

## Szemantikai igeosztályok tesztelése az MNSz-ben

Gábor Kata<sup>1</sup>, Héja Enikő<sup>1</sup>

<sup>1</sup> MTA Nyelvtudományi Intézet, Korpusznyelvészeti osztály, Postafiók 701/518,  
H-1399 Budapest, Magyarország

{gkata, eheja}@nytud.hu

**Kivonat:** A nyolcvanas évektől megnőtt a lexikon szerepe a szintaxiselméletekben. Egyre több megközelítés feltételezte, hogy az ige lexikális jelentése nagymértékben meghatározza azt a szintaktikai környezetet, amelyben az ige előfordulhat. Ezt az előfeltevést elfogadva, hosszú távú célunk egy olyan szabályalapú szintaktikai elemző fejlesztése, amely az igével egy tagmondatban szereplő esetragos főnévi csoportok annotálása során az ige releváns szemantikai tulajdonságaira támaszkodik. Jelen cikkben a Magyar Nemzeti Szövegtár adatai alapján azt vizsgáltuk, hogy milyen mértékben alkalmazhatók a magyarra Levin [6] szemantikai metapredikátumokkal definiált angol igeosztályai.

### 1. Bevezetés

Cikkünkben arra a kérdésre keressük a választ, hogy alkalmazhatók-e Levin [6] igeosztályai egy szabályalapú magyar szintaktikai elemzőben. Munkánk arra a feltevésre épít, hogy az angol vonzatkeret-alternációkat univerzálisan érvényes szemantikai metapredikátumok irányítják, melyek – legalábbis bizonyos szempontok szerint – a magyar nyelvben is egymáshoz hasonló viselkedésű ige csoportokat jelölnek ki. Levin az angol igék argumentumrealizációs lehetőségeit vizsgálva az igék szintaktikai tulajdonságaiból indul ki, és ebből következtet közös lexikai szemantikai tulajdonságokra. Mi ezt az irányt megfordítva arra voltunk kíváncsiak, hogy az általa felépített lexikai szemantikai reprezentációt magyar igékre átültetve megkapjuk-e azokat az igeosztályokat, melyek azonos bővítménykeretekkel rendelkeznek. A bővítménykeretet [3] alapján Levinnél tágabban értelmezzük, mivel véleményünk szerint az igék lexikális jelentés-komponensei nemcsak kötelező vonzataik, hanem a mellettük kitéhető szabad határozók körét és morfoszintaktikai tulajdonságait befolyásolják. Ezért nemcsak a feltételezett vonzatokat, hanem az igéket módosító, NP kategóriájú szabad határozók egy részét is az ige lexikai szemantikai reprezentációja által meghatározottnak tekintjük.

Hipotézisünket a Magyar Nemzeti Szövegtárból vett adatokon teszteltük. Kiválasztottunk két angol ige csoportot, és a korpuszból olyan mondatokat gyűjtöttünk, melyekben az igék magyar megfelelői állítmányként szerepelnek. Ezután az INTEX/NooJ korpuszfeldolgozó eszköz segítségével a mondatokat tagmondatokra bontottuk, és futtattuk a szövegen a legfelsőbb szintű NP-eket felismerő nyelvtanunkat. Az eredményül kapott tagmondat-vázakon azt vizsgáltuk, hogy melyek azok az

esetragok, melyek az ige csoportok mellett szereplő felső szintű NP-ken leggyakrabban előfordulnak.

A továbbiakban bemutatjuk az argumentumrealizációs elméletek közös előfeltevéseit és a számunkra releváns *predikátumdekompozíciós elmélet* módszereit [2.fej.], majd ismertetjük, hogy milyen szerepet tulajdonítunk a predikátumdekompozíciós jelentésleírásnak a szintaktikai elemzésben [3.fej.]. Ezután bemutatjuk a hipotézis ellenőrzéséhez használt módszerünket [4.fej.] és a vizsgálat eredményeit két, általunk választott ige csoportra [5.fej.]. Végül szót ejtünk a munka lehetséges további irányairól [6.fej.].

## 2 Az igei jelentés szintaktikai vonatkozásai

A nyolcvanas évektől megnőtt a lexikon szerepe a szintaxiselméletekben. Egyre több megközelítés feltételezte (pl.: LFG [4], GB [1], Role and Reference Grammar [2]), hogy az ige jelentése szerepet játszik az argumentumainak felszíni szintaktikai realizációjában, azaz az igei jelentés nagymértékben meghatározza azt a szintaktikai környezetet, amelyben az ige előfordulhat. Ez a feltevés abból a megfigyelésből indult ki, hogy a hasonló jelentésű igék vonzatkerete hasonló mintát mutat.

A szóban forgó elméletek közös előfeltevése, hogy az ige szemantikai reprezentációjában szereplő argumentumok a szintaxisban vonzatokként realizálódnak. Céljuk olyan szabályszerűségek megállapítása, amelyek alapján megjósolhatóvá válik egy adott argumentum szintaktikai funkciója. Így az argumentumok szintaxisban betöltött szerepét megjósoló elméletek sikeressége elsősorban két tényezőn múlik: egyfelől az ige megfelelő szemantikai reprezentációján, másfelől pedig a leképezési eljárás (*linking/mapping theory*) megfelelő kialakításán.

Az ige szemantikai reprezentációja szempontjából az argumentumrealizációval foglalkozó elméletek (legalább) két nagy csoportra oszthatók. Az első csoportba tartozó elméletek feltételezik, hogy az ige szintaktikai funkcióját az ige által kiosztott tematikus vagy szemantikai szerepek határozzák meg. A másik csoportba tartozó elméletek ezzel szemben azt feltételezik, hogy az ige jelentése tovább bontható, és az igei jelentést karakterizáló metapredikátumok felelősek azért, hogy az argumentumok hogyan realizálódnak. Ezeket az elméleteket az alábbiakban predikátumdekompozíciós elméleteknek fogjuk nevezni. Levin [5]<sup>60</sup> a predikátumdekompozíciós elméletek mellett érvel a tematikus szerepeket használó elméletekkel szemben.

A predikátumdekompozíciós eljárás célja, hogy az igék jelentésében megtalálja azokat a közös jelentéskomponenseket vagy metapredikátumokat, amelyek az adott igékre jellemző szintaktikai viselkedésért felelősek. Mára általánosan elfogadottá vált az a nézőpont, hogy az igék reprezentációjában el kell választani egymástól a több igére jellemző általánosabb jelentéskomponenseket és az ige jelentésének idioszinkratikus részét. Így tehát az ige szemantikai reprezentációja az általánosabb, metapredikátumokkal karakterizált eseménytípus és a beágyazott idioszinkratikus rész kompozíciójaként áll elő. Ekkor azonban már nem tartható az a feltevés, hogy az ige jelentése *kizárólag* az ige argumentumainak – a vonzatoknak – a szintaktikai megvalósulásáért felelős. Mint tudjuk, a predikátum jelentése, pontosabban a predi-

<sup>60</sup> 2.2, 5. fej., 6. fej.

kátum jelentéséből elvonható eseményséma meghatározhatja, hogy milyen típusú időhatározók jelenhetnek meg a mondatban.

Azt gondoljuk tehát, hogy az ige jelentése nemcsak a vonzatainak felszíni megjelenését determinálja, hanem – az eseménysémán keresztül – a szabad határozók esetleges megjelenését is az ige engedélyezi. Egy ige akkor engedélyezi egy szabad határozó megjelenését a mondatban, ha az igei jelentésben van egy olyan metapredikátum, amely *kompatibilis* a szabad határozó jelentésével. Az elképzelés alapján nemcsak az időhatározók disztribúcióját jósolhatjuk meg. További példát szolgáltatnak az állapotváltozást jelentő mozzanatos igék, amelyek mellett a *-ra* esetragos főnév megadhatja az ige által leírt cselekvés közvetlen előzményét („*puskálövésre elindult*”). Ezzel szemben, ha a mozzanatoság metapredikátuma nincs jelen az ige szemantikai reprezentációjában, akkor a szóban forgó szemantikai szerep sem jelenhet meg, hiszen nincsen olyan metapredikátum, amely engedélyezné. (Pl.: \**puskálövésre üldögélt*”).

A predikátumdekompozíciós elméletek hátránya, hogy nehéz megtalálni az ige szintaktikai viselkedése szempontjából releváns metapredikátumokat. Levin[1993] szerint a metapredikátumokat az igecsoportok szintaktikai viselkedéséből kiindulva lehet feltérképezni. Előfeltévése szerint azok az igék rendelkeznek azonos metapredikátumokkal, amelyek azonos vonzatkeret-alternációkban vesznek részt.

Ezt az elképzelést [5] a kauzatív alternációval szemlélteti.

<i>The wind opened the door.</i> A szél kinyitotta az ajtót.	<i>The door opened.</i> Az ajtó kinyílt.
---	---

<i>Mary broke the window.</i> Mari betört az ablakot.	<i>The window broke.</i> Az ablak betört.
--	--

<i>Bill cooled the soup.</i> Bill lehűtötte a levest.	<i>The soup cooled.</i> A leves lehűlt.
--	--

Levin szerint a kauzatív alternáció arra a szemantikai tényre vezethető vissza, hogy vannak olyan igék, amelyek jelentésreprezentációjában az ige által leírt esemény külső okának van argumentumhelye. Külső ok alatt Levin az állapotváltozást elszenvedő dologtól független entitást ért, amely közvetlen hatást gyakorol a tárgy által jelölt entitásra (a fenti mondatokban ‘*a szél*’, ‘*Bill*’ és ‘*Mary*’).

Ezzel párhuzamosan Levin azt is feltételezi, hogy ha van egy olyan – alternációk által kijelölt – igeosztályunk, amely bizonyos alternációkban egyezik, ám más alternációkban eltér egy másik osztálytól, akkor van olyan jelentéskomponens, amely mindkét igeosztály tagjaiban közös. Ezt az állítást a hangkibocsátó igék ([6] 43.2, 234.o.) példájával szemlélteti. A csoportba tartozó igék egy része részt vesz a kauzatív alternációban, míg egy másik része nem:

<i>The train rumbled.</i> A teherautó zakatol.	* <i>Peter rumbled the train.</i> *Péter zakatoltatta a vonatot.
---	---

<i>The tea kettle whistled.</i> A teáskanna fütyült.	* <i>The boiling water whistled the tea kettle.</i> *A víz fütyültette a teáskannát.
---	---

*The teacups clattered.*  
A csészék csörögtek.

*I clattered the cups as I loaded the sink.*  
Csörgettem a csészéket, ahogy ... .

*The windows rattled.*  
Az ablak zörgött.

*The storm rattled the windows.*  
A szél zörgette az ablakot.

Ha a kauzatív alternációt megalapozó szemantikai bázis helyes volt, akkor annak alapján magyarázhatónak kell lenniük a fenti adatoknak is. Ha az alternációt az ige által jelölt esemény külső ok szerepű argumentuma teszi lehetővé, akkor az elmélet azt jósolja, hogy azon igék esetében nem lesz jól formált az alternáció tranzitív tagja, amelyeknél a hangkibocsátás oka az eseményben résztvevő entitás belső, inherens tulajdonságainak a következménye. A fenti példák pontosan ezt igazolják.

### 3 A jelentésrepresentáció szerepe a szintaktikai elemzésben

Munkánk kiindulópontjaként Levin [6] szolgál, melyben a fenti módszer alapján meghatározott angol igeosztályok szerepelnek. A vizsgálat eredményét egy szabályalapú magyar szintaktikai elemző [9] továbbfejlesztésében szeretnénk hasznosítani, elsősorban a legfelsőbb szintű mondat-konstituensek és az igei állítmány közti dependencia-viszonyok automatikus felismeréséhez. Mint azt [3]-ban ismertettük, a szintaktikai elemzés nélkülözhetetlen részének tartjuk a szemantikai igeosztályokra való hivatkozást, mert véleményünk szerint a magyarban a névszói bővítmény és az ige közti *szintaktikai* függőségi reláció jellege (azaz a névszó vonzat/adjunktum státusza) sem dönthető el az ige bizonyos szemantikai tulajdonságainak ismerete nélkül. Míg az adjunktumok közös tulajdonsága, hogy egy-egy igecsoport mellett ugyanolyan formai jegyekkel (esetraggal) jelölve ugyanazt a szemantikai szerepet töltik be, a vonzatok szemantikai szerepe mindig egyedi, és nem határozható meg az ige jelentésére való hivatkozás nélkül. A cikkben megválaszolandó fő kérdés, hogy a Levin által meghatározott – végső soron szemantikailag definiált – angol igeosztályok mennyiben mutatnak a magyarban is hasonló szintaktikai viselkedést, azaz mennyire hasonlít az a szintaktikai környezet, amelyben előfordulnak. A *szintaktikai környezet* fogalmát tágabban értjük, mint Levin: az igével előforduló valamennyi főnévi csoportot megvizsgáljuk. Mint a 2. fejezetben már említettük, véleményünk szerint az ige szemantikai tulajdonságai, vagyis az igei jelentés releváns komponensei elsősorban éppen az adjunktum szerepű NP-k disztribúciójáért felelősek. Amennyiben bizonyosodnak, hogy a Levin által használt igeosztályok magyar megfelelői is osztoznak szintaktikai tulajdonságokban, úgy az igeosztályokra érvényes általánosításokat fogalmazhatunk meg szabályok formájában, melyeket beépíthetünk a szintaktikai elemzőbe.

## 4 A vizsgálat módszere

Kiinduló hipotézisünk szerint tehát a közös jelentéskomponensekkel (metapredikátumokkal) jellemzett igék egymáshoz hasonló szintaktikai környezetekben fordulnak elő. Két további feltételezéssel élünk:

1) Levin igeosztályait *univerzálisan* érvényes szemantikai metapredikátumok határozzák meg;

2) Ezek az igeosztályok átültethetők magyar nyelvre, vagyis az adott igék jelentésének megfelelő magyar igék szintén egységes csoportot fognak kijelölni.

Természetesen várható, hogy egyes, az angolban egységes csoportok a magyarban különbözőképp viselkedő alcsoportokra oszlanak, illetve a szemantikailag homogén osztályokon belül is számíthatunk lexikai kivételekre. További kérdést jelent, hogyan értelmezzünk a korpusz adatait annak megítélésében, hogy bizonyos szerkezetek *lehetségesek* a magyarban, hiszen abból, hogy egy szerkezet nem fordul elő az MNSz-ben, nyilvánvalóan nem következtethetünk arra, hogy nem is létezik a nyelvben. Mindazonáltal a jelen cikkben leírt munka kereteiben azt vizsgáltuk, hogy a metapredikátumokkal definiált igeosztályok elemei jellemzően egymáshoz hasonló (illetve a többi csoporttól eltérő) szintaktikai környezetben fordulnak-e elő a korpuszban. Így csak azokat a szintaktikai mintákat vettük figyelembe, melyek egy-egy predikátum előfordulásainak legalább tíz százalékában megjelennek.

A feltételezéseink alapján azt várjuk, hogy ha az igeosztályok a magyarban is releváns csoportokat alkotnak, akkor elemeik osztoznak egyfelől legalább egy jelentéskomponensben, másrészt – ebből következően – engedélyezik egy vagy több, az osztályra specifikus szemantikai szerepű és esetragú NP megjelenését. Ha a csoportok igéi valóban egységes viselkedést mutatnak, akkor a szintaktikai elemzőt olyan szabályokkal bővíthetjük, melyek kimenete nemcsak felismeri, hanem szemantikai szerepet jelölő címkékkel is ellátja az igék egyes bővítményeit.

A hipotézist a Magyar Nemzeti Szövegtár [7] adatai alapján teszteljük. Az első szakaszban az ige bővítményeként szereplő esetragos főnévi csoportok disztribúcióját vetettük össze.

Az MSZNY 2005 konferencián elhangzott előadásunkban [3] az instrumentális (-val) esetrag lehetséges szemantikai szerepeit vizsgáltuk. Ezek között találtuk – többek között – a *mentális állapotváltozás oka* ('*megdöbbsent vkit vmive'l*') és az általunk *nem default eszközhöz* ('*beszennyez vmit vmivel'*') nevezett szemantikai szerepeket. Az alábbiakban azt a két igeosztályt fogjuk megvizsgálni, amelyek ezeket a szemantikai szerepeket engedélyezik a mondatban. A korpuszból lekérdeztük azokat a mondatokat, melyek az igék valamelyikét tartalmazzák, és olyan mintát állítottunk össze, mely igénként legfeljebb 3,000 mondatot tartalmaz. Az eredményül kapott szöveget az Intex korpuszfeldolgozó eszköz [7] magyar moduljával [9] részleges szintaktikai elemzésnek vetettük alá. A mondatokat tagmondatokra bontottuk, a legfelsőbb szintű NP-eket tagokkal helyettesítettük, melyek csak az NP esetragját/névutóját kódolják. Az így kapott tagmondat-vázakból azokat vettük figyelembe, melyek állítmányként (finit alakban) tartalmazzák az adott igeosztályt valamelyik elemét. Ezután minden ige-re megszámoztuk a vele azonos tagmondatban szereplő esetragos NP-k előfordulásait.

## 5 Eredmények

### 5.1 Mentális állapotváltozást jelentő igék

A mentális állapotváltozást jelentő igéket [6] 2.13.4 igeosztályával azonosítottuk, amely osztályt a tranzitív birtokos alany alternáció definiálja:

*Mark terrified me with his singlemindedness.*

Mark megijesztett engem a korlátoltságával.

*Mark's singlemindedness terrified me.*

Mark korlátoltsága megijesztett.

A 221 elemű angol igeosztály elemeit összesen 77 magyar igének feleltettük meg. Az osztályba tartozó angol igék tranzitívak. Magyarra fordításukkor azonban úgy döntöttünk, hogy átalakítjuk őket intranszítívvá, mert azt találtuk, hogy a magyar intranszítív alakok morfológiailag egyszerűbbek, a tranzitív alak belőlük képezhető (pl. *megijed-megijeszt*). Az igék fordításánál kritérium volt, hogy az igékhez tartozó eseményséma minél hasonlóbb legyen az eredetihez. Ha több lehetőség közül kellett választani, a nem kauzatív, befejezett aspektusú változatot részesítettük előnyben.

A korpuszból egy 13,023 mondatból álló mintát vettünk, melyek mindegyike tartalmazza az osztály valamelyik igéjét. Az igék többségénél azt tapasztaltuk, hogy leggyakrabban bővítmény nélkül, illetve egy (nominatívuszi) bővítménnyel fordulnak elő. Ez feltehetőleg annak tudható be, hogy a mentális állapotváltozást jelentő igék leginkább személyes közlésekben, vagyis az MNSz internetes fórum alkorpuszában szerepelnek, melyre a sok rövid tagmondatból álló, sok igét tartalmazó, hosszú mondatok jellemzőek. Az esetragok gyakoriságának vizsgálata során azokat az igei lemmákat vettük figyelembe, amelyek ötnél többször fordultak elő a korpuszban. Az esetragok között sorrendet állítottunk fel annak alapján, hogy hány különböző igei lemmával szerepelnek együtt a lemma előfordulásainak legalább 10 százalékában. (A nominatívuszt ezen a ponton nem vizsgáltuk.)

Ennek eredményeképpen az alábbi sorrendet kaptuk:

1. Táblázat: Gyakori esetragok a mentális igék mellett

Eset	Igék száma
INE	12
SUP	12
ABL	10
SUB	7
INS	5
DAT	2

Amint a táblázat mutatja, a leggyakoribb esetragok is csak az igék 15 százaléka mellett fordulnak elő legalább az esetek 10%-ában, tehát nem állíthatjuk, hogy egy-egy viselkedésű csoportot találtunk. Ezért megvizsgáltuk, hogy az esetraggyakoriság alapján milyen lehetséges alcsoportokat állapíthatunk meg, és hogy ezek szemantikailag koherens osztályokat alkotnak-e. Ha az igei lemmákat a mellettük

előforduló leggyakoribb esetrag szerint csoportosítjuk, akkor több esetragra (SUP, ABL, DAT) is olyan alosztályokat kapunk, ahol a szóban forgó esetrag szemantikai szerepe ugyanaz: minden osztály tagjai az ige jelentésében foglalt állapotváltozás okát kódolják az adott raggal<sup>61</sup>.

Inesszívusz (-ban):

*elfásul, kifárad, kimerül, elfárad*

Datívusz (-nak):

*megörül, megörvend*

Ablatívusz (-tól) és szuperesszívusz (-on) váltakozik:

*megrémül, felélénkül, elborzad, megrészegül, megijed, fellelkesül, elkábul, elhül, meghökken, meglepődik, elképed, megsértődik, megütközik, stb.*  
(Össz.: 25 ige)

Az adatok tehát arra utalnak, hogy a mentális állapotváltozást jelentő igék a magyarban szűk alcsoportokra oszthatók az alapján, hogy milyen esetraggal fejezik ki az *ok* szemantikai szerepű bővítményüket. Ezek közül az INE és a DAT esetraggal járó osztályok egy-egy szinonimacsoportot alkotnak, mag az ABL/SUP váltakozással jellemzett osztály jelentés szempontjából is heterogén. További vizsgálatra szorul, hogy milyen igei jelentéskomponensek irányítják az esetragok közti választást, és vajon ezek a jelentéskomponensek a magyarban tényleg csak az adott alcsoportot határozzák-e meg.

## 5.2 “Spray/load” igék

Az alcímben szereplő igeosztály talán a vonzatkeret alternációkkal foglalkozó szakirodalom legtöbbet tárgyalt jelensége. Levin felosztásában ez a 2.3.1 osztály. Az alternációt az alábbi példák szemléltetik:

- |   |  |
|---|--|
| a) <i>John loaded the truck with hay .</i><br>János megrakta a szekeret szénával. | <i>John loaded hay on the truck.</i><br>János szénát rakott a szekérrre. |
| b) <i>John sprayed the wall with paint.</i><br>János befújta a falat festékkel.   | <i>John sprayed paint on the wall.</i><br>János festéket fújt a falra.   |

Amint a példamondatok magyar megfelelői is mutatják, az eltérő bővítményszerkezet a magyar igék esetében gyakran igezőtő megjelenésével jár együtt. Ezért feltételeztük, hogy az angol igék magyar megfelelőinek vizsgálatába érdemes bevonni az igék igezőtős változatait is. A 49 elemű angol igeosztálynak 51 igezőtő nélküli magyar igét feleltettünk meg, és ezek igezőtős változatainak előfordulásait is lekérdeztük. Az igezőtős igék bevonása mellett szól továbbá, hogy az igecsoport elemeinek jelentése tartalmazza a *helyváltoztatás* jelentéskomponenst, ami alapján feltételezhetjük, hogy az irányt jelölő igezőtők – legalábbis sok esetben – produktívan kapcsolódnak az igehez, átlátszó szerkezetet eredményezve.

<sup>61</sup> Az INE todalék az esetek többségében idő vagy hely szerepű szabad határozót jelöl, de kijelöl egy jól körülhatárolt szemantikai alosztályt is, amely mellett *okot* fejezhet ki.

Az igekötős változatokkal együtt összesen 295 ige előfordulásait kérdeztük le, így 59,941 mondatot vizsgáltunk. Várakozásunk az volt, hogy az angol load-spray alternáció magyar megfelelője nagy számban fog előfordulni:

V	N.ACCi	INSj	<i>Vki tölt vmit vmivel.</i>
V	N.ACCj	LOCi	<i>Vki tölt vmit vhová.</i>

Ahol a LOC irányt jelölő esetrag (ILL, SUB vagy ALL (-hoz)).

Az igék osztályként jelentős eltérést mutatnak a mentális állapotváltozást jelentő igékhez képest. Ezt úgy vizsgáltuk, hogy ebben a csoportban is kiválasztottuk azokat az esetragokat, amelyek valamely ige mellett legalább az előfordulásaink a 10 százalékaiban szerepelnek, és rangsort készítettünk olyan szempontból, hogy hány ige mellett érik el ezt a gyakoriságot. Nem foglalkoztunk az egyetlen ige mellett szereplő esetragokkal, valamint az alany- és tárgyesettel sem. Várakozásunknak megfelelően azt találtuk, hogy az INS a leggyakoribb esetrag, de majdnem ugyanilyen gyakori az ILL és a SUB együttvéve. Érdekes, hogy a harmadik LOC szerepű rag, az ALL sehol nem éri el a 10% gyakoriságot.

## 2. Táblázat: Gyakori esetragok a spray/load igék mellett

Eset	Igék száma
INS	98
SUB+ILL	96
INE	57
SUB	53
ILL	43
SUP	41

A táblázat első két sora arra utal, hogy az angol spray-load alternáció magyar megfelelője jellemző erre az igecsoportra. Érdekes azonban azt is megvizsgálni, hogy a két szintaktikai minta ugyanazon igék mellett váltakozik-e, vagy a különböző bővítmenykeretek az adott ige igekötős változataihoz kötődnek. Azt mondhatjuk, hogy 28 alternáló ígét találtunk (amely mindkét mintában szerepel), 70 ige csak az ACC+INS, míg 51 csak az ACC+LOC mintában fordul elő. Ahol az alapige és igekötős változatai is megjelennek a mintázatban, ott a legtöbb esetben az alapige az egyetlen, amely mindkét mintával kompatibilis és 10% fölötti gyakorisággal elő is fordul bennük (pl. *locsol, dobál, töm, ken, permetez, borít*)<sup>62</sup>. Igekötős változataik a két minta közül legfeljebb egyben fordulnak elő 10%-nál gyakrabban. A *meg* és a *tele* – mint nem irányt jelölő igekötők – mindig az ACC+INS szerkezetet hívják elő. Az irányt jelölő igekötők esetében nem általánosítható, hogy melyik szerkezetet részesítik előnyben, például a *ki* a *borít* igével az ACC+LOC, a *töm* igével az ACC+INS mintát, a *be* a *ken* mellett az ACC+INS, a *szúr* mellett az ACC+LOC mintát hívja elő. Kérdés, hogy

<sup>62</sup> Az INS és a LOC típusú bővítmeny váltakozása jellemző a *szúr* igére is, de abban különbözik az említettektől, hogy a LOC típusú bővítmeny az INS-es szerkezetben nem tárgyként, hanem ILL (-ba) esetben jelenik meg.

az alapigék szemantikai tulajdonságai alapján megjósolhatók-e az igekötő-használat szabályai, vagy az igekötők nem átlátszó összetétellel kapcsolódnak ezekhez az igékhez.

## 6 Konklúzió és további teendők

A két csoport igéi mellett előforduló főnevek esetragjainak gyakorisága a csoportok között jelentős eltérést mutat, ami igazolni látszik azt a feltevést, hogy az igék lexikai szemantikai tulajdonságai befolyásolják a felszíni szintaktikai környezetüket. Mindazonáltal a leggyakoribb esetragokról sem mondható el, hogy a csoport igéinek többsége mellett előfordulna. Különösen igaz ez a mentális igék csoportjára, melyek mellett tagmondatonként csak átlagosan egy NP szerepel a korpuszban (az alanyesetűket is beleértve). Az esetragok előfordulásainak alacsony száma arra is rávilágít, hogy a magyarban kevés a kötelező vonzat<sup>63</sup>. A mentális állapotváltozást jelentő igékről bebizonyosodott, hogy nem érdemes őket szintaktikailag egységes csoportként kezelni, ám találhatunk szűkebb, szemantikailag és szintaktikailag is koherens alcsoportokat. Kérdés, hogy ezek jelentéséből elvonható-e olyan metapredikátum, melyet érdemes beépíteni a lexikai szemantikai reprezentációba, mert más szintaktikai jelenségek leírásában is használhatók, vagy az igék csak idioszinkratikus szemantikai és szintaktikai tulajdonságaikban hasonlítanak. Az általánosítás lehetőségét az veti fel, hogy az igeosztály alcsoportjai ugyanazt a szemantikai szerepet fejezik ki a mellettük leggyakoribb esetragokkal.

A spray/load igék 10 százaléka részt vesz az angol csoportra jellemző alternációban, és összesen több mint 50 százalékra jellemző a két minta legalább egyike. Ahhoz, hogy igeosztályként elfogadjuk ezt a csoportot, először meg kell vizsgálnunk, miért nem mutat hasonló adatokat a csoport igéinek másik része: valóban nem vehetnek részt az alternációban, vagy esetleges másfajta használatuk/másik jelentésük torzította a korpusz adatait?

A szemantikai igeosztályok tesztelésekor elsősorban arra voltunk kíváncsiak, hogy találunk-e szintaktikai hasonlóságot az angol adatok alapján egy osztályba sorolt igék között. A két igeosztály eredményeinek összehasonlítása alapján erre igenel válaszolhatunk, ám az osztályokon belüli egyes igék közti eltérések és a minták előfordulásainak csekély száma mindenképp azt vonja maga után, hogy az osztályok kialakításánál kézi ellenőrzésre is szükség van. Tervezzük egyéb igeosztályok magyarra fordítását és a csoportok szintaktikai viselkedésének tesztelését is. A vizsgálatoktól azt várjuk, hogy kiderüljön, vannak-e és melyek azok az igei jelentéskomponensek, amik a magyar igék szintaxisának leírásában hasznos általánosítások megragadására szolgálhatnak. Ha bebizonyosodik az, hogy az angol igeosztályokat karakterizáló metapredikátumok a magyarban is szerepet játszanak az ige szintaktikai környezetének meghatározásában, akkor a szintaktikai elemző számára kulcsfontosságú igeosztályok létrehozásakor támaszkodhatunk az angol csoportokra, és ezáltal szükségtelemmé válik a magyar alternációk részletes feltérképezése.

<sup>63</sup> Természetesen találtunk kivételeket, például a “megbékél” ige mellett előfordulásainak több, mint 40 százalékában, a “megbékül” mellett pedig 51 százalékban szerepel az INS esetrag.

## Bibliográfia

1. Chomsky, N.: *Lectures on Government and Binding*, Foris, Dordrecht 1981
2. Foley, W. A., R. D. Van Valin, Jr. *Functional Syntax and Universal Grammar*, Cambridge University Press, Cambridge 1984
3. Gábor, K., Héja, E.: Vonatok és szabad határozók szabályalapú kezelése. *In: Alexin Z., Csendes D. (szerk): A Harmadik Magyar Számítógépes Nyelvészeti Konferencia Előadásainak kötete*, Szeged Egyetemi nyomda, 2005. Szeged, pp. 245-257 .
4. Kaplan, R. M. and J. Bresnan: Lexical-Functional Grammar: a Formal System for Grammatical Representations. In J. Bresnan (ed.) pp. 173-281. [1982]
5. Levin, B., M. R. Hovav: *Argument Realization*. Cambridge University Press, Cambridge. 2005
6. Levin, B.: English Verb Classes and Alternations: A Preliminary Investigation. *Int. J. Digit. Libr.* 1 (1997) 108–121
7. Silberztein, M.: *Dictionnaires électroniques et analyse automatique de textes: Le systeme Intex*. Masson, Paris 1993.
8. Váradi, T.: The Hungarian National Corpus. *Proceedings of the Third International Conference on Language Resources and Evaluation*, 2002. Las Palmas pp.385-389
9. Váradi, T., Gábor, K.: A magyar INTEX fejlesztésről. *In: Alexin Z., Csendes D. (szerk): A Második Magyar Számítógépes Nyelvészeti Konferencia Előadásainak kötete*, Szeged Egyetemi nyomda, 2004. Szeged, pp. 3-11.