

## Metaforaazonosítás magyar nyelvű szövegekben: egy módszer adaptálásáról

### 1. Bevezetés

#### Kérdések a metafora körül

Ha azokat a problémaköröket szeretnénk röviden áttekinteni, amelyek mentén napjaink metaforakutatása kibontakozik, mindenekelőtt a következő kérdés adódik: létezik-e egyáltalán a metafora kategóriája. Az egyik szélsőséges válasz a relevanciaelmélet keretei között fogalmazódott meg: „mi továbbra is csak a szó szerinti, laza és hiperbolikus értelmezéseket magában foglaló kontinuum egyik felén elhelyezkedő esetek összességéként látjuk a metaforát” – jelenti ki Dan Sperber és Deirdre Wilson (2013: 314). Tézisüket a következőképpen részletezik: „A metaforikus értelmezésekhez pontosan ugyanolyan módon jutunk el, mint a többi értelmezéshez. Nincsen specifikusan a metaforákhoz kapcsolódó mechanizmus, vagy különleges, csak rájuk jellemző általánosítás. Más szavakkal: a nyelvi metafora nem természetes fajta és a „metafora” a nyelvi kommunikáció tanulmányozásának sem elméleti jelentőséggel bíró fogalma.” Ezzel szemben a kognitív nyelvészet fogalmi metaforaelmélete éppen a metaforában látja a megismerés alapvető tényezőjét: „A metafora általános elmélete – írja George Lakoff (2006/1993: 185–6) – a tartományok közötti leképezések jellemzésével adható meg. És a folyamatban a mindennapi absztrakt fogalmak mint idő, állapotok, változás, okozás és cél szintén metaforikusnak bizonyulnak. Ennek eredményeképpen a metafora (vagyis a tartományok közötti leképezés) teljes mértékben központi a hétköznapi természetes nyelv jelentéstanában, az irodalmi metafora tanulmányozása pedig a mindennapi metafora tanulmányozásának kiterjesztése. A mindennapi metafora tartományközi leképezések ezreinek hatalmas rendszerével jellemezhető, és ezt a rendszert alkalmazzuk az újszerű metaforában is.”

A figyelmes olvasó számára nem csupán az tűnhet fel, hogy a két idézett álláspont ellentétes a metafora létét illetően, hanem az is, mennyire rímelnék egymással a metafora létmódját tekintve. Noha egyik tagadja, másik állítja a metafora létét, közösek abban, hogy azt nem a nyelv szerkezetének tekintik. Sperber és Wilson egyértelműsíti, hogy a metafora kutatása semmilyen belátással nem gazdagítja a természetes nyelvek leírását, és bár Lakoff a szemantika tárgykörébe sorolja a jelenséget, kimondja ugyanakkor azt is, hogy „a metafora helye egyáltalán nem a nyelv, hanem az a mód, ahogyan konceptualizálunk egy mentális tartományt egy másikkal megadva” (Lakoff 2006/1993: 185). Amennyiben tehát a kognitív nyelvészet eredményeit elismerve elfogadjuk is, hogy a metafora mindent átható jellegzetessége az emberi megismerésnek, továbbra is kérdéses marad,

van-e egyáltalán a nyelvnek szerepe a metaforikus jelentés kialakulásában, és ha igen, mi ez a szerep.<sup>1</sup>

A legmarkánsabban mindez idáig talán Gerard Steen (2007: 108) fogalmazta meg a metafora fogalmi és nyelvi aspektusának megkülönböztetését: az adatgyűjtés módszereit tekintve a mentális reprezentáció feltárása (*metaphor in thought*) a pszichológia diszciplináris területére tartozik, míg a nyelvi szerveződés vizsgálata (*metaphor in language*) a nyelvészet feladata. Az elmúlt bő évtized korpusznyelvészeti kutatásai valóban azt mutatják, hogy a nyelvnek – a fogalmi metafora-elmélet alaptéziséhez képest – van szerepe a metaforák használatában. Mind a diskurzus szerveződése, mind az egyes szintaktikai szerkezetek szintjén „gyakori és feltehetően szabályos különbségek vannak ugyanazon szó metaforikus és szó szerinti használata között”, továbbá számos metaforikus használatba vétel „grammatikailag korlátozottnak” tűnik (l. Deignan 2005: 145). E korlátozottság (illetve a rendszerszerű különbség) mindenekelőtt a kollokálódás specifikus mintázataiban, illetve esetenként az idiómaszerű előfordulásban ragadható meg (Deignan 2005: 155), amelyet a magyar nyelvre irányuló kutatások is igazolni látszanak (Babarczy–Simon 2012; Simon 2018).

Ezen a ponton válik lényegessé a vizsgálat módszertana. Immár nem az a kérdés, hogy a természetes nyelvben létezik-e a metafora kategóriája (igen, hiszen sajátos mintázatokba rendeződnek a metaforikus jelentések), és nem is az, hogy ennek a kategóriának a kutatása gazdagíthatja-e a nyelvről való ismereteinket (igen, hiszen ráirányítja a figyelmet az együttes előfordulások gyakoriságának a jelentőségére a rendszer alakulásában). A legfontosabb továbbvezető kérdés az lesz, hogyan, milyen módszerrel vizsgáljuk a nyelvi metaforákat. Ahhoz ugyanis, hogy a metaforikus nyelvhasználat jellemző sablonjait feltárjuk, azonosítani kell ezeket a korpuszban. Erre pedig a manuális kiválogatáson túl a forrástartományra utaló kifejezések keresése (pl. Deignan, Hanks), a céltartományra utaló szerkezetek gyűjtése (pl. Stefanowitsch), illetve olyan mondatok szűrése kínálkozik, amelyekben mindkét tartomány explikálódik nyelvileg (pl. Martin; a módszerekről összefoglalóan l. Stefanowitsch 2006: 2–5). Ezek a bevett módszerek korpuszalapúnak tekinthetők, amelyek tehát a metafora elméleti modelljét (ez esetben a tartományközi leképezést) veszik alapul az adatgyűjtésben (Deignan 2008: 155; a korpuszalapúságról l. Biber 2010; a metafora kapcsán l. Simon 2018). Ezért bár a metaforikus jelentések nyelvi szerveződéséről nyújtanak adatokat, e jelentés fogalmi természetét előfeltételezik a vizsgálat során, és így csak azokra a találatokra terjed ki a korpuszalapú elemzés, amelyek megfeleltethetők valamely absztrakt forrás- vagy céltartomány összetevőinek.

Jól látható, hogy a metafora korpuszalapú kutatásának egyik megoldandó problémája a metaforikus kifejezések azonosításában rejlik. Ez a folyamat komoly kihívást támaszt a nyelvtechnológia számára, mert a standard korpuszon be-tanított, fogalmi leképezéseken alapuló metaforakereső program nem éri el a kézi

<sup>1</sup> Természetesen hasonlóan nagy jelentőségű kérdés fogalmazható meg a metaforikus jelentések neurális reprezentálásáról és agyi feldolgozásáról (l. Forgács 2013), mivel azonban a jelen kutatás a metaforikus kifejezések szövegbeli azonosításának problémájára irányul, a neurolingvisztikai eredmények bemutatásától ezúttal eltekintünk.

kiválogatás határfokát (Babarczy–Simon 2012: 237). További nehézséget támaszt a fogalmi tartományok komplex szimbolizálódása a nyelvi kifejezésekben, például a *fel van dobódva* szerkezet egyszerre tartalmazza a FELFELE IRÁNYULÁS és a MOZGÁS forrástartományait (komplexen megjelenítve az igekötő és az igető együttesében), miközben a céltartományt (POZITÍV ÉRZELMI ÁLLAPOT) a szerkezet egésze jeleníti meg, utalva egyszersmind (az igeképzővel) az állapot külső ok általi kiváltódására is. Vagyis a tartományok közötti leképezés modellje alapján nem minden ponton világos még ebben az egyszerű kifejezésben sem, melyik összetevő milyen módon/mértékben járul hozzá a metaforikus minősítéshez.

Olyan azonosítási módszerre van tehát szükség, amely alkalmas egy szöveg vagy szövegrészlet körültekintő nyelvi elemzésére, és ezzel összekapcsolva mutatja fel eredményként a szöveg(részlet) metaforikus jelentéseinek a nyelvi szerkezeteit. Mindezt úgy, hogy a szerkezetek elemzését nem egyik vagy másik kitüntetett metaforaelméleti kiindulópontból végzi el, hanem lehetőleg semlegesen, minimális előfeltevéseket érvényesítve a metafora mibenlétééről.

## 2. A metaforaazonosítás módszere: az MIPVU

Ebben a részben röviden bemutatjuk a metaforikus kifejezések azonosításához alapul vett eljárást, amely természetesen maga sem előzmények nélküli: a kétezres évek második felében a Pragglejaz Csoport alakított ki először módszert a metafora korpuszbeli annotálásához, amely az MIP<sup>2</sup> (*Metaphor Identification Process*) nevet kapta. Ugyanakkor a módszer gyakorlati alkalmazása nagy mennyiségű nyelvi anyagon már a kezdetektől fogva finomításokat tett szükségessé; ezeket Gerard Steen munkatársaival együtt végezte el, új eljárás keretei közé szervezve az egyes innovációkat – ezzel állt elő az évtized végére az MIPVU (*Metaphor Identification Process Vrije Universiteit*, a két eljárásrend részletes összevetésére l. Steen et al. 2011: 5–20).

A metaforaazonosítás módszerének kidolgozását két célkitűzés motiválja: (i) annotálási eljárást – és ezzel a pontos mérés lehetőségét – kínálni a metafora kutatásának elméleti kérdéseire; (ii) a mindenkori nyelvhasználó nyelvi tudásához igazítani a metaforikus kifejezések azonosítását. Az első szempont megvalósítását szolgálja, hogy a metaforát a lehető legtágabban, indirekt jelentésalkotás eseteként határozzák meg, amelyben az elsődleges és a kontextuális jelentés közötti közvetett kapcsolatot a hasonlóság (tágabban a tartományok közötti leképezés) biztosítja.<sup>3</sup> Steen és munkatársai minimális elméleti előfeltevésekre alapozzák tehát a metafora azonosítását, elkerülve ezzel az előzetes elméleti szűkítés prob-

<sup>2</sup> A módszert hat évig fejlesztették a csoport tagjai: Lynne Cameron, Alan Cienki, Peter Crisp, Alice Deignan, Ray Gibbs, Joe Grady, Kövecses Zoltán, Graham Low, Elena Semino és Gerard Steen (l. Steen et al. 2011: 4).

<sup>3</sup> Ezzel a definícióval teoretikusan szűkítik a metafora kategóriáját a hasonlóság eseteire, de mint látni fogjuk, lexikális szemantikai keretben, továbbá az eredeti módszer grammatikai szemléletű kiterjesztésével a módszer alkalmas a metaforikus kifejezések tág körének az azonosítására.

lémáját és lehetővé téve az elméletsemleges mérések megvalósítását. A második kritériumnak az eljárás szótáralapú, lexikális szemantikai szemlélete felel meg: a jelentés metaforikusságát (indirekt jellegét) mindig a lexikális egységek szintjén elemzik (nem alkalmazva grammatikai analízist), mégpedig az angol nyelv korpuszalapú (tehát az aktuális nyelvhasználat tényeire épülő) nagyszótári alapján.<sup>4</sup> Lehetőség szerint kizárják tehát az azonosításból az elemzői intuíciót, arra törekedve, hogy ott jelöljenek metaforikusságot a vizsgált szövegekben, ahol napjaink angol anyanyelvű nyelvhasználója is tenné az eljárás ismeretében.

A módszer a következő alapvető lépésekből áll (Steen et al. 2011: 25–6 alapján).

1. Tagoljuk a vizsgált szöveget lexikális egységekre.
2. Ha a szó aktuális használata indirekt jellegű, és ez potenciálisan magyarázható az alapjelentésből eredő tartományközi leképezésekkel, azonosítsuk metaforikusként.
3. Ha a kifejezés közvetlen jelentésében fordul elő a szövegben, és ez potenciálisan magyarázható a szöveg egyik alapvetőbb referensére vagy topikjára irányuló tartományközi leképezésekkel, azonosítsuk a kifejezést közvetlen (direkt) metaforaként.
4. Ha egy szó lexikogrammatikai helyettesítés céljából jelenik meg a szövegben (mint pl. a harmadik személyű személyes névmás), vagy ha elliptikus egy szöveghely, ugyanakkor ezek az esetek potenciális tartományközi leképezések direkt vagy indirekt eseteiként magyarázhatók, azonosítsuk a szót implicit metaforaként.
5. Ha egy kifejezés a tartományközi leképezések jelölőjeként funkcionál, azonosítsuk metaforázászlóként.
6. Ha a kifejezés újszerű formációnak tekinthető, vizsgáljuk meg komponenseit az előző lépések alapján.

Amilyen kevés és átfogó előfeltevést fogalmaznak meg tehát a módszer kialakítói a metaforikus jelentés elméleti modelljét illetően, olyannyira differenciáltan kezelik a nyelvi megvalósulást. A szószerintiség, a kifejtettség és a metaforizáció jelölése egyaránt az elemzés szempontjaiként jelennek meg, és a módszer arra is alkalmas, hogy komplex kifejezéseket (az összetételektől az idiómáig) is kezelni lehessen vele. Lényeges továbbá, hogy Steen és munkatársai nem azonosítják a metafora jelenségét a szöveg egyes szavaival, kifejezéseivel: az annotálás során alkalmazott terminusuk a *metaforához kapcsolódó szó* (*metaphor related word*, MRW, a szó kategóriájára l. 2.1.), vagyis az MIPVU segítségével azokat a szöveghelyeket jelölhetjük meg, amelyek biztosan közreműködnek a metaforikus jelentés kialakításában. A továbbiakban az eljárást a lépései mentén részletezzük (Steen et al. 2011: 26–42-re alapozva).

<sup>4</sup> Elsődlegesen a The Macmillan English Dictionary for Advanced Learners 2002-es kiadására támaszkodnak (erre a kézikönyvre a továbbiakban szótárként utalunk magunk is a tanulmány jelen egységében), ám kontrolleszközként alkalmazzák a The Longman Dictionary of Contemporary English (2003) kiadványt, és az egyes jelentések közötti kapcsolat árnyalására használják még az Oxford English Dictionary (1974) jelentésleírásait is.

## 2.1. Lexikális egységek a szövegben

Az MIPVU a szóra, pontosabban a szóértékű kifejezések körére összpontosít. Ezek meghatározásához a módszer kidolgozói a BNC (*British National Corpus*) elemzési listájára támaszkodtak: azt tekintik önálló lexikális egységnek, amely POS (*part of speech*) taget kap a BNC-szövegek annotálása során. Következésképpen sem a szóalaknál kisebb nyelvi elemeket (lemmák, morfémák), sem a szintaktikai szerkezeteket nem kezeli az eljárás. A többszavas kifejezések esetén is igazodtak a kidolgozók a BNC gyakorlatához, amely tartalmazza ezek véges elemszámú listáját (*polyword* státus alá sorolva). Még így is felmerültek azonban sajátos esetek. A vonzatos igéket (a BNC eljárásával szemben) egy lexikális egységnek tekintik, megkülönböztetve ezeket a kötött prepozícióval járó igei szerkezetektől,<sup>5</sup> ugyanakkor e két kategória elhatárolásában is a korpuszbeli annotálás a kiindulópont, a szótár csak az ellenőrzést segíti. A szóösszetételeket az eljárás szerint egy lexikális egységnek kell tekinteni (még kötőjeles írásmód esetén is, ekkor a szótári címszó a döntő). Csak abban az esetben bontható a szerkezet komponenseire, ha új, még nem konvencionális szerkezet (a szótár adataiból kiindulva), illetve ha a szótár kettős hangsúlymintázatot közöl, például: *„nuclear ‘power*. (Vagyis az eredeti módszer egyértelműen elhatárolódik mind a grammatikai konstrukciók, mind az összetételek részletekbe menő elemzésétől, ez az adaptálás során lesz lényeges szempont.) A tulajdonnevek érdemelnek még külön figyelmet: azok minden eleme önálló lexikális egységnek minősül, kivéve, ha a szótár már állandósult egységként kezeli a teljes nevet, minden elemével együtt, például: *Labour Party*, *Pulitzer Prize*.

## 2.2. Az indirektség megállapítása: kontextuális jelentés, alapjelentés

A következő módszertani lépés annak megállapítása, tételezhető-e indirekt jelentés az adott szöveghely kapcsán, és amennyiben igen, milyen jellegű ez. Ennek négy lépése van.

- a) Azonosítsuk a kontextuális jelentést.
- b) Állapítsuk meg, van-e elsődlegesebb alapjelentése a lexikális egységnek.
- c) Vizsgáljuk meg, kellően elkülöníthető/elkülönülő jelentések-e ezek.
- d) Vizsgáljuk meg, kapcsolatba hozhatók-e a jelentések a hasonlóságra alapozva.

Mivel az indirektség meghatározása tekinthető a módszer központi műveletének, célszerű ezeket a lépéseket részletesebben tárgyalni.

<sup>5</sup> Tesztelve a mondatbeli elmozdíthatóságot, a kérdéses komponens beférkőző jellegét (adverbiumok esetén), a kérdő névmás elé helyezhetőséget, a közvetlen tárgyhoz (direct object) képesti szórendi lehetőségeket, továbbá az anaforikus névmási helyettesítés lehetséges mondatbeli pozícióit.

### 2.2.1. A kontextuális jelentés meghatározása

Ez a jelentés lehet újszerű és konvencionális is, az előbbi esetben az elemzői jelentésadás az adekvát művelet, az utóbbi esetben azonban a szótári szócikkből lehet kiindulni. (Az általunk vizsgált esetek között az első esetre kevés példa volt.) A kontextuális jelentés meghatározása akkor igazán problematikus, ha a kifejezés töredékes (szóbeli diskurzusok esetében), ha nincs elegendő kontextuális információ a meghatározáshoz (pl. megfigyelhető irányjelentés, de nem egyértelmű, hogy fizikai, metonimikus vagy metaforikus értelemben jelenik meg), illetve ha olyan specifikus terminusokról van szó, amelyek nem találhatók meg egy általános szótárban. Míg az első esetben a lexikális egységet ki kell zárni az elemzésből (*discarded from metaphor analysis*, DfMA), addig a másik két esetben bent kell hagyni, de jelölni kell a kétséges státust (*when in doubt, leave it in* – WIDLII címkével ellátva). Végül külön kezelik Steen és munkatársai a megszemélyesítéseket, amelyek lehetnek metonimikusak, ám ekkor is egyértelműen metaforikusnak jelölik a lexikális egységet.

### 2.2.2. Az alapjelentés meghatározása

E lépés kiindulópontja és legfőbb elméleti előfeltevése, hogy az alapjelentés mindig konkrét, specifikus és humánorientált jelentés. (Tehát nem lehet egy kifejezés jelentése alapjelentés, ha nem szerepel a szótárban.) A kategória előtagja (alap-) ugyanakkor nem a történetileg elsődleges jelentésre utal, hanem a nyelvhasználó közösség számára alapbeállítású (default) jelentésre, amelyet a szótár is első jelentésként tüntet fel. Az alapjelentés megállapításának számos kikötése van, a legfontosabbak a szófajiságot, a funkcionálást és a grammatikai konstrukciókat érintik. Az MIPVU értelmében az alapjelentés a szófajisághoz tartozik: amennyiben egy lexikális egység több szófaji kategória tagja, ezek mind önálló alapjelentésként értelmezendők, és a kontextuális jelentés szófajiságához igazítva kell egyiküket kiválasztani. Ha a lexikális egységnek több funkciója is van (pl. az igealak segédigeként is megjelenik), azokat is külön alapjelentésekként kell kezelni. És amennyiben több eltérő szerkezetben is megjelenik (pl. tranzitív és intranszítív argumentumkerete is van egy igének, vagy megszámlálható és megszámlálhatatlan jelentésben is használatos egy főnév), azok is külön alapjelentéseknek számítanak. A kiindulópont minden esetben a szótári jelentésadás, illetve az abban meghatározott, elkülönített alkategóriák. A kiindulóponthoz képest előfordulhat olyan lexikális egység is, amelynek az alapjelentése általánosabb és konkrétabb, míg a kontextuális jelentése specifikusabb, de absztraktabb (ilyen pl. a *father* főnév általános 'apa' és specifikus 'Isten' jelentése), ekkor is a konkrétabb jelentés lesz az alapjelentés. Terminusoknál gyakori, hogy nagyon specifikus az alapjelentés, amelyhez viszonyítva az absztraktabb kontextuális jelentés lehet metaforikus. Végül abban az esetben, ha mind az alap-, mind a kontextuális jelentés egyformán specifikus/absztrakt, a jelentés eredete dönti el, melyik az alapjelentés (ilyen esetek a hangutánzó kifejezések, amelyeknél az eredeti prototipikus referens gyakran

fizikai tárgy vagy állat, így ezek hangadása lesz az alapjelentés, nem pedig a hasonlóan konkrét emberi hang).<sup>6</sup>

### 2.2.3. Az egyes jelentések elkülöníthetősége

Alapesetben a szótári szócikk számokkal elkülönített jelentéseire alapul az elemzés, így ha a két azonosított jelentés külön szám alatt jelenik meg, a jelentések kellő mértékben elkülönülnek. Ha azonban a szótárban csak egy jelentés van, akkor az tekintendő alapjelentésnek, és minden attól való kontextuális elmozdulást (jelentéskiterjesztést) kellően elkülöníthető esetnek kell tekinteni.

### 2.2.4. A jelentések közötti kapcsolat elemzése

A fő kérdés ennél a lépésnél, hogy megállapítható-e hasonlóságon (tartományközi leképezéseken) nyugvó kapcsolat az elkülönített jelentések között. A hasonlóság lehet funkcionális vagy szemléletbeli, és lehet sematikus vagy gazdagon kidolgozott, ezek a fokozati és minőségbeli különbségek azonban az annotálás során nem lényegesek. Problematikusabb azonban, ha a kontextuális jelentés homályos, absztrakt, szemantikailag kiüresedett (kifakult, *bleached*): ekkor felmerül a metaforikusság, a jelentés általánossá válása (generalizáció) azonban nem tekinthető annak. (Pl. az *appeal* 'folyamodik' jogi jelentéséből általánosuló 'fordul vkihez, kér' jelentés nem metaforikus.) Amennyiben nem állapítható meg egyértelműen a hasonlóság a két jelentés között, az etimológia válik mérvadóvá: az Oxford English Dictionary alapján ellenőrizték a módszer kidolgozóit, van-e közös előzménye a két jelentésnek, és amennyiben igen, akkor két önálló (származtatott) alapjelentésként kezelték a két jelentést. Végül, ha azt találjuk, hogy a kapcsolat alapvetően metonimikus, ez Steenék szerint még nem zárja ki a metaforikusság lehetőségét, ezért ezeket az eseteket nem zárták ki automatikusan. (E tekintetben az adaptálás során szigorúbban jártunk el, lásd 3.1.) Összegezve: ha a két jelentés közötti kapcsolat nem generalizáció, nem etimologikus (közös előzményre visszautaló), és nem kizárólag metonimikus, akkor minősíthető a lexikális egység metaforához kapcsolódó szóként (MRW-ként).

## 2.3. A direkt metaforikus kifejezések

Azokat a kifejezéseket, amelyek alapjelentésükben, de a szöveg jelentésében valamilyen analógia alapjaként jelennek meg, direkt metaforákként kezeli az MIPVU módszere. Ekkor a lexikális egység referense és a szöveg(rész) topikja, központi referense, témája között tartományközi leképezések hozhatók létre. Ennek megál-

<sup>6</sup> Amennyiben az egyes szótárak jelentésadása között eltérés mutatkozik, Steen és munkatársai a leginkább gazdag jelentésleírásra hagytakoztak.

lapításához azonosítani kell az adott szövegrész központi témáját, illetve a téma-váltásokat, majd vizsgálni kell, hogy az újonnan bevezetett topik beilleszthető-e hasonlóság alapján egy átfogó referenciális keretbe. Amennyiben igen, és az ösz-szevetés nem szó szerinti (pl. nem a konkrét, fizikai hasonlóságra utal, mint az *olyan vagy, mint az apád* szerkezetben), akkor metaforához kapcsolódó direkt szó-ként kell annotálni az adott kifejezést.

#### 2.4. Implicit metaforikus kifejezések

Ebbe a kategóriába azok a kifejezések tartoznak, amelyek az alapjelentéssel való ösz-szevetés mentén nem minősülnének metaforikusnak, ám azáltal, hogy koreferens kapcsolatban állnak egy másik metaforikus kifejezéssel, vagy ellipszisként metaforikusan egészíthetők ki, mégis kapcsolatba kerülnek a szöveg metaforikus-ságával. Ilyen esetek példaként a következők:<sup>7</sup> *természetesen nekivágni egy ilyen lépésnek nem jelenti szükségszerűen az azonnali sikeres megvalósítását*, ahol a *megvalósítását* határozott tárgyú inflexiója anaforikusan a *lépésnek* nominálisra utal vissza, amely metaforikus; vagy a [Nem gondolnám, hogy mindenki ignoráns disznó,] *de ő az* [mármint ignoráns disznó] diskurzusfordulóban a mutató névmás jelentése önmagában nem metaforikus (különösen, hogy diskurzusdeixisként is értelmezhető), ám az a szövegelőzmény, amelyre (rámutatva) utal, annak minősül. Ezek az esetek az implicit metaforikusság címkéjét kapják.

#### 2.5. A metaforikusság jelölői és új szerkezetei

Egyes kifejezések funkciója éppen az, hogy tartományok közötti leképezéseket jelöljenek explicit módon. Ilyenek a hasonlatok nyelvi fordulatai (*úgy, mint, több, kevesebb, vedd össze, hasonlít, hasonlóság, ugyanaz, analóg, analógia*), illetve a komplex mentális predikátumok (*tekintsük valamilyennek, lássuk olyannak, képzeld el, gondol, beszél, viselkedik mint*), vagyis a metafora „zászlói” (*metaphor flags*, MFlag). Kizárták azonban ebből a kategóriából a generikusság jelölőit (*valamilyen fajta, valamilyen típusú*) és a topik fogalmi tartományának jelölését (amilyen az *intellektuális stagnálás* szerkezetben a jelző). A módszer áttekintését érdemes azzal zárni, hogy az újszerű kifejezések esetén Steen és munkatársai is javasolják a komponensekre bontást, vagyis annak meghatározását, grammatikailag meddig terjed ki a szerkezet, és összetevői mennyiben tekinthetők önmaguk is metaforához kapcsolódó kifejezésnek. (Az adaptálás során ezt az elvet terjesztet-tük ki a magyar nyelv tipológiai sajátosságai miatt.)

<sup>7</sup> A példák Steen et al. (2011) magyarra fordított szerkezetei. A fordítás egyúttal ráirányítja a figyelmet a magyar nyelv agglutináló jellegéből adódó sajátosságokra is.



### 3. Alkalmazás és alakítás: metaforaazonosítás magyar nyelvű szövegekben

Az MIPVU egyfelől következetesen kidolgozott módszertani alapokat kínál bármely szöveg metaforikus szerkezeteinek az azonosításához és annotálásához. Másfelől azonban minden rugalmassága mellett is felmerülnek a metódus korlátai, hiszen alapvetően angol és holland nyelvű szövegek elemzésére alakították ki, továbbá a hasonlóság szemantikai tényezőjének középpontba helyezésével Steen és munkatársai kiterjesztették a metaforikusság kategóriáját, beleértve a hasonlatokat és helyet hagyva benne a metonimikus szerkezeteknek is. Amennyiben tehát magyar nyelvű szövegekre szeretnénk alkalmazni a módszert, bizonyos pontokon módosítanunk kell, illetve ki kell egészítenünk az eljárást, tekintettel a magyar nyelv agglutináló jellegére. Mindezekon túl akkor is alakítani kell a felkínált módszertani kereteken, ha az eredeti célkitűzésekhez képest szigorúbban kívánjuk kijelölni a vizsgált szövegek metaforikus szerkezeteit, azaz pontosabban szeretnénk azonosítani a szövegben a potenciális metaforákat. A továbbiakban az MIPVU adaptálása során felmerülő elméleti (3.1) és módszertani (3.2) problémákat, illetve az azokra javasolt megoldásainkat vesszük sorra, majd tömören összefoglaljuk az általunk alkalmazott eljárást (3.3). Ezt követően (3.4) a próbaannotálások során feldolgozott kutatói korpuszból mutatunk be tipikus és kevésbé tipikus, sőt nehezen kezelhető példákat, hogy rövid elemzésekkel segítsük a módszer iránt érdeklődő olvasókat az önálló alkalmazásban.

#### 3.1. Az adaptálás elméleti kérdései

Ezeket a kérdéseket mindenekelőtt az veti fel, hogy a magyar nyelv gazdag morfológiája lehetővé teszi a szóalakon belüli metaforizálás megvalósítását (gondoljunk az *észben tart* vagy a *fejébe vett valamit* kifejezésekre, ahol a TARTÁLY kognitív modellje az esetragon keresztül motiválja a metaforikus jelentést). Magyar nyelvű szövegeken végzett metaforaazonosítási vizsgálatok során tehát nem követhetjük a lexikális egységekre alapozott eljárást, ha teljes pontossággal fel akarjuk mérni a metaforikus kifejezéseket. Ezért saját elemzéseink során morfológiai elemzést is végeztünk, mégpedig az e-magyar online digitális nyelvtechnológiai elemzőfelülettel (Váradi et al. 2017).<sup>8</sup> Azokban az esetekben, amikor a szótóhoz szófajtartó képző társult, azt nem tekintettük az elemzés önálló egységének, hiszen az MIPVU módszere szerint a szófaji határokon belül kell megállapítani az elsődleges és a kontextuális jelentést (ugyanakkor a morfológiai szegmentálást ezekben az esetekben is elvégeztük a későbbi adatbázis-építés és egy tervezett online lekérdezőfelület könnyebb kialakítása céljából). Természetesen a lexikalizálódott szerkezetek esetében (mint amilyen az összefüggő melléknév) a szóelemekre bontást elhagytuk, hiszen a metaforaazonosításnak nem célja etimológiai, grammatikalizálódási ösvények nyomon követése.

<sup>8</sup> <https://e-magyar.hu/hu/parser> (Utolsó elérés: 2018. november 19.)

Az inflexiós morfémákhoz hasonlóan a szóösszetételeket is komponenseikre bontva elemeztük. Ezt a döntést az motiválta, hogy az összetételi tagok külön is metaforizálódhatnak, noha egymástól korántsem függetlenül.

- (1) Apámat és mostohaanyámat az irodában találtam: szűk, akváriumszerűen kivilágított üvegalicka [...].

Az (1) példában (amely Kertész Imre: Sorstalanság című regényének részlete) az *üvegalicka* kifejezés érdemel ezúttal figyelmet: míg az előtag alapjelentésével ('homok, mészkő, szóda megolvasztott keverékéből gyártott rideg, kemény [átlátszó] anyag') járul hozzá a szerkezet egészéhez, addig az utótag esetében már egyértelműen megkülönböztethető az alapjelentés ('madár v. kisebb állat fogságban tartására való, drótból, vesszőből stb. készült rácsos kis [hordozható] házikó') és a kontextuális jelentés ('szűk fülke'), amelyek között hasonlóságon alapuló kiterjesztés a kapcsolat. Az utótag tehát metaforikus, míg az előtag nem. Mindezek alapján az összetételi tagokat külön elemzési egységekként vettük fel, kivéve a teljes mértékben lexikalizálódott szerkezeteket (mint amilyen a *szerecsmalac*).

További problematikus pontnak bizonyult az igekötők kezelése: sok esetben ugyanis még napjaink nyelvhasználói számára is hozzáférhető az igekötők elsődleges irányjelentése, amely gyakori kiindulópontja a metaforizálódásnak (erre példa a *kiereszti a gőzt* vagy a *lecsendesít* kifejezések). Ugyanakkor számos olyan eset is előfordult (különösen a *meg-* vagy az *el-* esetében, l. *megír;* *elmosogat*), amikor az igekötő irányjelentése a mai beszélő perspektívájából már nem aktív, és a komponens fő funkciója az aspektualitás kifejezésében van. Elemzéseink során ezért az igekötőket egy kontinuum mentén vizsgáltuk, amelynek egyik végpontja az irányjelentés szerkezeten belüli aktiváltsága (ekkor a metaforikusság potencialitása fennáll), a másik végpontja a tisztán aspektuális funkcionálás irányjelentés nélkül (ekkor nem beszélhetünk metaforikusságról). A két végpont között pedig többé vagy kevésbé aktív az irányjelentés, amely relatív metaforikusságot von maga után (ezt az annotálás során is jelöltük, lásd erről később).

A nyelvi rendszerből adódó adaptálási mozzanatok mellett döntéseket kellett hoznunk a metaforikusság pontosabb körülhatárolása érdekében is, elsősorban metafora és a metonímia megkülönböztetése kapcsán. Amennyiben az alapjelentés és a kontextuális jelentés elkülöníthetőnek bizonyult, ám a két jelentés kapcsolata metonimikusan is magyarázható, az eredeti módszerrel szemben nem jelöltük a metaforikusság potencialitását a szerkezeten, még ha a metonimikusság nem is zárja ki egy esetleges metaforizálódás lehetőségét. Erre példa a *piacra ment* szerkezet, ahol a nominális fő elsődleges jelentése az ÉKsz.<sup>2</sup> szerint 'napi szükségleti cikkeknek, kül. élelmiszereknek erre kijelölt helyen, téren folyó rendszeres adásvétele', ez pedig metonimikus kapcsolatban áll a kontextuális 'piactér' jelentéssel, ezért a szerkezet egésze nem – miként az esetrag sem – metaforizálódik elemzésünkben.

Ha egy szóalak inflexiós morfémája minősíthető metaforikusnak, ez természetesen nem hagyja érintetlenül a szótő jelentését sem, másként fogalmazva, míg az észben tart nominális kompozitumában az esetrag tekinthető egyértel-

műen metaforizáló komponensnek, ez egyúttal a tövel kifejezett főnévi jelentés metaforizálódását is maga után vonja. Ezért külön címkével láttuk el a metaforikus komponenseket: azokat a szerkezettagokat, amelyek nem kezdeményezik a metaforizálódást, de kiterjed rájuk annak hatása. Ugyanezen megfontolásból külön kezeltük a birtokos szerkezetekben megvalósuló metaforikusságot is. A *lépcső tövében* szerkezet elemzése során például a *tövében* nominális tekinthető elsődlegesen metaforikus szerkezetnek, hiszen a *tő* elsődleges jelentése 'növény szárának a földben, ill. közvetlenül a föld felett levő része', ám ebben a példában kontextuális jelentése metaforikus ('építmények, képződmények stb. legalsó része', saját jelentésadás). Ez ugyanakkor maga után vonja a *lépcső* metaforizálódását is, amely a szerkezet egészében bontakozik ki, a birtokviszony referenciapont-szerkezetként történő értelmezése következtében (Langacker 2008: 83–5). Mivel a célpont metaforikussága kihat a referenciapont jelentésére is, indokolt annak metaforikus komponensként történő jelölése.

Szerkezeti szempontból leginkább kiterjedtnek azok az esetek tekinthetők, amelyekben egy lehorgonyzott igealak metaforizálódik, ez ugyanis az ige bővít-ményeinek egy részére is kiterjesztheti a metaforizálódást. A *kapcsolatot teremthet* témájának rokonsága szerkezetben példának okáért a *teremt* igető elsődlegese jelentése ('<Isten> a semmiből létrehoz') kellően specifikus ahhoz, hogy megkülönböztessük az általánosabb kontextuális jelentéstől ('<helyzetet, állapot> (gyorsan) előidéz, létrehoz'), és metaforikusságot jelöljünk az igealakon. Mivel azonban a *teremt* igetőnek a Magyar igei szerkezetek szótára alapján (Sass et al. 2011, MISz; az adaptálás során alkalmazott segédeszközökről részletesen a 3.2. szakaszban lesz szó) jellemző szerkezete a KAPCSOLAT -t -vAl, a *kapcsolatot* nominálisnál is jelölni kell a metaforizálódás szerkezetben belüli kiterjedését. Lényeges azonban, hogy nem minden bővítményre feltétlenül érvényes ez a megállapítás, ugyanis az általunk konstruált *kapcsolatot teremtett azonnal az intézménnyel* szerkezet határozószója nem része a konstrukciónak, így metaforikussága nem áll fenn. Ezért az MIPVU eljárását ezen a ponton azzal bővítettük, hogy szótár segítségével megvizsgáltuk, vajon az adott bővítmény jellemző argumentuma, gyakori bővítménye, azaz konstrukciós összetevője-e az igei szerkezetnek, és amennyiben igen, úgy az argumentumon is jelöltük a metaforikusságot.

Végül a próbaannotálás során felmerültek olyan esetek is, amelyek egyik-másik összetevője egyértelműen metaforikus, ugyanakkor a szerkezet egésze idiomatikus, állandósult lexikális egységnek tekinthető. Erre példa a *táncot jár* szerkezet: igei komponense az alapjelentéstől megkülönböztethető, és hasonlósággal ahhoz kapcsolható, azaz metaforikus kifejezés, ám ez a metaforikusság közvetlen összefüggésben áll az állandósult szókapcsolat egészével, hiszen azon kívül a *jár* igealak nem metaforikus. Következésképpen megkülönböztettük a metaforizálódás tipikus eseteitől az idiomatikus előfordulásokat, s ezeket külön címkével láttuk el.

Részben már módszertani jellegű az a döntésünk, hogy nem alkalmaztuk a WIDLII-címke magyar megfelelőjét, azaz a kétséges, nehezen megmagyarázható eseteket (mint amilyenek a lexikalizálódott igekötők, illetve azok aspektuális jelentése) nem jelöltük metaforikusnak. Csupán azoknál a kifejezéseknél azonosí-

tottunk tehát metaforikusságot, amelyek esetében a komponensek egyértelműsíthetők voltak, továbbá az alapjelentés a kontextuális jelentés közötti hasonlóságon alapuló viszony mellett a kutatócsoport tagjai egyöntetűen érvelni tudtak. (Ez természetesen nem jelenti, hogy ne szembesültünk volna problémás esetekkel, l. 3.4.) Mindebből az is következik, hogy a metonimikus jelentéseket – noha nem zárják ki a metaforizálódást – nem jelöltük metaforához kapcsolódó kifejezéseként.

## 3.2. Az adaptálás módszertana

A magyar nyelvre adaptálandó metaforaazonosítási módszer kialakításakor a fent részletesen ismertetett elméleti döntések meghozatala párhuzamosan zajlott a próbaannotálások elvégzésével, a metaforikus kifejezések magyar nyelvű szövegekben történő címkézésével. A továbbiakban azt mutatjuk be részletesen, hogy az elméleti döntések meghozatalával összhangban, azokra mintegy folyamatosan reflektálva milyen módszertani döntéseket hoztunk meg a metaforikus kifejezések azonosításakor.

### 3.2.1. A próbaannotálások előkészítése. Az annotátorok munkája

Az adaptált módszert két rövid szöveg, lexikonszócikk (32 és 58 token) és egy regényszövegrészlet (324 token) próbaannotálásával alakítottuk ki. Törekedtünk arra, hogy az elemzési metódus finomításához feltételezett egyszerűség és kezelhetőség miatt a saját stílustulajdonításunk alapján a semleges stílusminőséghez közelítő szövegeket válasszunk, ezért kézenfekvőnek tűnt egy ismeretterjesztő lexikonszócikk (JNÜSz.), illetve fogalomtári meghatározás (MIK) részletének a kijelölése. A második fázisban elemzett szöveg ugyan ez előzőektől eltérően szépirodalmi (Sorstalanság), stílusa semlegesnek mondható. Ezek a stilisztikai megfontolások azért voltak lényegesek, hogy a módszertani adaptálás első fázisában ne szembesüljünk mindjárt a metaforák újszerűségének és poétikusságának a dilemmáival, és lehetőség szerint a konvencionális szótári jelentésadásokból kiindulva határozhassuk meg, mi tekinthető metaforikusnak, ne pedig a szépirodalmi szövegekben előforduló kreatív metaforizációhoz viszonyítsunk.

A két rövid szövegen a kutatásban részt vevő hat annotátor kezdetben egyénileg dolgozott (közös annotálás). A morfológiai elemzés az e-magyar szövegelemző rendszer feldolgozása alapján történt, de az eredmény elfogadása, illetve az egységek, tokenek meghatározása önálló döntés volt minden elemzőnél. Ezután a kialakított egységeket az annotátorok szintén önállóan címkézték. Az egyéni munkát megbeszélés követte, a felvetődő problémák kapcsán konszenzust alakítottunk ki. Ennek során megállapítottuk az alapvető metaforizációs tendenciákat, és rögzítettük az annotációs címkéket (utóbbiakról részletesen l. a 3.2.4. részt).

Az előző munkafolyamatban meghozott döntések tükrében a munkafolyamat második fázisában a kijelölt szöveg annotátoronként felosztott elemzését végeztük el, ekkor tehát a részletből kiválasztott különböző tokeneken különböző

annotátor dolgozott (párhuzamos annotálás). Ebben a fázisban is a problematikus helyek megvitatásával véglegesítettük döntéseinket, valamint amennyiben szükség volt rá, bevezettünk újabb címkéket, hasznosságuk mérlegelése után. A közös munkának ebben a szakaszában hoztunk döntést arról, hogy az idiomatikus és vonzatnak minősíthető előfordulások megjelölésekor szótári adatbázisra támaszkodunk (l. a 3.3. részt), hogy megbizonyosodhassunk róla, hogy bevett kifejezésről vagy pedig egyszeri előfordulásról van szó.

Az annotátorok csoportos megbeszélése elősegítette, hogy vitatható esetekben elkülönítsük a metaforizációt a metonimizációtól és a generalizációtól. Nagyban hozzájárult a közös annotálás szakaszában együttesen elvégzett munka a magyar nyelv sajátosságaiból adódó különböző lehetőségek alapos mérlegeléséhez, a második fázis megbeszélései pedig a szükséges címkék pontosításához. A későbbiekben az adaptálás módszertanának további finomításához tervezzük új, elemzett szöveg alapján az annotátorközi egyezés számítását. Ehhez azonban már online annotálási felületet fogunk alkalmazni, amelyen a kutatócsoport tagjai egymástól függetlenül, de összevethetően végzik majd a kiválasztott szöveg kifejezéseinek a minősítését.

### 3.2.2. A próbaszövegek előkészítése az annotációra

Az annotáció előkészítéséhez és kezdeti megvalósításához, továbbá megelőlegezve a későbbi nyelvtechnológiai felhasználást és egy informatikai lekérdezőfelület létrehozását, Excel-táblázatot alakítottunk ki. A táblázatot úgy építettük fel, hogy a cellatartalmak konzekvens hálózatot alakítsanak ki, amelyben nem okoz problémát az információk rendszerszerű egymáshoz rendelése.

Az első oszlopban a szöveg tokeneit tüntettük fel, megtartva a szövegbeli sorrendet. Az ismétlődő tokeneket értelemszerűen nem töröltük, hiszen a metaforikusság kontextusfüggő. A második oszlop elemekre bontva mutatja be a tokenek szerkezetét. A felosztásban követtük az e-magyar szegmentálását, illetve felhasználtuk annak kódrendszerét. Az inflexiós határokat és az összetételi határokat kötőjellel (-), a szófajváltó képzőket<sup>9</sup> aszteriszkkal (\*) jelöltük. A különbségtételre azért volt szükség, mert a következő oszlopokban a morfológiai kódolást tüntettük fel, méghozzá minden második oszlop felhasználásával, lehetőségét teremtve annak, hogy a lehetséges metaforikusságot közvetlenül a morfológiai kódok mellett tegyük jelöltté. A kötőjel alkalmazása pedig arra adott alkalmat, hogy közvetlenül rendelhessük az információkat a második oszlop celláinak az elemeihez. A szegmentálás oszlopának cellái tehát nem egy, hanem több információt tartalmaznak: az összes szegmentumot. A döntés mellett pusztán az az érv szólt, hogy kevesebb oszlopszámmal, kompaktabb táblázattal könnyebben dolgozhattunk, a konzekvens jelölés miatt azonban a későbbi cellákra bontás nem okozhat majd

<sup>9</sup> A szófajváltó képzők leválasztására elsősorban a morfológiai szerkezet ábrázolása, illetve későbbi vizsgálhatósága miatt volt szükség. A metaforizálódást a szófaji jelentéskategóriák keretein belül értelmezi az eredeti módszer, ezért amennyiben a metaforikus jelentés egyúttal szófajváltással is jár, ennek címkéjét a létrejövő képzett alakhoz rendeltük.

problémát. Ahány kötőjellel elválasztott egység van tehát a szegmentálás oszlopában, annyi morfológiai jelölést, illetve potenciális metaforikusságot tartalmazó oszlop követi azt a táblázatban. Ha az adott szegmentum nem metaforikus, akkor a hozzárendelt oszlop üresen marad. A kidolgozott táblázat felépítését a (2) példa kiemelt tokenjein szemléltetjük, a Sorstalanságból vett részlettel.

(2) nem az érzékenységet akarta érinteni ezzel az *intelmével*

Token	Szegmentálás	Morfológiai kód + metaforikusság				Argumentum	Szófaj	Alapjelentés	Kontextuális jelentés
érezékenységet	érzék*eny*ségemet	[/N]	MKA	[Poss.1Sg]	[acc]		N	1. 'Csekély ingerre is könnyen való reagálás.'	2. 'Vkinek, vminek lelkileg is érzékeny volta.'
érinteni	érint*eni	[/V] [Inf]	MKK			-t (-vAl)	Inf	1. 'Könnyedén hozzáér vkihez, vmihez.'	4. 'Röviden, mellesleg érint vmit.'
intelmével	intelem= intelm-é-vel	[/N]	MKA	[Poss.3Sg]	[Ins]		N	1. 'Erkölcsei célzatú figyelemztetés'	

A morfológiai kódolás és a metaforikusság után felvettünk egy oszlopot a vonzottság jelölésére. A metaforikus szerkezethez való hozzátartozásmegítélésében a Magyar igei szerkezetek szótárára (MISz.) támaszkodtunk.<sup>10</sup> Az argumentumok jelölésére azért volt szükség, mert az igei (igenévi) szerkezetek tagjai bizonyos mértékig együtt metaforizálódnak (a metaforikusság ige felőli kiterjedésével).<sup>11</sup> Ezután a következő oszlopban a token szófaját tüntettük fel az e-magyar besorolása alapján. Végül az elsődleges jelentéseknek, illetve a kontextuális jelentéseknek tartottunk fent egy-egy oszlopot. A cellákba az ÉKsz.<sup>2</sup> definícióit tüntettük fel, illetve a kontextuális jelentéshez saját meghatározás került, amennyiben a szótári jelentésadás nem volt alkalmazható az adott szövegszóra. Kialakítottunk egy segédmunkalapot is: ide felvettük az esetragok és az igekötők elsődleges jelentését szintén az ÉKsz.<sup>2</sup> alapján, amelyet mintegy szótárként tudtunk használni a munka során.

<sup>10</sup> Amennyiben a szótár nem tartalmazta az adott argumentumot gyakori szerkezettagként, az MNSz2-ben is megvizsgáltuk a szerkezet előfordulásait.

<sup>11</sup> Az argumentummal rendelkező igei (igenévi) szerkezetek együttes feldolgozása azért is hasznosnak bizonyult, mert így nem volt szükség az argumentumon megjelenő esetrag metaforikusságával külön foglalkozni (pl. *emlékszik vmire*). Ezekben az esetekben az esetrag metaforikusságát nem, az argumentumét azonban jelöltük.

### 3.2.3. A metaforikus jelentés megállapítása

A metaforaazonosítás szótárakra alapozott módszeréhez első lépésben szükséges volt a segédeszközök kijelölése, amelyek alapján elvégezhetővé vált ez a munkafázis. Az első nehézséget az okozta, hogy – szemben az MIPVU módszerének szótári apparátusával – a magyar nyelv korpuszalapú szótára még nem készült el, így más szótárt kellett bevonnunk a munkafolyamatba. A Magyar értelmező kéziszótár második, átdolgozott kiadására (ÉKsz.<sup>2</sup>) esett a választásunk, mert a magyar nyelvre kidolgozott szótárak közül ez a legfrissebb és a legkomplexebb, valamint a benne feltüntetett szógyakoriságot jelző értékeket a Magyar nemzeti szövegtár alapján állapították meg. A komplex igei szerkezetek feltérképezéséhez és az annotálásban történő feldolgozásához a Magyar igei szerkezetek (MISz.) szótárát használtuk kiindulópontként. Azokban az esetekben, amikor a korpuszbeli igei szerkezet nem szerepelt a szótárban, a Magyar nemzeti szövegtár (MNSz2 v2.0.5.) adataival egészítettük ki az elemzést. A frazémák, idiómák kezeléséhez Szemerkenyi Ágnes Szólások és közmondások (SzK.) című gyűjteményét vettük alapul a feldolgozáskor.

A metaforizáció megállapítását két lépésben végeztük el az eredeti módszer alapján. Az alapjelentés meghatározását minden esetben az ÉKsz.<sup>2</sup> szótár által megadott elsődleges jelentés alapján kíséreltük meg. Ahol azonban indokoltta vált, a szótári elsődleges jelentést felül kellett bírálnunk, és így egy másik, általa feltüntetett jelentést vettünk alapul az annotálás során.<sup>12</sup> A kontextuális jelentés megadását ehhez az annotátori táblázatba rögzített alapjelentéshez képest tettük meg. Ehhez először szintén az ÉKsz.<sup>2</sup> által elkülönített jelentéseket vettük alapul, ám ahol problematikusnak bizonyult a szótárból való jelentés kiválasztása (mert például nem tartalmazta a szótár azt a jelentést, amely megítélésünk szerint abban a kontextusban aktivizálódott), ott a Magyar nemzeti szövegtárból vételeztünk mintát, és ez alapján rendeltük hozzá a kontextuális jelentést a táblázatban.

Ezt követően történt meg annak a megállapítása, hogy valóban a tartományok közötti leképzés történik meg az alapjelentés és a kontextuális jelentés között, vagy pedig más művelet valósul meg (pl. generalizáció vagy metonimizáció); az erről való döntést rendre egyeztetették egymás között az annotátorok. A folyamat során nem arra tettünk kísérletet, hogy megállapítsuk, a szövegek idealizált befogadója hogyan és milyen fokú tudatossággal érzékelné a tartományok közötti leképezést, a metaforikusságot. Azokat az eseteket tettük kizárólagosan jelöltté, amelyek kontextuális jelentésük révén potenciálisan metaforikus jelentésként értelmeződhetnek.

<sup>12</sup> Ilyen volt a *sárga csillagomat* szerkezet nominálisa, amelynél a szótár által felkinált elsődleges jelentés az égitestre vonatkozik, és csak a 2. pont alatt sorolja fel az alakra vonatkozó jelentést ('egy középpontból szabályosan szerteágazó 3-5 vagy több vonalból, egy kör peremén így elhelyezkedő csúcsokból álló ábra, idom'), ugyanakkor a kontextuális jelentés megállapítása során az utóbbiból indultunk ki, hogy ne az égitest és a konvencionálódott csillag alakzat közötti hasonlóság váljon a metafora azonosításának fő kérdésévé.

### 3.2.4. A metaforikus kifejezések azonosítását szolgáló annotációs címkék kialakítása

Az annotációs címkék kialakításakor arra törekedtünk, a címkék kellő árnyaltsággal tükrözzék a metaforikus jelentés létrehozásának összetett lehetőségeit. A morfémaalapú annotációval egyfelől kezelhetővé tudtuk tenni az inflexiók, illetve argumentumok metaforizálódásának a jelenségét is, jelölhetővé téve azt is, hogy egy nagyobb konstrukció elemei közül melyik kezdeményezi a metaforikus jelentést. Másfelől a kialakított annotációs címkék a metaforikus jelentésképzés olyan specifikus eseteire is kiterjednek, mint a direkt és az implicit metaforák, továbbá az MIPVU-ban nem kezelt idiomatikus kifejezésekhez kapcsolódó metaforizáció, valamint a hasonlóságon alapuló indirekt jelentéseket explicitte tevő kifejezések. Ugyanakkor az annotációs kategóriarendszer adaptálásakor és bővítésekor igyekeztünk csupán annyi címkét létrehozni, amennyit a próbaannotálások tapasztalatai valóban indokolttá tettek, szerettük volna elkerülni a kezelhetetlen mennyiségű annotációs címke veszélyét.

Mіндеzen elveket szem előtt tartva a következő annotációs címkéket alakítottuk ki a metaforikus kifejezések annotálására:

- MKK = *metaforához kapcsolódó kifejezés*: azon kifejezés jelölésére szolgál, amelynél különbség van a szótár alapján megállapítható elsődleges jelentés, valamint a kontextuális jelentés között, a jelentések elkülönítésének az alapja pedig valamilyen hasonlóság, azaz tartományok közötti leképezés.
- (3) [...] magamtól is képes vagyok felfogni a bennünket ért *csapás* súlyosságát [...]

A (3) példában a *csapás* szó az MKK címkét kapta, mivel annak a szótárban meghatározott elsődleges jelentése 'erőteljes ütés', az adott szövegkörnyezetben viszont a 'súlyos baj, megpróbáltatás' kontextuális jelentésben fordul elő, amely a szótári jelentésadás szerinti harmadik jelentése a *csapás* szónak.

- MKI = *metaforához kapcsolódó inflexió*: azon inflexiók jelölésére szolgál, amelyek nem valamely konkrét, specifikus, humánorientált jelentésviszony kifejezését, hanem ehhez képest hasonlósági alapon értelmezhető konceptualizációt tesznek lehetővé.
- (4) Azt mondta, hogy ezen a számunkra oly szomorú napon az én „megfelelő viselkedésemre is számíthat”.

A (4) példában az *-on* superessivus rag konkrét, specifikus, humánorientált jelentése szerint valakin vagy valamin való rajtalevést, illetve valaminek a tetején, felületén való elhelyezkedést jelent. Adott kontextusban azonban a szerkezet egy időtartam jelölésére szolgáló egészet felületként konceptualizál, a szimbolizált



folyamat időbeli megvalósulását pedig az *-on* superessivus rag fejezi ki rajtalevésként.

- MKA = *metaforához kapcsolódó argumentum*: azon argumentumok jelölését szolgálja, amelyek igei alaptagjuk metaforizálódása miatt hozzájárulnak a hasonlóságon alapuló jelentésselkülönüléshez, azonban önmaguk nem kezdeményezői a tartományok közötti leképezésnek.

A korábbi (2) példában az érzékenység és az *intelem* tömorfémák is MKA címkét kaptak, mivel ezek a kifejezések a MISz. szerint az MKK címkével ellátott érinteni ige argumentumszerkezetének a részei. Abban az esetben pedig, ha egy nagyobb szerkezet alaptagja/profilmeghatározója tartományközi leképezést kezdeményez, ez hatással van az alptag bővítményeire is, tehát ezekre az esetekre szükséges egy metaforához kapcsolódó argumentumcímke felvétele.

- MKKomp = *metaforához kapcsolódó komponens*: ez a címke azoknak a szerkezettagnak a jelölését szolgálja, amelyek ugyan nem kezdeményeznek metaforizálódást, de mivel a szerkezet egy másik eleme metaforizálódik, ennek a hatása kiterjedhet a szerkezet többi elemére is.

(6) még visszacsapódhatnék a *kabátom* szárnya

A fenti példában a *szárnya* főnévi konstrukció tekinthető a metaforikus jelentés kezdeményezőjének, hiszen a szárny főnév elsődleges jelentése ('állatnak, képzeletbeli lénynek a repülésre való testrésze') helyett a 'szétnyitható ruhadarab nyitható része' kontextuális jelentés érvényesül. Így a nominális MKK címkét kap az annotálás során. A *kabátom* birtokos személyjellel ellátott főnév viszont birtokviszonyt tesz hozzáférhetővé egy referenciapont-szerkezet részeként, ezáltal a metaforához kapcsolódó komponensként hozzájárul a metaforikus jelentés kibontakozásához.

- MKKid = *metaforához kapcsolódó kifejezés idióma részeként*: ezt a címkét azon kifejezések jelölésére alakítottuk ki, amelyek idiomatikus szerkezet részeként kezdeményeznek metaforikusságot, tehát maga a szerkezet rendelkezik egy konvencionizált metaforikus jelentéssel.

(7) Sütő úr ugyanis, mivel faji tekintetben őneki teljesen rendben a *szénája*, nem visel sárga csillagot

A (7) példában a *rendben a szénája* idiomatikus kifejezés szerepel (vö. SzK.), amely szerkezet tagjai közül a *szénája* nominális egyértelműen a kezdeményezője a tartományközi leképezésnek, hiszen a fenti kontextusban nem a *széna* 'lekaszált és megszártított fű v. más takarmánynövény mint állati eledel' jelentésében szerepel, hanem körülményekre, dolgokra, érzelmekre, a saját maga előidézte fejleményekre vonatkozik (I. SzK.). Ezért a *szénája* nominálist az annotáció során az MKKid címkével láttuk el.

- MKAid = *metaforához kapcsolódó argumentum idioma részeként*: ez a címke jelöli azokat az argumentumokat, amelyek egyfelől egy metaforizálódó igealakhoz kapcsolódnak, másfelől az igealakkal idiomatikus kifejezést alkotnak, azaz a szerkezet egésze rendelkezik egy konvencionizált metaforikus jelentéssel.

(8) az ő *kedvét* kell nézmem

A (8) szövegrészlet arra mutat példát, hogy bizonyos esetekben az idiomatikus kifejezésben előforduló metaforizálódó összetevő jellege is pontosítható: a *kedvét* nominális a *kedvét nézi/keresi valakinek* szerkezet állandósult komponense, egyúttal az igealak (illetve a példában a főnévi igenév) argumentuma is. A származtatott MKAid címkével mindkét funkció együtt jelölhető.

- MZ = *metaforazászló*: azon morfémák, szerkezettagok jelölésére szolgál, amelyek a tartományközi leképezéseket explicitté teszik, mintegy reflektálva a metaforikus jelentésre, az eredeti módszert követve a metaforazászló-címkét hoztuk létre.

(9) szűk, akváriumszerűen kivilágított üvegkalicka

A fentiek alapján a (9) példában a *-szerű* szerkezettag kapott MZ címkét, mivel kifejtette teszi, hogy a szerkezet másik tagja hasonlóságon alapuló tartományközi leképezést valósít meg a jelentés tekintetében. A *-szerű* morféma tehát önmagában nem kezdeményez metaforikusságot, ellenben ráirányítja a figyelmet arra, hogy a konstrukció másik tagja metaforához kapcsolódó kifejezés.

- dMKK = *direkt metaforához kapcsolódó kifejezés*: ez a címke azon szerkezettagok jelölésére szolgál, amelyek egy konstrukción belül a nyelvi reflexió tárgyává váló metaforikusságot kezdeményezik.

A dMKK címke sajátossága, hogy valamennyi esetben csak a MZ címkével együttesen alkalmazható, ahogyan arra a (9) példa is felhívja a figyelmet. A példában az *akvárium* szerkezettag kezdeményezi a tartományközi leképezést, majd a *-szerű* morféma „zászlóként” teszi jelöltté, hogy az *akvárium* főnév direkt figuratív összehasonlítás alapjává válik az adott kontextusban.

- MKKimp = *metaforához kapcsolódó implicit kifejezés*: ezt a címkét annak jelölésére vezettük be, amikor olyan kifejezéssel találkoztunk, amely koreferens viszonyban van valamely metaforikus kifejezéssel, ilyen módon tehát részt vesz a metaforához kapcsolódó szerkezet kialakításában, nem kezdeményezője azonban a tartományközi leképezésnek.

(10) [...] hogy örködhessen az ottani javainkon, no meg hogy *ezalatt* a bevételről se kelljen egészen lemondanunk.

A példa kiemelt kifejezése névmási anaforaként utal vissza az anyagi javakon való örökösre mint antecedensre. Az örökdhessen kifejezés a szövegelőzményben MKK címkét kapott, mert kontextuális a jelentése ('Félt gonddal vigyáz vmire') metaforikus kiterjesztése az alapjelentésnek ('Örként szolgálatot teljesít vhol'). Következésképpen az *ezalatt* anafora implicit metaforikus kifejezésnek minősíthető, ezért az annotálás során MKKimp címkével jelöltük.

### 3.3. Metaforaazonosítás magyar nyelven: a módszer lépései

A részletes tárgyalás, továbbá az annotálás során alkalmazott címkék és segédeszközök bemutatása után ebben a pontban röviden összefoglaljuk a metaforaazonosítás magyar nyelvre adaptált módszerét. Az eljárás lépései a következők.

1. Tagoljuk a vizsgált szöveget morfológiai egységekre.
2. Ha egy morféma kontextuális jelentése indirekt jellegű, és ez potenciálisan magyarázható az alapjelentésből eredő tartományközi leképezésekkel, azonosítsuk metaforikusként.
  - a) Az alap- és a kontextuális jelentés megállapításánál támaszkodjunk a Magyar értelmező kéziszótárra.
  - b) Amennyiben a szótári jelentésadás nem pontos vagy nem alkalmazható az aktuális jelentéshez, saját jelentésadással pótoljuk.
3. Ha egy igei szerkezet alaptagja metaforizálódik, a hozzá kapcsolódó argumentumot, amely hozzájárul a hasonlóságon alapuló jelentéskiterjesztéshez, szintén azonosítsuk metaforikusként.
  - a) Az igei metafora argumentumainak megállapításához alkalmazzuk a Magyar igei szerkezetek szótárát.
  - b) Amennyiben a szótár nem tartalmazza a szövegben előforduló ige + argumentum szerkezetet, a Magyar nemzeti szövegtár 2 korpusza alapján állapítsuk meg az együttes előfordulás gyakoriságát.
4. Ha egy összetett szerkezet (esetraggal ellátott nominális, igekötős ige, birtokos szerkezet) egyik eleme metaforizálódik, és ennek hatása kiterjed a teljes szerkezet más elemére is, utóbbit szintén azonosítsuk metaforikusként.
5. Ha egy kifejezés koreferens viszonyban áll valamely metaforikus kifejezéssel, és ezáltal részt vesz a metaforához kapcsolódó jelentés aktívan tartásában, azonosítsuk metaforikusként.
6. Ha egy morféma vagy egy szerkezet komponense a tartományközi leképezések jelölőjeként funkcionál, azonosítsuk metaforazászlóként.
7. Ha egy metaforizációt kezdeményező kifejezés egyben idiomatikus szerkezet komponense is, a teljes szerkezetet azonosítsuk metaforikusként.
  - a) Az idiomatikus megállapításához alkalmazzuk a Szólások és közmondások szótárát.
  - b) Amennyiben az idiomatikus szerkezet ige + argumentum szerkezetként jelenik meg, az alaptagot és az argumentumo(ka)t a 3. lépésnek megfelelően jelöljük.

### 3.4. Problémás esetek

Az MIPVU-módszer magyar nyelvre történő adaptálása során nem csupán azokkal az elméleti és módszertani kérdésekkel szembesültünk, amelyek egyfelől az eredeti módszer döntéseiből következtek (mint amilyen a metafora és a metonímia együttes annotálása), másfelől a magyar nyelv sajátosságai tették tárgyalásukat szükségessé (mint amilyen az esetragok vagy az igekötők kezelése). Számos olyan token is megjelent már a próbakorpuszban is, amelyek metaforikussága a mégoly körültekintő annotálási eljárás során sem vált egyértelművé. Gerard Steen és munkatársai ezekre az adatokra vezették be a WIDLII (*When In Doubt, Leave It In*) címkét, mi egyszerűen problémás esetekként kezeltük ezeket, amelyek jelölésére azonban nem alkalmaztunk külön címét. A teljesség igénye nélkül szeretnénk a továbbiakban bemutatni belőlük néhányat, ez ugyanis az annotálás módszerének jövőbeli felülvizsgálatát és finomítását is segítheti.

(11) Hajdúszoboszlón a szilveszteri *zajcsapáshoz* mondai hagyományt fűznek

A fenti példában a *zajcsapáshoz* nominális bizonyult problematikusnak: a *csapás* utótag alapjelentése 'erőteljes ütés', az összetétel azonban az ütés keltette hanghatásra vonatkozik. Vagyis indirekt jelentésre látunk példát, ám ez sokkal inkább metonimikus (A CSELEKVÉS MÓDJA A CSELEKVÉS EREDMÉNYE HELYETT), nem pedig hasonlóságon alapuló metaforizálódás. Így az adaptált módszer értelmében az összetételi fő nem minősül metaforikusnak.

(12) az egész csak amolyan üzleti *cselfogás* tulajdonképpen

Ugyancsak problematikusnak bizonyult a *cselfogás* összetétel: az egyes tagok jelentése ('mások rászédésére kieszelt ravasz ötlet', illetve 'cselekvésben, eljárásban alkalmazott célszerű, ügyes, gyak. ravasz mód[ozat]') eleve elbizonytalanítja a metaforikusságot, hiszen a két jelentés átfedi egymást, a ravaszság mindkettőnek elsődleges része. Eszerint a (12) példa is inkább metonimikus. Az igazi problémát azonban az okozza, hogy az eljárás során alkalmazott szótár tartalmaz *cselfogás* lemmát is ('csel, fortély' alapjelentéssel), vagyis a szerkezetet lexikalizálódottként mutatja be. Mindez egyfelől arra utal, hogy a lexikalizálódás megnehezíti a metaforikusság (vagy egyéb figuratív jelentés) megállapítását (a lexikalizálódott szerkezet ugyanis jelentésében is egységként gyakorlódik be). Másfelől azt is mutatja a példa, hogy az összetételek körültekintő kezelése a lexikalizálódás vizsgálatára is ki kell terjedjen.

(13) A Toldiban a cseh vitézt, a Toldi estéjében az olaszt *győzi le*

Munkánk során szembesültünk az igekötők metaforizálódásának kérdésével is. Metaforikusságot ezekben az esetekben is az elsődlegestől eltérő jelentéseknél – tehát az igekötők kapcsán a fizikai irányulástól való elmozdulásoknál – jelöltünk. Az elmozdulás teoretikus végpontja a kizárólagos aspektusjelölés, amikor tehát

már nem érvényesül az irányulás. A hétköznapi nyelvhasználó számára ugyanakkor a pusztá aspektualitás sem feltétlenül tűnik metaforikusnak. A kiemelt példában (13) ráadásul felvetődik, mennyiben aspektualitásról és mennyiben irányjelentésről van szó, hiszen a *legyőz* kifejezésben érzékelhető egy olyan jelentésárnyalat, amely a fizikai küzdelemben a 'földre kerülés' mozzanatát sugallja. A *fel* és a *le* esetében mindenképpen számolnunk kell azzal, hogy metaforizálódott szerkezetekben is jóval erőteljesebben érzékelhető az irányjelentés, mint más igeekötőknél (Szili 2009: 183). Hangsúlyozni kell azonban, hogy az adaptált módszer nem támaszkodik a szavak keletkezésekor meglévő motiváltsága; a történeti mozzanatokat az elemzés a háttérben hagyja. A kérdéses esetben a metaforikusság mellett szól, hogy az irányjelentés együtt érvényesül az aspektualitás kifejezésével, illetve megjelenik a totalitás akcióminősége is: ezt támasztja alá a *legyőz* szótári jelentése („teljes győzelmet arat valakin”). Szili a szót *AZ ALÁVETETTSÉG / A VERESÉG LEFELÉ IRÁNYULTSÁG* kategória példajaként hozza, tehát szintén metaforikusnak tekinti (2009: 185). Az igeekötők annotálása során általános elvünk ké vált, hogy akkor tekintünk metaforikusnak egy igeekötőt, ha a fizikai irányulás jelentése elmozdul a nem fizikai tartományba, ugyanakkor nem válik kizárólagossá az aspektus kifejezése.

#### (14) tematikailag összefüggő műalkotás

Az összefüggő token esetében felmerült, hogy a *függ* ('fix pontról lóg') jelentése az összefügg lexikalizálódott szóalakban még aktíválódik-e. Az összefügg elsődleges jelentése a szótár szerint: 'egymással összefüggésben lévő, összefüggést alkotó', vagyis éppen a konvencionalizált metaforikus jelentés, amely már elszakadt a *függ* eredeti jelentésétől. Természetesen az igeekötő irányjelentése is módosít az ige jelentésén. A lexikalizálódás intenzitása azonban meghaladta azt a pontot, amelytől már nem felismerhető az eredeti jelentés, és bár nem egyértelműen zárható ki a metaforikusság, ezt a tokent is problémás esetnek tekintettük, így a metaforikusságot nem jelöltük.

#### (15) Egy mozdulatot is *megindított* énfelém a kezével

Nem csupán a kontextuális jelentés megállapítása, illetve az alapjelentéssel való viszonya okozhat elemzési nehézségeket – maga az alapjelentés sem minden esetben segíti a metaforikusság megállapítását. A (15) példában a *megindít* ige főalapelentése ('<gépet, szerkezetet, géperejű járművet> mozgásba, működésbe kezd hozni') kellően specifikus és konkrét a metaforizálódás kiindulópontjává váláshoz, az annotátor által kialakított kontextuális jelentés elvontabb ('fizikai entitás mozgását, ill. állapotának változását lehetővé teszi'). A problémát az okozza, hogy a két jelentésadás közötti távolság (azaz a jelentések elkülöníthetősége) nem határozható meg kellő pontossággal, részben mert a kontextuális jelentés az annotátor munkájának eredménye (a szótár nem tartalmaz a példának megfelelő jelentést), részben pedig, mert az analógia nem állapítható meg egyértelműen: generalizálódásként is értelmezhető a két jelentés kapcsolata. Az MIPVU ezekben az esetek-

ben az etimológiát ellenőrzi: amennyiben a két jelentésnek közös forrása van, két származtatott alapjelentésként kezeli ezeket. Ha azonban az egyik jelentés bizonyul etimológiailag elsődlegesnek, jelentéskiterjesztésről beszélhetünk. (TESz.)

(16) A fapincénk itt a közelben, egy *mellékutcában* van.

Előfordultak olyan esetek, amikor az inflexiók metaforizálódásának megállapításakor szembesültünk nehézségekkel, amelyek a konstrukció másik szerkezettagjának, a tőnek a jelentésével hozhatóak összefüggésbe. A (16) példában szereplő *-ben*, illetve *-ban* inessivus rag alapjelentése 'bent lévő helyen, belül', tehát a BENTLEVÉS állapotát, azaz valamely körülhatárolt fizikai tárgynak egy másik, szintén körülhatárolt fizikai tárgyban való elhelyezkedését jelenti (Tolcsvai Nagy 2017: 373). A példában azonban mind a *közel* határozószóról, mind a *mellékutca* főnévről nehezen állapítható meg, hogy az általuk hozzáférhetővé tett fizikai tér milyen mértékben tekinthető körülhatolhatóknak, amelyben – mintegy tartályban – lehetséges a BENTLEVÉS, ebből adódóan pedig kérdésessé válik, hogy a rag jelentéstartományok közötti leképezést valósít-e meg, vagy sem. Mivel azonban mindkét szerkezet alaptagjának az elsődleges jelentése ('kis távolságban van', illetve 'városban, falun kisebb közlekedési útvonal, amelyet legalább részben házak szegélyeznek') a fizikai tér egy specifikumát teszi hozzáférhetővé, így az MKI címke mellőzése mellett döntöttünk.

(17) Így mondjuk legalábbis.

További problémát okozott az annotálás során azoknak az igei szerkezeteknek elemzése, amelyekről az MISz. nem adott egyértelmű információt. Ilyen konstrukció volt a (17) példa is. Ezeknél az eseteknél a Magyar nemzeti szövegtár (MNSz2 v2.0.5.) adatbázisából kértünk le gyakorisági listát, tehát a (17) példában prezentálódó *Így mondjuk* igei konstrukció esetében is így jártunk el. Az MNSz2 keresőfelületén a *mond* szótőre kerestünk, és ezen elemre vonatkozóan kértünk le gyakorisági listát, amelyben hozzáférhetővé vált az, hogy a *mond* szótövet tartalmazó lexémától balra első helyen az *így* határozói névmás 4879 gyakorisági értékben fordult elő a korpuszban, míg az összes *mond* szótő előfordulása 2 131 269 volt, a *mondjuk* szóalak pedig 175 655 előfordulással jelenik meg. Vagyis a névmás közel 3%-ban jellemzően együtt fordul elő az igealakkal, ezért, bár az argumentumokat listázó szótári adatbázis nem tartalmazza (az ugyanis igei lemmákat vizsgál), mégis gyakori argumentumnak tekinthető a konkrét szóalak mellett. Tehát azoknál az eseteknél, amelyeknél az igei argumentumszerkezetek korpuszalapú feldolgozása nem igazított el a szerkezet elemzésében, a Magyar nemzeti szövegtár alkalmazhatóan bizonyult.

(18) Az iskolától az üzemünkig gyalogosan *tettem meg az utat*.

Végül a (18) idézet kevésbé egy problémás esetet, sokkal inkább egy továbbmutató kutatási irányt példáz. A példában a *tettem* ige MKK címkével láttuk el, mert nem

a 'cselekszik' alapjelentés érvényesül az adott kotextusban, mivel pedig az *utat* nominális a *tesz* ige argumentuma, azt az MKA címkével láttuk el. Összességében tehát egy olyan igei konstrukciót annotáltunk, amelyben a metaforikus jelentés kialakításában az igei alaptag rendelkezik profilmeghatározó szereppel, ugyanakkor a példa felhívja a figyelmet arra is, hogy a *tettem meg az utat* szerkezetet célszerű lehet kollokációnak tekinteni (az MNSz2. 1318 *tettem meg* előfordulásából 102 esetben jelenik meg a szerkezet előtt vagy mögött [5 szónyi távolságban]) az *utat* nominális, amely 10 fölötti MI-értékkel és 5,6-os logDice értékkel válik kollokációjelöltté), amelynek a szerkezettagjai között szorosabb, tartományközi leképzéssel összefüggésbe hozható jelentésvizony van. Az annotációs eljárás tehát *utat* nyithat a kollokálódó metaforikus szerkezetek szisztematikus vizsgálata felé is.

#### 4. Összegzés és kitekintés

A tanulmány több, egymással szorosan összefüggő célt tűzött ki maga elé. A metaforikus jelentés fogalmi szerveződése helyett annak nyelvi megvalósulására irányította a figyelmet, összhangban a kognitív nyelvészet egyre erősödő nemzetközi tendenciájával, amely a fogalmi metaforákat heurisztikus modelleknek tekinti, és a korpuszokra alapozódó vizsgálatokat szorgalmazza. Ugyanakkor a korpuszalapú metaforaelemzésnek is egyetlen fázisát, a metaforikus kifejezések korpuszbeli azonosítását és jelölését állította a középpontba, hogy a nyelvi metafora annotálásának megvalósításával további korpusznyelvészeti kutatásokat kezdeményezzen. Végül, a tanulmány legszűkebb fókusza a magyar nyelvű metaforikus szerkezetek annotálhatóságának a kérdése volt, amelyre egy, a nemzetközi kutatásokban már eredményesen alkalmazott eljárás (MIPVU) adaptálásával próbáltunk válaszokat kínálni.

A metaforaazonosítás problémájának körülhatárolása, majd az eredeti módszer részletes bemutatása után az adaptálás kérdéseit tárgyaltuk. A magyar nyelv sajátosságai tették szükségessé a szóalapú annotálás elhagyását, az inflexiók vizsgálatát, továbbá az összetételek és az igekötők megkülönböztetett kezelését. Az adaptálás további iránya egy pontos, a metaforizálódás nyelvi szerkezeti tényezőiről is számot adó módszer kidolgozása volt, ezért a metonimikus kifejezéseket nem jelöltük potenciális metaforaként, ugyanakkor bevezettük a metaforikus komponens és a metaforikus argumentum kategóriáját, valamint az idiomatikus szerkezetekre is külön jelölést alakítottunk ki.

A tanulmányban nem csupán módszertani javaslatokkal éltünk, hanem az adaptálás folyamatát segítő próbaannotálás menetéről is beszámoltunk. Egy több mint 400 szövegszóból álló, három szövegrészletet tartalmazó próbakorpusz kialakításával és elemzésével finomítottuk az átvett, illetve kialakított címkéket. Jelenleg ennek a korpusznak az online annotálása zajlik, emellett új szövegrészlettel bővítjük a próbakorpuszt, amely így már kvalitatív elemzési szempontok érvényesítésére is alkalmasnak mutatkozhat.

A kutatás távlati lehetőségei sokrétűek. Mindenekelőtt szükségesnek mutatkozik az azonosítási eljárás hatékonyságának a megállapítása (és fokozása). Ezt az online annotálás technikája lehetővé teszi, de annotátorközi egyezés szá-

mítását is tervezzük a közeljövőben. A kellően hatékony módszer lehetővé teszi majd minél több és minél változatosabb szövegek annotálását, amely már egy metaforakorpusz kialakításának az irányába mutat, alkorpuszokkal, és (távoli célként) a kereshetőség biztosításával. Ám nem csupán extenzív bővítésre ad módot a módszertani adaptálás: a magyar metaforikus kifejezések bizonyos szerkezeti jellemzői is feltárhatóvá válnak már a kibővített próbakorpuszban is, mint amilyen az igecentrikusság, a metafora szerkezeti hatóköre vagy a névszói metaforák mintázatai. Bízunk abban, hogy kutatásunkkal nem csupán új módszert adhatunk az érdeklődő kutató kezébe, de a metaforizálódással kapcsolatos nyitott kérdésekre is ráirányíthatjuk a figyelmet.

## FORRÁSOK

- ÉKsz.<sup>2</sup> = Puztai Ferenc (főszerk.) 2003. *Magyar értelmező kéziszótár*. Akadémiai Kiadó, Budapest.
- JNÜSz. = Tátrai Zsuzsanna – Karácsony Erika 1997. *Jeles napok, ünnepi szokások*. Mezőgazda Kiadó, Budapest.
- MGr. = Keszler Borbála (szerk.) 2000. *Magyar grammatika*. Nemzeti Tankönyvkiadó, Budapest.
- MIK. = Devecsery László 2002. *Mesélő irodalmi kislexikon*. Tóth Könyvkereskedés és Kiadó Kft., Budapest.
- MISz. = Sass Bálint – Váradi Tamás – Pajzs Júlia – Kiss Margit (szerk.) 2011. *Magyar igei szerkezetek. A leggyakoribb vonzatok és szókapcsolatok szótára*. Tinta Könyvkiadó, Budapest.
- MNSz2. = *Magyar Nemzeti Szövegtár 2*. Magyar Tudományos Akadémia Nyelvtudományi Intézet, Budapest. Nszt. = Ittész Nóra (főszerk.) 2006–. *A magyar nyelv nagyszótára I–*. MTA Nyelvtudományi Intézete, Budapest.
- Sorstalanság = Kertész Imre 2016. *Sorstalanság*. Magvető Kiadó, Budapest.
- SzK. = Szemerkenyi Ágnes 2009. *Szólások és közmondások*. Osiris Kiadó, Budapest.

## SZAKIRODALOM

- Babarczy Anna – Simon Eszter 2012. A fogalmi metaforák és a szövegstatistika szerepe a metaforák felismerésében. In: Prószték Gábor – Váradi Tamás (szerk.): *Nyelvtechnológiai kutatások. Általános Nyelvészeti Tanulmányok XXIV*. Akadémiai Kiadó, Budapest, 243–64.
- Biber, Douglas 2010. Corpus-based and corpus-driven analysis of language variation and use. In: Heine, Bernd – Narrog, Heiko (eds.): *The Oxford Handbook of Linguistic Analysis*. Oxford University Press, Oxford, 159–92.
- Deignan, Alice 2005. *Metaphor and Corpus Linguistics*. John Benjamins, Amsterdam, Philadelphia.
- Deignan, Alice 2008. Corpus linguistic data and conceptual metaphor theory. In: Zanutto, Maria Sophia – Cameron, Lynne – Cavalcanti, Marilda C. (eds.): *Confronting metaphor in use. An applied linguistic approach*. John Benjamins, Amsterdam, Philadelphia, 149–62.
- Forgács Bálint 2013. Agy és megismerés a figuratív nyelv megértésében. In: Pléh Csaba (szerk.): *A kognitív szempont a nyelv pszichológiájában. Általános Nyelvészeti Tanulmányok XXV*. Akadémiai Kiadó, Budapest, 383–410.
- Lakoff, George 2006/1993. Conceptual metaphor. The contemporary theory of metaphor. In: Geeraerts, Dirk (ed.): *Cognitive Linguistics: Basic Readings*. Mouton de Gruyter, Berlin, New York, 185–238.
- Sass Bálint – Váradi Tamás – Pajzs Júlia – Kiss Margit 2011. *Magyar igei szerkezetek*. Tinta Könyvkiadó, Budapest.
- Simon Gábor 2018. Az igei jelentés metaforizációjának mintázatai. *Nyelvtan- és korpuszvezérelt esettanulmányok. Jelentés és Nyelvhasználat 5*: 1–36.



- Sperber, Dan – Wilson, Deirdre 2013. A metafora leleplezése. In: Pléh Csaba (szerk.): A kognitív szempont a nyelv pszichológiájában. *Általános Nyelvészeti Tanulmányok XXV.* Akadémiai Kiadó, Budapest, 313–38.
- Steen, Gerard 2007. *Finding Metaphor in Grammar and Usage. A methodological analysis of theory and research.* John Benjamins, Amsterdam, Philadelphia.
- Steen, Gerard J. – Dorst, Aletta G. – Herrmann, Berenike J. – Kaal, Anna A. – Krennmayr, Tina – Pasma, Trijntje 2011. *A Method for Linguistic Metaphor Identification. From MIP to MIPVU.* John Benjamins, Amsterdam, Philadelphia.
- Stefanowitsch, Anatol 2006. Corpus-based approaches to metaphor and metonymy. In: Stefanowitsch, Anatol – Gries, Stefan Th. (eds.): *Corpus-Based Approaches to Metaphor and Metonymy.* Mouton de Gruyter. Berlin, New York, 1–16.
- Szili Katalin 2009. A *fel, le* és egyéb igekötős igék formai-szemantikai viszonyának kérdéséhez. *Magyar Nyelv* 105: 175–88.
- Tolcsvai Nagy Gábor 2017. Jelentéstan. In: Tolcsvai Nagy Gábor (szerk.) *Nyelvtan.* Osiris Kiadó, Budapest, 207–499.
- Váradi Tamás – Simon Eszter – Sass Bálint – Geröcs Mátyás – Mittelholtz Iván – Novák Attila – Indig Balázs – Prószéky Gábor – Farkas Richárd – Vincze Veronika 2017. Az e-magyar digitális nyelvfeldolgozó rendszer. In: Vincze Veronika (szerk.): *XIII. Magyar Számítógépes Nyelvészeti Konferencia.* Szegedi Tudományegyetem, Szeged, 49–60.

Simon Gábor      Szlávich Eszter  
egyetemi adjunktus      PhD-hallgató

Bajzát Tímea – Ballagó Júlia – Havasi Zsuzsanna – Roskó Mira  
egyetemi hallgató  
Eötvös Loránd Tudományegyetem

#### SUMMARY

*Simon Gábor – Bajzát Tímea – Ballagó Júlia –  
Havasi Zsuzsanna – Roskó Mira – Szlávich Eszter*

#### **Metaphor identification in Hungarian texts: a methodological adaptation**

The study aims at elaborating a method for metaphor identification in Hungarian. The vantage point of the endeavour is the so called MIPVU (Metaphor Identification Process Vrije Universiteit, University of Amsterdam). It was developed for metaphor annotation in English and Dutch, being based on corpus-based dictionaries. Consequently, the main challenge – beside the missing corpus-based concise dictionary of Hungarian – was to adopt the original methodology for an agglutinative language, as well as to make it morpheme-based instead of word-based. Moreover, we have made additional changes in the process in order to specify not only the structure of a metaphorical expression, but also the semantic relations in it. The adapted method has been tested and evaluated in the course of annotating a small-scale research corpus. The paper details some of the problematic cases as well.

**Keywords:** metaphor, identification, semantic relations, annotation