO 2000 többnyelvű teaurusz általános ismeretkört ölel fel öt nyelven. A Jorna nevéhez fűződő InfoDEFT pedig háromnyelvű, csak az információ-visszakeresés kifejezéseit feltáró teaurusz. Mindkettő elektronikus formájú. A két terminológiai szótár azért érdekes számunkra, mert a saját területükön legyőzik a fogalomtárak leggyakoribb nehézségeit. A legnagyobb különbség ezen példányok és a korábbi teauruszok között az, hogy túlmutatnak az alapvető szemantikai viszonyokon, a tágabb, a szűkebb és a kapcsolódó kifejezések szempontjából. Az InfoDEFT facettás alapú, míg a Schmitz-Esser-féle teaurusz a hierarchikus kapcsolatokban 13 alosztást különböztet meg (24. táblázat).

1. Nem-faj kapcsolat: növény – kókuszdió	7. Ok-okozati kapcsolat: erdőirtás- elsivatagosodás
2. Fizikai és elméleti felosztás: gazdaság – versenyszféra	8. Előnyös kapcsolat: alfabetizáció – magasabb életszínvonal

3. Jog és joghatósági kapcsolat: Európai Unió - Franciaország	9. Hátrányos kapcsolat: túlhalászás – tengeri élővilág pusztulása
4. Földrajzi és helyrajzi felosztás: Latin-Amerika – Dél-Amerika	10. Anyagi kapcsolat futószalag - gumi
5. Származási kapcsolat: nagyapa – apa	11. Forma és megjelenés kapcsolat: kettős spirál - kromoszóma
6. Műszeres kapcsolat: hegesztőpáka - hegesztés	12. Folyamat kapcsolat: futószalag - gyártás
13. Állapot kapcsolat víz – jég	

**24. táblázat**

*Az EXPO 2000 alosztásai: (s. k.)*

Schmitz-Esser bevezetett egy alapvető szemantikai referencia szerkezetet a bizonytalan asszociatív kapcsolatok helyett, amely hat fő kategóriát tartalmaz: a fogalom, a név, az esemény, a hely, a szempont és az idő megkülönböztéseit. Ez a rendszerezés részben megegyezik az InfoDEFT öt fő kategóriájával, mely a következő kifejezések alapján válogat: dolgok, emberek, tevékenységek, helyek és környezet. Mindkét tezausz tehát azt mutatja be, milyen értelemben kapcsolódnak egymáshoz a kifejezések.

Jorna ugyan egyetért Schmitz-Esser besorolási rendszerével az általános tárgykörű tezauszban, azzal azonban nem, hogy az EXPO 2000-ben a lexikális hézagokat parafrázisokkal töltik ki. Mivel Jorna szerint a lexikai hiányosságokra inkább fel kell hívni a figyelmet, mint azokat elrejtetni. A felhasználói tudatosság fokozásával ugyanis, nő a fogalmi megértés és javul a kommunikáció.

Az InfoDEFT azonban, egy második és lényegesebb tekintetben is különbözik az EXPO 2000-től. Utóbbiban ugyanis minden deskriptor egy kétdimenziós rácsszerkezetben kerül elhelyezésre, melynek egyik dimenziójában a tizenhárom hierarchia típus, a másikban pedig a hat alapvető szemantikai referencia található. Ezzel szemben az InfoDEFT-ben facetták vannak, amelyek a gráfok felépítését követik, vagyis számos horizontális kapcsolatot tartalmaznak, melyek a fák ágainak felelnek meg. A Schmitz-Esserféle táblázatos megjelenítésnek az az előnye, hogy átláthatóbb és következetesebb a felhasználók szá-

mára, vagyis inkább enciklopédikus áttekintést nyújt, amely gyakran megy a nyelvi pontosság rovására. Az InfoDEFT-re ezzel ellentétben az intuitív megközelítés a jellemző, a cél a kifejezések jelentésének jobb megértése a különböző nyelveken, amely az enciklopédikus áttekintés helyett. Végeredményben megállapítható, hogy a tezauszok különböző funkciói indokolják ezeket az eltérő modern módszereket.

## 5.2. Az interoperabilitás a jövő tezauszaiiban

Az interoperabilitás a kulcsa az egyidejű keresés és visszakeresés képességének két vagy több adatbázisban egy vagy több nyelven. Az interoperabilitás a különböző rendszerek és szervezetek képessége az együttműködésre. A kifejezést általában technikai rendszerekben mérnöki értelemben használják, vagy alternatív módon, tágabb értelemben véve a társadalmi, politikai és szervezeti tényezők figyelembevételével. Kezdetben az interoperabilitást informatikai rendszerek definiálták, a rendszerek közti korlátozott információcserére. A szemantikus interoperabilitás a rendszer azon képessége, hogy értelmesen és pontosan feldolgozza a kicserélt információt a rendszer végfelhasználóinak érdekében. Alapfeltétele, hogy mind az átadó és a befogadó oldalon is alkalmaznia kell a referenciamodellt, vagyis azt, hogy amit elküldünk, egyezzen meg azzal, amit megértünk belőle.<sup>20</sup>

A rendszerek interoperabilitása, egy másik modern megközelítése annak, hogy az eszköz használata egyszerűbbé váljon a felhasználó számára. Régen, miután megnéztünk egy nyomtatott tezauszt, továbbhaladtunk a kiválasztott kifejezések nyomán az indexelő rendszerben, vagy a kereső felületeken. Manapság a kifejezéseket csak át kell másolni az elektronikus tezauszba, vagy gyakrabban csak rájuk kattintani. A kereső rendszer csak akkor lesz képes kapcsolatba lépni a tezausz adatbázisával, ha az interoperabilitás fennáll. Vagyis a felhasználók böngészhetnek és miközben keresnek, kiválaszthatják a tágabb kifejezéseket, egy általánosabb keresésre és így tovább. Ezek az átvitelek nem bonyolultak, feltéve, ha alkalmazzuk a szabványos adatsere formátumokat és protokollokat.

A tezauszok vagy más információkereső rendszerek interoperabilitása, több létfonosságú folyamatot tesz lehetővé. Az egyik az, hogy biztosítja az adatok szabványos formában való behozatalát és használatát más rendszerekben. A másik az, hogy lehetővé teszi a

rendszerek között a kifejezések leképezését, valamint biztosítja a bármilyen más típusú információcserét a rendszerek között.<sup>21</sup>

Az interoperabilitást elérni, azonban nem olyan könnyű. A szemantikus webben az ontológiák esetében szintén ilyenfajta együttműködésre van szükség. Hamarosan arra is igény lesz, hogy a tezauszokat összekapcsolják az ontológiákkal és/vagy a taxonómiákkal. Jelenleg ennek megoldásával foglalkozik az ISO 25964 szabvány<sup>22</sup> második „*Interoperabilitás más szótárakkal*” c. részén dolgozó munkacsoport is. Így ez lehet a többnyelvű tezauszok jövője.

## 6. Következtetések

A tezausz teljes fogalmi és terminológiai készletet kínál minden képviselt nyelven, a többnyelvűség hozzáadásával pedig biztosíthatjuk az információhoz való hozzáférés széles skáláját. Ezen eszközök tekintetében is rendkívül fontos azonban, hogy a struktúrában a relációk és a fogalmak azonos értékek legyenek, illetve hogy a fogalmi értelmezés középpontjában a válogatás helyett a terminológiai ellenőrzés álljon. Egy jó tezausz alapvetően szemantikailag strukturálnak kell lennie, persze, magában foglalhat egy betűrendes részt is, de fontos, hogy ne ez legyen benne a fő rendezési elv.

A többnyelvű tezauszok kialakítási eljárásai közül az ab initio nyelvpárokkal dolgozó eljárását preferálják a szakirodalomban, a szabványok alapvetően ehhez a folyamathoz nyújtanak iránymutatást. Ez a módszer szimmetrikus tezauszstruktúrára törekszik és a takarékos szógyűjtés jellemzi. Az alulról-felfelé építkezés során sorolják az ekvivalenciákat a legtöbb kategóriába, vagyis a legjobban kidolgozott az egyenértékűségek megkülönböztetése is. Éppen azért egyértelmű, hogy az így kapott tezauszok lesznek a legjobbak minden szempontból. Azonban az eljárás legnagyobb kritikája az, hogy komoly szakembergárdát és költségbeli ráfordításokat kíván.

A már kialakított tezauszok és/vagy tárgyszólisták összekapcsolásának előnye az, hogy a felhasználók továbbra is azzal az ismert szókinccsel dolgozhatnak, amivel már korábban is találkoztak, vagyis egy ismerős szókészlethez egy új funkciót rendelnek hozzá. A kialakítás ezen típusa kevesebb ekvivalencia típust használ, mivel a szókészletek „összefésülése” során alapvetően egy létező struktúrához kell alkalmazkodni, így nincs meg az a szabadság, mint az alulról-felfelé történő építkezés során. Hátránya, hogy sokkal több kifejezés gyűlik össze és aszim-

metrikus szerkezetet eredményez. Ez a módszer – a párhuzamos szerkesztéshez hasonlóan – megfelelő szakembergárdát és költségbeli ráfordításokat követel, azonban nem jár akkora pontossággal.

A fordítás mind közül a legvitatottabb eljárás, egyes vélemények szerint nem is lehet ilyen formában jó tezauszt kialakítani. A szabványok szerzői sem preferálják ezt az eljárást, a gyakorlat azonban ennek ellentmond, mivel számos esetben így pl. az *EuroVocnál* is kezdetben ugyan követték az alulról-felfelé építkezést, később azonban az egyes tagországok csatlakozásával már fordítást használtak. Az eljáráson belül megkülönböztetjük a szakemberek által kialakított és a mechanikus eszközökkel létrehozott szótárakat (utóbbi nagy minőségbeli hiányosságokat eredményezhet szigorúbb emberi ellenőrzés nélkül). Azonban ezt az eljárást is lehet jól kivitelezni, amennyiben emberi erőforrás végzi a munka nagy részét, illetve ha párosítjuk a Keränen által is javasolt módszerekkel (a társszó analízissel, a szótársítási módszerrel vagy a szemantikus alkotórész analízissel). Ezekre azért van szükség, hogy a megfelelő szókészlet feltérképezésén túl, az adott nyelv kifejezései közötti kapcsolatok is átláthatóbbá váljanak a fordító számára.

A tezauszokat bírálók kritikája általában az, hogy ezen eszközök előállítása költségesebb és nehéz a bővítés. Azonban egy jól kialakított tezauszban már a kidolgozás előtt biztosítanak arra módszereket, hogy a frissítés folyamatos legyen. Tény az is, hogy a tezauszalapú keresőrendszerek rendelkeznek a legjobb pontossági aránnyal, így hiba volna nem használni őket. A véleményem az, hogy a kialakítási eljárást olyan alternatívákkal kell kiegészíteni, amelyek áthidalják a kulturális és nyelvi különbségekből adódó problémákat. Ezek lehetnek statisztikai módszerek, egy egyedi facettás struktúra alkalmazása, illetve az elektronikus rendszerek átjárhatóságát biztosító lehetőségek. A lényeg az, hogy a befektetett idő és energia megtérül, így bátran kombináljuk a klasszikus eljárásokat más újszerű lehetőségekkel!

## Irodalom

1. A téma kutatása során hatalmas segítséget jelentett *Barátné dr. habil. Hajdu Ágnes*, tanszékvezető egyetemi docens (ELTE BTK, Könyvtár- és Információtudományi Intézet) biztatása és szakmai útmutatása, amiért ezúton is szeretném hálás köszönetemet kifejezni.
2. BARÁTNÉ HAJDU Ágnes – BABICZKY Béla: Bevezetés az

- információkereső nyelvek elméletébe és gyakorlatába. Budapest, Universitas, 1998. 224 p.
3. HUDON, Michele: Multilingual thesaurus construction: integrating the views of different cultures in one gateway to knowledge and concepts = *Information Services & Use*, 17. vol. 1997. 2–3. no. 111–123. p.
  4. HUDON, Michele: Relationships in multilingual thesauri. In: *Relationships in the organization of knowledge*. Ed. Carol A. Bean, Rebecca Green. Dordrecht, Kluwer Academic Publishers, 2001. 67–80. p.
  5. Guidelines for multilingual thesauri. Hague, Working Group on Guidelines for Multilingual Thesauri, IFLA Classification and Indexing Section, 2009. 30 p. <http://www.ifla.org/files/assets/hq/publications/professional-report/115.pdf> (Az utolsó megtekintés ideje: 2015. október 8.)
  6. UNGVÁRY Rudolf: A két- és többnyelvű tezauszok irodalma. Kritikai szemle és annotált bibliográfia. Szerk. Ungváry Rudolf, Orbán Éva, Budapest, OSZK KMK, 1981. 123 p.
  7. DAHLBERG, Ingetraut: Concept and definition theory. In: *Classification theory in the computer age. Conversations across disciplines*. New York, Rockefeller College Press, University at Albany, State University of New York, 1989. 12–24. p.
  8. DOERR, Martin: Semantic problems of thesaurus mapping = *Journal of Digital Information*, 1. vol. 2001. 8. no. <https://journals.tdl.org/jodi/index.php/jodi/article/view/31/32> (Az utolsó megtekintés ideje: 2015. október 8.)
  9. NIDA, Eugene Albert – REYBURN, William: *Meaning across cultures*. New York, Orbis Books, 1981. 90 p.
  10. KERÄNEN, Susanna: Multilingual thesaurus construction. *Homemakers in Finnish and British discourses*. Barcelona, ISKO, Capitulo Espanol, Barcelona, 2005. 357–379 p. <http://bd.ub.edu/isko2005/keranen.pdf> (Az utolsó megtekintés ideje: 2015. október 8.)
  11. VEHMAS-LEHTO, Inkeri: Kopiointia vai kommunikointia. *Johdatus käännösteoriaan* = *Helsingin sanomat*, 9. vol. 1999. 7. no. 122–232 p.
  12. VERMEER, Hans, J.: Skopos and commission in translational action. In: *Readings in translation theory*. Ed. Chesterman, Andrew, Loimaa: Oy, Finn Lectura Ab, 1989. 173–187. p.
  13. REISS, Katharina – VERMEER, Hans J.: *Mitä kääntäminen on*. Helsinki, Gaudeamus, 1986. 132. p.
  14. NIELSEN, Marianne Lykke. *The word association method. A gateway to work-task based retrieval*. Doctoral dissertation. Åbo, Åbo Akademi University Press, 2002. 342. p.
  15. JORNA, Kerstin – DAVIES, Sylvie: „Multilingual thesauri for the modern world. „No ideal solution?“ = *Journal of Documentation*, 57. évf. 2001. 2. sz. 284–295. p.
  16. VON HAHN, Walther – ANGELOVA, Galia. *Combining terminology, lexical semantics and knowledge representation in Machine Aided Translation*. In: *TKE'96, Terminology and Knowledge Engineering*. Frankfurt, INDEKS-Verlag, 1996. 304–331 p.
  17. SCMITZ-ESSER, Winfried: *Thesaurus and beyond. An advanced formula for linguistic engineering and information retrieval* = *Knowledge Organization*, 26. vol. 1999. 1. no. 10–22. p.
  18. GRADDOL, David. *Will English be enough? In: Where are we going with languages?* Ed. Moys, A. London, Nuffield Foundation, 1998. 24–33. p.
  19. CABRÉ, Maria Teresa: *A hypermedia project on terminology*. In: *TKE'96, Terminology and Knowledge Engineering*. Frankfurt, INDEKS-Verlag, 1996. 424–430. p.
  20. NEELAMEGHAN, Arashanipalai. – LALITHA, S. K.: *Multilingual thesaurus and interoperability* = *Journal of Library & Information Technology*, 33. vol. 2013. 4. no. 289–294 p.
  21. AITCHISON, Jean – CLARKE, Stella J. Dextre: *The thesaurus. A historical viewpoint, with a look to the future* = *Cataloging & Classification Quarterly*, 37. vol. 2004. 3–4. no. 5–21. p.
  22. ISO 25964: *The international standard for thesauri and interoperability with other vocabularies*. International Organization for Standardization. 2011. <http://www.iso.org/schemas/iso25964/#summary> (Az utolsó megtekintés ideje: 2015. október 8.)

*Beérkezett: 2015. október 19.*