

**KALIVODA ÁGNES**

**IGEKÖTŐS SZERKEZETEK A MAGYARBAN**

**Doktori (PhD) értekezés**

**Pázmány Péter Katolikus Egyetem  
Bölcsészet- és Társadalomtudományi Kar  
Nyelvtudományi Doktori Iskola**

**Vezetője: Dr. Surányi Balázs**

egyetemi tanár, az MTA doktora

**Nyelvtechnológia Műhely**

**Témavezető**

**Dr. Prószéky Gábor**

egyetemi tanár, az MTA doktora

Budapest, 2020

# Tartalomjegyzék

<b>A fonológiai jelölések jegyzéke</b>	<b>5</b>
<b>1. Bevezetés</b>	<b>6</b>
<b>2. Módszertan</b>	<b>8</b>
2.1. Korpuszvezéreltség . . . . .	8
2.2. A kutatás alapjául szolgáló erőforrások . . . . .	9
2.2.1. Magyar Nemzeti Szövegtár . . . . .	9
2.2.2. A PREVLEX táblázat . . . . .	12
2.3. Összegzés . . . . .	15
<b>3. Az igekötő-állomány meghatározása</b>	<b>16</b>
3.1. A meghatározás problémái . . . . .	16
3.2. Az igekötők prototípus-elméleti keretben . . . . .	18
3.3. Az igekötőség mérhető jegyei . . . . .	20
3.3.1. Morfológiai produktivitás . . . . .	21
3.3.2. Reduplikáció . . . . .	24
3.3.3. Elválás deverbális főnévtől tagadás esetén . . . . .	28
3.3.4. A <i>-fele/felé</i> szócska csatolhatósága . . . . .	29
3.3.5. Fokozhatóság . . . . .	30
3.3.6. Személyragozhatóság . . . . .	32
3.3.7. Általános jegyek: szerkesztettség, szótagszám és gyakoriság . . . . .	33
3.4. Igekötő $\times$ Jegy mátrixok . . . . .	35
3.5. Jegyek közötti korrelációk . . . . .	37
3.6. Az igekötők jegyalapú osztályozása . . . . .	42
3.6.1. A jegyek súlyozásának kérdései . . . . .	42
3.6.2. Egyszerű matematikai műveletek . . . . .	44
3.6.3. Klaszteranalízis . . . . .	46
3.6.4. A módszerek közötti választás . . . . .	48

3.7. Összegzés . . . . .	50
<b>4. Az igekötős szerkezetek típusai és szintaktikai jellemzői</b>	<b>54</b>
4.1. Szinkrón vizsgálat: a mai magyar nyelvállapot . . . . .	55
4.1.1. A kutatás módszere . . . . .	55
4.1.2. Finit ige . . . . .	57
4.1.3. Infinitívusz . . . . .	66
4.1.4. Határozói igenév . . . . .	70
4.1.5. Melléknévi igenév . . . . .	73
4.1.6. Deverbális főnév . . . . .	81
4.2. Diakrón vizsgálat: az ómagyar kortól napjainkig . . . . .	83
4.2.1. A kutatás módszere . . . . .	84
4.2.2. A prototipikus igekötők pozíció szerinti eloszlása . . . . .	86
4.2.3. A kétféle tagadó szórend arányának változása . . . . .	89
4.3. Összegzés . . . . .	91
<b>5. Az igekötők produktív kapcsolódási mintái</b>	<b>94</b>
5.1. A kutatás módszere . . . . .	95
5.2. Az igealkotás produktív esetei . . . . .	96
5.2.1. Névszó + igeképző . . . . .	97
5.2.2. Ige + igeképző . . . . .	100
5.2.3. Hangzásséma . . . . .	101
5.2.4. A vizsgált igealkotási módok arányai az igekötős szerkezetekben . . . . .	105
5.3. PREVCONS: Igekötős szerkezetek táblázata . . . . .	106
5.4. Az igekötők ontológiája . . . . .	111
5.4.1. A poliszémiából adódó nehézségek . . . . .	113
5.4.2. Ontológiai entitások: igekötők és jelentések . . . . .	115
5.4.3. Szemantikai relációk . . . . .	117
5.5. Összegzés . . . . .	120
<b>6. Összefoglalás, új tudományos eredmények</b>	<b>122</b>

<b>Hivatkozások</b>	<b>127</b>
<b>Mellékletek</b>	<b>137</b>
I. számú melléklet . . . . .	137
II. számú melléklet . . . . .	138
<b>Magyar nyelvű összefoglaló</b>	<b>155</b>
<b>Summary in English</b>	<b>156</b>

## A fonológiai jelölések jegyzéke

C	mássalhangzó
C:	hosszú mássalhangzó
V	<i>o, ö, e</i> vagy <i>a</i>
V <sub>1</sub> ... V <sub>1</sub>	azonos magánhangzók ( <i>o, ö, e</i> vagy <i>a</i> )
A	<i>a</i> vagy <i>e</i>
O	<i>o, ö</i> vagy <i>e</i>
U	<i>u</i> vagy <i>ü</i>
Ó	<i>ó</i> vagy <i>ő</i>
/.../	fonéma (pl. /j/ → <i>j</i> vagy <i>ly</i> )
(...)	nem mindig megjelenő elem

# 1. Bevezetés

„A tárgy [...] rendkívül aktuális, újszerű és hatalmas szakirodalma van, ugyanakkor még senki nem foglalkozott vele.”

(Kálmán C. György: Módszertani útmutató opponensi véleményhez)

A mottóként idézett önellentmondás viccből hangzott el, mégis nagyon illik ehhez a dolgozathoz. A magyar nyelv igekötős szerkezeteivel sokan és sokféleképpen foglalkoztak már: elsőként Sylvester Jánost említhetjük, aki az 1539-es *Grammatica Hungarolatina* című művében *előlvető* névvel tárgyalta az igekötőt a kor szellemének megfelelően, a latin prepozíciók mintájára (Jakab 1982: 11-12). A későbbi nyelvtanírók fokozatosan távolodtak a latin nyelvtani szemlélettől, és felismerték az igekötők olyan – főként mondattani – tulajdonságait, amelyek alapján nem rokoníthatók a latin prepozíciókkal. Ezeket a felismeréseket tükrözik a jelenség későbbi megnevezései is: például a *Debreceni Grammatika* (1795) *elegyes előlutóljáró*, Virág Benedek (1810) *elő- s utol-vető* műszóval hivatkozik az igekötőre (Jakab 1982: 68-69).

Napjaink nyelvészeti szakirodalmában is élénk érdeklődés mutatkozik az igekötők iránt. Kerekes (2011: 110) a következőt írja erről: „A szakirodalom egységes abban, hogy egy nagyon egyedi, mégis összetett, mi több, ellentmondásos kategóriával van dolgunk.” Ezzel lényegében össze is foglalta azt, amiben a szakirodalom az igekötőket tekintve egységes. Máskülönben nagyon eltérő nézetekkel találkozunk, például azokban a kérdésekben, hogy az igekötő önálló szófaj-e vagy sem, önálló szó-e vagy képzőszerű elem, és egyáltalán melyik szavakat soroljuk az igekötők közé.

A mottó második fele azért állja meg itt a helyét, mert az igekötők állományának és az igekötős szerkezeteknek korábban nem készült korpuszvezérelt leírása. A dolgozatomban ennek a megvalósítására törekszem. A korpuszvezéreltség egy hangsúlyozottan adatközpontú szemlélet, amelyben a kutatás nagy mennyiségű szöveganyag automatikus elemzéséből indul ki, és elsődleges célja az introspekcióval nem feltárható jelenségek kimutatása. A kutatásom során a legtöbb ötletet a kognitív nyelvészetből merítem, de nem köteleződök el egyetlen nyelvészeti irányzat mellett sem. Ezzel kapcsolatban Ladányi (2007: 13) gondolatát szeretném idézni – hozzá hasonlóan én is vállalom inkább „az eklekticizmus vádját, mintsem egyetlen választott elméleti keret (esetleg a szükségesnél szűkebbre szabott) korlátait.”

A dolgozat következő fejezetében a kutatás módszertani háttérét ismertetem. Ezután egy-egy nagy fejezetben foglalkozom a főbb kutatási kérdéseimmel, amelyek a következők:

1. Mely lexikai elemeket soroljuk az igekötők közé, és mi alapján végezzük a besorolást?
2. Milyen mintázattípusokat mutatnak az igekötős szerkezetek, mikor és mennyire távolodhat el az igekötő az igétől, igenévtől, igei származéktól?
3. Hogyan írhatjuk le az igekötők produktív kapcsolódási mintáit, és ez alapján milyen megállapítások tehetők az igekötők szemantikájáról?

A dolgozatot a három kérdésre adott válaszaim összefoglalásával és a téziseim megfogalmazásával zárom.

## 2. Módszertan

„Adatod lehet információ nélkül,  
de információd nem lehet adat nélkül.”

(Daniel Keys Moran)

Ebben a fejezetben röviden ismertetem azt az erősen adatközpontú szemléletet, amely a dolgozatom egészét jellemzi. Ezután bemutatom azokat az erőforrásokat, amelyeket a teljes kutatás során felhasználtam. Az egyes kutatási kérdések megválaszolásához használt eszközöket és módszereket mindig a releváns fejezeteknél tárgyalom.

### 2.1. Korpuszvezéreltség

Egy-két évtizede még élénk viták folytak arról, hogy a korpuszok – különféle szempontok szerint összeállított szövegtárak – szükségesek-e a nyelvi jelenségek vizsgálatához, mára viszont a nyelvészet legtöbb területén magától értetődővé vált ezek használata. A kompetenciát, vagyis az anyanyelvi beszélő nyelvi képességét kutató nyelvészek körében is egyre jellemzőbb, hogy nem kizárólag introspekcióval dolgoznak, hanem korpuszadaton is ellenőrzik a hipotéziseiket.

A korpuszok használatának természetesen nem az az ideális módja, amikor a korpuszt csak arra használjuk fel, hogy az elméletünket alátámasztó példákat gyűjtsünk belőle. A korpusznyelvészet túl akar lépni ezen: törekszik olyan jelenségek leírására is, amelyekre – egyelőre – nincs magyarázat (Teubert 2005: 125-126). Így aztán fontosnak tartja a performancia-hibákat, valamint minden olyan jelenséget is, amely a kompetencia szempontjából anomáliák sorának tűnhet. Ilyenekre bőven lesz példa ebben a dolgozatban.

Francis (1993) óta kétféle módszert szokás megkülönböztetni a korpuszvizsgálatokban. Az egyik a korpuszalapú (corpus-based), amellyel egy kialakult elméletet tesztelünk, és az adatok alátámasztják vagy cáfolják az elméletet. A másik a korpuszvezérelt (corpus-driven), amely kifejezetten azt hangsúlyozza, hogy előzetes elmélet nélkül kezdjük a kutatást, introspekcióval nem feltárható jelenségek kimutatása a cél, és tisztán az adatból kiindulva, indukción keresztül jutunk eredményre. A két módszer alapos leírását adja Sass (2011: 15-17).

Stubbs (2013) úgy véli, hogy a korpuszvezérelt módszert jellemzően objektívebbnek tartjuk a korpuszalapúnál, de gyakran túlhangsúlyozzuk az indukciónak a szigorúságát. Tanulmányában kifejti,



hogy a tiszta indukciót már több mint négyszáz éve megkérdőjelezzük, az elmélettől mentes indukció lehetőségét pedig – legalábbis a társadalomtudományokban – általánosan elutasítják. Arra szeretné felhívni a figyelmet, hogy egy jelenség vizsgálatánál tulajdonképpen sosem a semmiből indulunk ki. Ha nincs legalább minimális elképzelésünk arról, hogy mit keresünk, akkor kevés esélyünk van arra, hogy megtaláljuk.

Az általam alkalmazott módszer alapvetően korpuszvezérelt, de nem mentes bizonyos előfeltevésektől. Feltételezem például, hogy az igekötő bármi olyan szótól elválhat, amely igekötős igéből képzett. Természetesen ez ilyen formában nem lesz igaz, de jó kiindulópontja a korpuszméréseknek, és segíti az olyan szerkezetek feltárását, amelyekkel a szakirodalom ritkán vagy esetleg egyáltalán nem foglalkozik. Emellett több esetben elméleti forrás adja az ötletet egy korpuszvizsgálomhoz (ilyen például a növekvő összetevők törvényével kapcsolatos korpuszmérés, bővebben ld. a 4.1.2. fejezetben). Úgy vélem, hogy ennek a kutatásnak előnyére válik az induktív és deduktív módszerek keveredése.

## **2.2. A kutatás alapjául szolgáló erőforrások**

A továbbiakban arról a két erőforrásról lesz szó, amelyet a disszertációhoz kapcsolódó összes kutatásomban felhasználtam. Elsőként bemutatom a Magyar Nemzeti Szövegtár 2.0.4 verzióját (Oravecz et al. 2014), amelyen a korpuszméréseim alapulnak. Ezután egy általam létrehozott lexikai erőforrásról lesz szó, a PREVLEX táblázatról (Kalivoda 2019). A PREVLEX 53 503 igekötős igét tartalmaz a gyakorisági adataikkal együtt. Manuálisan ellenőriztem, így a minősége és a mérete is alkalmassá tesz arra, hogy a különböző korpuszmérések során a hibás találatok jelentős részét ki lehessen szűrni a segítségével.

### **2.2.1. Magyar Nemzeti Szövegtár**

A 20-21. századi szövegeket tartalmazó Magyar Nemzeti Szövegtár 2.0.4 verziója – a továbbiakban MNSZ2 – összesen 1,348 milliárd tokenből áll, amely írásjelek nélkül számolva 1,04 milliárd szövegszót jelent.<sup>1</sup> A korpusz hat stílusrétegből tartalmaz szövegeket: sajtó, szépirodalom, tudo-

<sup>1</sup> Az MNSZ korpusznak 2018. június 8-án megjelent egy újabb verziója, a 2.0.5-ös. Ez az összetételét tekintve nem változott az általam használt 2.0.4-es verzióhoz képest, viszont a meglévő szöveganyagban történtek javítások. Erről bővebben itt lehet tájékozódni: <http://clara.nytud.hu/mnsz2-dev/hirek.html>

mányos, hivatalos, személyes, beszélnyelvi. Ezen belül öt regionális nyelvváltozatra tagolódik: magyarországi, szlovákiai, kárpátaljai, erdélyi, vajdasági.<sup>2</sup>

A teljes szöveganyag mondatokra bontott, lemmatizált – tehát a szóalakok vissza vannak vezetve szótári alakokra –, valamint morfológiailag és szófajilag elemzett. Ennél magasabb szintű annotációt – pl. szintaktikai elemzést – a korpusz nem tartalmaz.

Két okból döntöttem az MNSZ2 használata mellett. Az első az volt, hogy a szövegek gazdagon el vannak látva meta-adatokkal is, amelyek közül többet – főleg a stílusréteggel kapcsolatos adatokat – külön is vizsgáltam, amikor ez a korpuszmérések eredménye alapján indokoltnak tűnt. A másik ok teljesen gyakorlati jellegű: lehetőséget kaptam arra, hogy közvetlenül az MNSZ2 forrásfájllal dolgozzak, a webes lekérdező felület használata nélkül. Ez egyrészt meggyorsította a különféle kereséseket, másrészt lehetővé tette azt, hogy a korpusz annotációját a kutatás igényeihez igazodva módosítsam. Erre a következő esetekben volt szükség:

1. Minden igekötős finit igénél és igenévnél, ahol ugyan látszik az annotációban, hogy igekötős a szó (az elemzés IK.-val kezdődik), de sem a szóalakban, sem a lemmában nem látszik, hogy hol van az igekötő és az igei rész közötti határ. Így például automatikus adatfeldolgozás szempontjából a *felültettek* szó lehet *felül+tettek* és *fel+ültettek* is.
2. Olyan igekötős igéknél, igeneveknél és igei származékoknál, amelyek UNKNOWN elemzést kapnak, pedig egyébként létező szavak (pl. *bedepizni*, *összelájkolnotok*), erről bővebben lesz szó következő fejezetben.
3. Igei származékoknál – amelyek nem finit igeik vagy igenevek – nem derül ki az MNSZ2 jelenlegi annotációjából, hogy a képzés bemenetül szolgáló szónak van-e igekötője vagy sem. Tehát például a *megértettség* szó elemzése FN.NOM (alanyesetű főnév), az *elvárhatóan* elemzése HA (határozószó). Ezek mind helyes annotációk, de kevésnek bizonyulnak akkor, ha az igekötő szintaktikai viselkedését akarjuk tanulmányozni – és a példaként hozott két szerkezet állhat elvált igekötővel is: *meg nem értettség*, *el is várhatóan*.
4. A -hAtÓ végződésű igeneveknél, amelyek mindig melléknévként annotáltak, pedig állítmányi helyzetben nagyon hasonlít a viselkedésük a finik igeikéhez: az igekötőjük elválhat, le-

<sup>2</sup> Az egyes alkorpuszok méretére vonatkozó adatok ezen a linken érhetők el: <http://clara.nytud.hu/mnsz2-dev/stat.html>

het hátravetett, és posztverbálisan messzire is kerülhet az igenevektől (erről bővebben ld. a 4.1.5. fejezetet).

Az említett módosításokon túl három szűrést végeztem a forrásfájlon annak érdekében, hogy a lehető legjobb minőségű szöveganyagot kapjam. Egyrészt kiszűrtem a verseket, mivel sokuk nem természetes nyelvhasználatot tükröz. Másrészt – amennyire csak lehetett – eltávolítottam az idegen nyelvű, valamint a magyar, de ékezet nélkül írt mondatokat, mert torzíthatták volna a keresések eredményét. Például az ékezetet eleve nem tartalmazó igekötős igék sokkal gyakoribbnak tűntek volna, mint az ékezetesek. Ehhez azt a heurisztikát alkalmaztam, hogy töröltem minden olyan mondatot, amelyben a tokenek 80%-a UNKNOWN vagy SKIP elemzést kapott. Ez a módszer inkább a pontosságnak, mintsem a fedésnek kedvezett. Végül igyekeztem kiszűrni a korpuszban található duplumokat. Itt is a pontosságot tartottam szem előtt. Csak a nyolc tokennél hosszabb mondatokat vettem figyelembe a szűrésnél, feltételezve, hogy ennél rövidebb mondatoknál – pl. köszönéseknél – természetes lehet a többszörös jelenlét. Még ezzel az óvatos módszerrel is rendkívül magasnak bizonyult a duplumok aránya (20,12%), a személyes alkorpuszon belül akadt olyan – meglehetősen hosszú – mondat, amely száznál többször ismétlődött. A szűrések eredményét az 1. táblázat foglalja össze. Amikor az MNSZ2-re hivatkozok a dolgozat további részeiben, mindig az itt bemutatott, módosított verziót értem ezalatt.

<b>korpusz</b>	<b>token</b>	<b>százalék</b>
<b>eredeti MNSZ2</b>	<b>1 348 000 000</b>	<b>100</b>
versek	5 661 000	0,42
UNKNOWN/SKIP	26 825 200	1,99
duplumok	271 217 600	20,12
<b>módosított MNSZ2</b>	<b>1 044 296 200</b>	<b>77,47</b>

**1. táblázat.** Az MNSZ 2.0.4 mérete a versek, értelmes elemzés nélküli mondatok, valamint a duplumok szűrése előtt és után. A tokenszám írásjelekkel együtt értendő.

Úgy ítélt meg, hogy az MNSZ2 jelenleg az egyik legjobb minőségű, mai magyar nyelvállapotot reprezentáló korpusz, és az itt bemutatott módosítások után alkalmasnak találtam a kutatás céljaira.

### 2.2.2. A PREVLEX táblázat

Az MNSZ2 előző fejezetben bemutatott, módosított verziójából hoztam létre a PREVLEX-et,<sup>3</sup> amely az igekötős igék jelenleg legbővebb (53 503 szavas), manuálisan ellenőrzött táblázata. Részt képezik a hapaxok – egyszer előforduló szavak –, valamint azok a szavak is, amelyek a korpuszban UNKNOWN címkével jelennek meg. Először arról lesz szó, hogy milyen munkafolyamattal, milyen megfontolások mentén készült a PREVLEX. Ezután az erőforrás felépítését ismertetem.

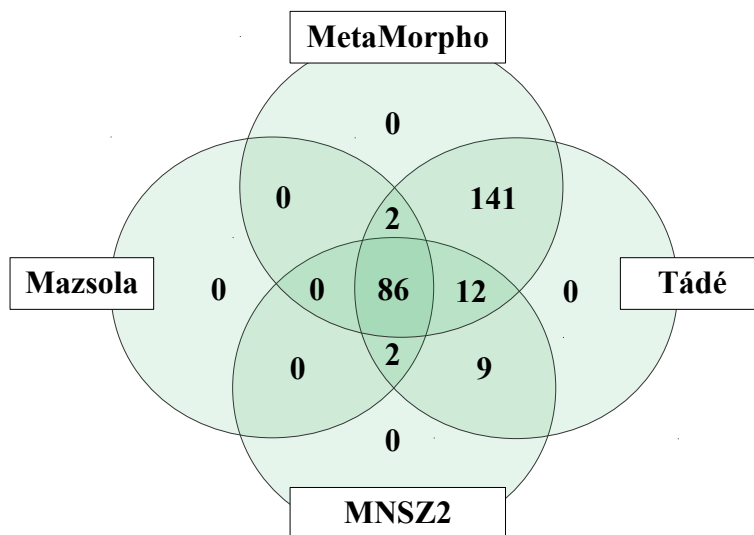
A PREVLEX előállításához először is szükség volt a lehetséges igekötők listájára. Úgy döntöttem, hogy ezt a Manócska<sup>4</sup> integrált igei vonzatkeret adatbázisból nyerem ki (Kalivoda et al. 2018). Ebben minden olyan lexikai elem szerepel, amelyet a magyar igei vonzatkerettárak közül legalább egy igekötőnek jelöl (a továbbiakban igekötőnek nevezem őket az egyszerűség végett). A Manócska készítése során a szerzőtársaimmal átnéztük ezeket a szavakat, és javítottuk az egyértelműnek tűnő hibákat (pl. a *vissz*, *nyug* igekötőnek jelölését a *visszhangoz*, *nyugdíjaz* szavak esetében). Ennek ellenére a végső lista hosszú – összesen 252 tagot számlál –, és több mint kétharmada esetében (pl. *szénné*, *pofon*, *zsebre*) az igekötői státusz erősen vitatható. Ez egyúttal jól tükrözi azt is, mennyire nincs egyetértés abban, hogy mely szavakat tekintjük igekötőnek (a problémakörrel részletesen foglalkozok a 3. fejezetben). A 1. ábra azt szemlélteti, hogy melyik erőforrás hány igekötőt nevez meg. A MetaMorpho (Prószéky et al. 2004) és a Tádé (Kornai et al. 2016) kezelik a legtágabban ezt a kategóriát. Az MNSZ2 alapján készült listák (Kalivoda 2016) feleannyi jelöltet sem tartalmaznak, a legszigorúbb pedig a Mazsola (Sass et al. 2010; Sass 2015), amelyben 90 igekötő szerepel.

A kiinduló, 252 szavas listában 17 hibát találtam (ilyen pl. a *vízi*, amely egyszerűen bizonyos szóösszetételek első tagja). Így végül 235 szó maradt, amely az ismertetett források valamelyike szerint igekötő, és a kutatásnak ebben a szakaszában én is fenntartottam ennek a lehetőségét.

A következő lépésben lekértem az MNSZ2-ből minden olyan finit igeként vagy UNKNOWN-ként annotált szót, amely egy adott igekötővel kezdődik. Ennek a döntésemnek két része is magyarázatra szorul. Először is az, hogy miért csak a finit igéket vettem figyelembe, amikor az igekötők például igenevekhez is kapcsolódhatnak. A korpuszvizsgálat során azt tapasztaltam, hogy az igenevek esetében erős a tendencia az igemódosító és az igenév egybeírására (pl. *jóltáplált vendég*, *földreszállt*

<sup>3</sup> <https://github.com/kagnes/prevlex>

<sup>4</sup> <https://github.com/ppke-nlpg/manocska>



**1. ábra.** A Manócskában szereplő erőforrások összesen 252 szót minősítenek igekötőnek. A halmazok azt mutatják, hogy az egyes erőforrások hány másikkal és hány darab szót illetően értenek egyet.

*angyal*), míg ugyanezeket az igemódosítókat a finit igével már kevésbé írják egybe. Valószínű, hogy az igenevek figyelembevétele nem változtatott volna jelentősen a PREVLEX összetételén, viszont sokkal több ellenőrizendő adathoz vezetett volna. Másodszor, az UNKNOWN szavakra azért volt szükség, mert sok jó találat csak így jelenik meg (pl. *visszacuccol*, *felstócol*, *benyammog*). Ugyanakkor az UNKNOWN szavak legnagyobb része hibás találat (elírt vagy idegen nyelvű szó), és a finit igék között is akadnak álpozitív találatok (pl. a *túlélősködik* mint igekötős ige). Emiatt az eredményül kapott, közel 178 000 szavas listát át kellett nézni.

Ez a munka körülbelül huszonegy órát vett igénybe. Először eltávolítottam a lehetséges igekötőket a szavak elejéről, és a megmaradó szórészeket néztem át aszerint, hogy egyáltalán igék-e vagy sem. Ezután a már jóval rövidebb listát átnéztem úgy, hogy az ige az adott igekötővel is megfelelő-e (ezen a szinten szűrtem ki pl. a *túlélősködik* és *feltűnősködik* igéket). Néhány olyan esetben, ahol az igekötő+ige kombináció nem volt értelmetlen, viszont nagyon valószínűtlennek tűnt, csak a konkrét szövegbeli előfordulások segítségével tudtam dönteni (pl. a *túltejesít*-ről így derült ki, hogy mindig a *túlteljesít* hibásan írt változata). Ezután lokálisan újraelemeztem a forrásfájlt a javított adatokkal (pl. a korábban UNKNOWN *hype-olok*, *hype-ol* szavakat összevontam egy lemmává). A javított korpuszból állt elő a PREVLEX végső változata.

Az adatok átnézése során többször felmerült a kérdés, hogy bizonyos szóalakokat nem kellene-e valahogyan normalizálni. Három esetben az igező okozott bizonytalanságot, mert (1) teljesen

azonos jelentésű igék történetileg eltérő tőváltozattal rendelkeznek (pl. *verekedik* – *verekszik*), (2) két igenek minimálisan eltérő töve van (pl. *gyömöszköl* – *gyömöcköl*), (3) egy-egy neologizmus többféle írásváltozatban létezik (pl. *dizájno*l – *design-ol* – *designol*). Egyedül az utóbbi csoport kapcsán voltam biztos abban, hogy a különbség csak ortográfiai jellegű. Ezeket a szavakat normalizáltam – rendszerint a magyar kiejtés szerint írt változatra –, mindenhol megőrizve az eredeti szóalakot is.

Elkülöníthető továbbá három olyan probléma, amely a képzőt érinti: amikor (1) két vagy több ige képzőjében csak a kötőhang tér el (pl. *feccel* – *feccöl* – *feccol*), (2) opcionálisan -ikes végződésű az ige (pl. *szörföz* – *szörfözik*), (3) ugyanaz az ige -Vz(ik) és -Vl képzővel is előfordul (pl. *offtopicol* – *offtopicoz*). Bár itt is szólhatnak érvek a normalizálás mellett, annyi biztos, hogy nem egyszerű ortográfiai különbségekről van szó. A (3)-asban látható példák egyelőre még ugyanazt jelentik, de elképzelhető, hogy idővel kis jelentésbeli eltérés kapcsolódik hozzájuk (ahogy azt pl. a *házal* – *házaz* párnál látjuk). A normalizálást ezekben az esetekben önkényesnek találtam, és nem vállalkoztam rá.

Az erőforrás egy TSV fájlként érhető el, amely öt oszlopból áll. Az első oszlopban szerepel az ige (igekötő+igelemma formában). Ezt követi az MNSZ2-ben mért tokengyakoriság. A harmadik oszlopban kétféle érték szerepelhet attól függően, hogy az ige kapott-e megfelelő annotációt az MNSZ2-ben (FIN, ha igen és UNKNOWN, ha nem). A negyedik oszlop azt jelzi, hogy az ige hány dokumentumban fordult elő. Ez fontos információ lehet akkor, ha a tokengyakoriság és a tartalmazó dokumentumok száma nincs arányban (pl. az ige százszor fordul elő, de mindössze egy dokumentumban). Utolsóként szerepel a normalizált alak, amely csak a neologizmusoknál térhet el az első oszlop tartalmától.

Bár az igekötős igék listája manuálisan ellenőrzött, a gyakorisági adatok fenntartással kezelendők. Néhány igealak ugyanis több lehetséges elemzéssel rendelkezik (pl. a *leszel* egyik lehetséges elemzése a *lenni* E/2. alakja, a másik a *leszel* igekötős ige). Ezek az elemzések sokszor eleve rosszak a forrásfájlban, így kissé torzíthatják a gyűjtött statisztikát.

A 2. táblázat számszerű áttekintést ad a PREVLEX-ről. A várakozásnak megfelelően az igekötős igék Zipf-eloszlást (Zipf 1932) mutatnak: néhány ige rendkívül nagy tokengyakorisággal bír, míg a hapaxok ritkák, de sokfélék.

<b>kategória</b>	<b>típus</b>	<b>token</b>
összes igekötős ige	53 503	12 149 637
hapaxok	21 081	21 081
UNKNOWN szavak	5 075	25 189
UNKNOWN hapaxok	3 508	3 508

**2. táblázat.** A PREVLEX számokban. Az értékek az eredeti igealakokra vonatkoznak, nem a normalizáltakra.

## 2.3. Összegzés

Ebben a fejezetben bemutattam az igekötős szerkezetekkel kapcsolatos kutatásaim módszertani hátterét. Röviden ismertettem a korpuszvezérelt megközelítést, és jeleztem, hogy bár alapvetően ez a szemlélet jellemzi a munkámat, a megállapításaim jelentős részéhez nem tisztán induktív általánosítással jutottam el, hanem különböző elméletek is inspiráltak. Ezután bemutattam a Magyar Nemzeti Szövegtár 2.0.4-es verzióját, amelyen a korpuszméréseim alapulnak, majd felsoroltam a módosításokat, amelyeket el kellett végezni a korpuszon ahhoz, hogy a kutatás céljaira alkalmas legyen. Végül beszámoltam az 53 503 szavas PREVLEX táblázatról, amelyet az MNSZ2 felhasználásával hoztam létre, és jelenleg az igekötős igék legbővebb gyűjteményének tekinthető.

### 3. Az igekötő-állomány meghatározása

„Ha lenne két nyelvész,  
aki mindenben egyetértene,  
az egyik fölösleges lenne.”

(Mádl Péter egy svéd nyelvórán)

Egy tudományos munkával szemben elvárás, hogy a szükséges fogalmakat, terminusokat először definiálja, majd a definíció szerint használja őket. Természetesen én sem írhatok meg egy igekötőkről szóló disszertációt úgy, hogy nem határozom meg, egyáltalán mit értek ‘igekötő’ alatt. Csakhogy erre a jelenségre nagyon ráillik a fenti mottó: a kapcsolódó szakirodalmat áttekintve alig találni arra példát, hogy két szerző azonos szempontok mentén határozná meg az igekötőket, és ugyanazokat a szavakat sorolná ebbe a kategóriába. Így a dolgozatomban jelen fejezetében két kérdés megválaszolására törekszem. Az első az, hogy mely lexikai elemek tartoznak az igekötőkhöz, a második – és ez a hangsúlyosabb –, hogy mi alapján sorolunk egy szót az igekötők közé.

Elsőként nagy vonalakban áttekintem az igekötők meghatározásának elméleti és gyakorlati problémáit. Ezután – immár az általam javasolt megoldáshoz közelítve – ismertetem a prototípus-elméletet, amely lehetővé teszi az elmosódott kategóriahatárok rugalmas kezelését. A továbbiakban azzal foglalkozom, hogy a prototípus-elmélet hogyan ültethető át a gyakorlatba. Ennek kulcsfontosságú pontja az igekötőség mérhető jegyeinek megállapítása, majd a jegyek vizsgálata korpuszadaton. A fejezet legfontosabb szakasza az egyes jegyek részletes leírása, valamint ezek együttállásainak, korrelációinak megállapítása. Ezt követően arról lesz szó, hogy ha egyszerű gyakorlati okokból meg akarjuk húzni a határvonalat az igekötők és más igemódosítók között, arra milyen lehetőségeink vannak – az erre irányuló módszereket jegyalapú osztályozásnak nevezem. A fejezetet az eredmények összefoglalása zárja.

#### 3.1. A meghatározás problémái

A definíciós problémák azzal a kérdéssel kezdődnek, hogy az igekötő önálló szófaj-e vagy sem. A hagyományos nyelvtanok az előbbi vallják, míg például a generatív szakirodalomban az a jellemző, hogy az igekötő terminus a határozószóknak egy alosztályát jelenti (mint ahogyan a tulajdonnevek a főneveknek egy alosztályát képviselik). Eszerint igekötők azok a határozószók, amelyek



semleges mondatban igemódosító pozícióban – közvetlenül a finit ige előtt – állnak.

Akár önálló szófajnak, akár a határozószók egyik típusának tekintjük az igekötőt, fennáll az a helyzet, hogy nincs éles kategóriahatár. Az első esetben az a kérdés, hogy az igekötőket hogyan lehet elhatárolni más szófajú szavaktól, amelyek hozzájuk hasonló funkciót töltenek be, a második esetben pedig az, hogy az igekötő-szerűvé vált, pusztá névszói igemódosítók szófaja névszó vagy határozószó. Ezért nem részletezem tovább a kérdést – látni fogjuk, hogy az általam javasolt megközelítésnek akkor is megvan a létjogosultsága, ha az igekötőt önálló szófajnak, és akkor is, ha a határozószó egyik típusának tekintjük.

Az igekötő-állomány meghatározását az teszi problémássá, hogy az igekötők jelentéstani és mondattani viselkedésük alapján is hasonlítanak más lexikai elemekhez, amelyeket a továbbiakban pusztá névszói igemódosítónak nevezek (Kiefer 2007: 226-227). A két kategória szintaktikai hasonlóságát a 3. táblázat szemlélteti. Ez a hasonlóság még sok másra is kiterjed a mondatbeli pozíción túl, például arra, hogy jellemzően milyen szavak férkőzhetnek be az ige és az igemódosítója közé. Erről részletesen lesz szó a 4. fejezetben.

Igekötő	Pusztá névszói igemódosító
<b>Kifestem</b> a kerítést.	<b>Kékre festem</b> a kerítést.
<b>Ki</b> is <b>festem</b> a kerítést.	<b>Kékre</b> is <b>festem</b> a kerítést.
Nem <b>festem ki</b> a kerítést.	Nem <b>festem kékre</b> a kerítést.

**3. táblázat.** Az igekötő és a pusztá névszói igemódosító szintaktikai hasonlósága.

A 3. táblázatban szereplő *kékre* nem szokott felmerülni kérdéses esetként, de mi a helyzet az olyan szavakkal, mint például az *észre* (*észre+vesz*) vagy a *pofon* (*pofon+vág*)? Jakab (1976: 9-10) hét forrásművet vizsgál abból a szempontból, hogy melyik milyen szavakat sorol az igekötők közé. Arra jut, hogy a forrásművek összesen 28 elemben értenek egyet, ez a vizsgált 75 elemű listának mindössze 37,3%-a.<sup>5</sup> Az egyetértés hiánya lényegében változatlanul fennáll Jakab 1976-os tanulmánya óta is, erről bővebben ld. Komlósy (1992) és Kiefer (2007) munkáját.

A bizonytalanság nemcsak elméleti szinten jelentkezik, hanem a gyakorlatban is megnyilvánul. Ezzel kapcsolatban visszautalnék a 2. fejezetben bemutatott az 1. ábrára, ahol látható, hogy kü-

<sup>5</sup> A 28 elem a következő: *abba, agyon, alá, be, bele, el, fel, félbe, félre, felül, fenn, hátra, helyre, hozzá, ide, keresztül, ki, körül, közbe, le, meg, neki, oda, össze, rá, rajta, túl, vissza*. Ha ezt a listát további szakirodalmakkal vetnénk össze, még inkább csökkenne a kollektívan igekötőnek minősített szavak aránya. Kiefer (2003b) például kizárná innen a *fenn*-t.

lönféle lexikai erőforrásokban, korpuszokban mekkora eltérések vannak az igekötőnek címkézett szavak mennyiségében. Itt említendő még az egybe- és különírás problémája is. Ha például az MNSZ2-ben bizonyos igemódosítók egybe vannak írva az igével – pl. *póruľjár, békénhagy* – akkor a teljes szó UNKNOWN elemzést kap, míg a Tádéban ezek igekötős igék.

Az igekötők és a pusztán névszói igemódosítók elhatárolásának nehézsége nem elszigetelt eset. A legtöbb nyelvi jelenségnél megfigyelhető, hogy a kategória-határok elmosódottak (Langacker 1987: 14), és bár ilyen a jelenségek természete, az azokat leírni próbáló modellek számára nem praktikus az, hogy elmosódottan tekintsék a kategória-határt. Jellemzőbb az, hogy szükséges és elégséges feltételek mentén elhatárolt kategóriákkal dolgoznak. Ennek a módszertani mellékhatásaként szemet kell hunyniuk az említett tapasztalati tények felett. A következő fejezetben egy olyan modellről lesz szó, amely megpróbálja elkerüľni ezt a mellékhatást, és az elmosódó kategória-határok rugalmas kezelésére törekszik: ez a prototípus-elmélet.

### 3.2. Az igekötők prototípus-elméleti keretben

A prototípus-elmélet a fogalmi tudás kategóriákba rendezésének fontos eszköze a kognitív tudományokban. Az elmélet részletes kidolgozása Eleanor Rosch nevéhez köthető (Rosch 1973), aki első sorban kognitív és antropológiai szempontból foglalkozott az emberek kategorizálási képességével és a prototipikus egyed fogalmával. A nyelvészetben csak később jelent meg a prototípus-elmélet, először a holisztikus kognitív szemantika eszköztárában. Az egyik legalaposabb, magyar nyelvű ismertetését Kiefer (2007: 84-114) adja.

Az elmélet szerint azért tekintünk valamit egy adott kategória tagjának, mert a tulajdonságai elég közel állnak annak a példánynak a tulajdonságaihoz, amelyről úgy véljük, hogy a kategóriát legjobban képviselő példány (ez a prototípus). A kategóriához tartozás fokozat kérdése: némely példányok jobban hasonlítanak a prototípusra, némelyek kevésbé. Ha egy magyar kultúrkörben nevelkedett beszélőnek fel kellene sorolnia az első három madarat, ami eszébe jut, valószínűleg legalább egyet említene ezek közül: veréb, fecske, galamb. Ezek a madár-kategória prototipikus példányai, mert rendelkeznek a „madárság” tipikus jegyeivel (ún. tipikalitási feltételeivel): tudnak repüľni, van szárnyuk, van tollazatuk, van csőrük, stb. A kategóriának kevésbé tipikus példánya a strucc vagy a pingvin, mert repüľni tudás híján mindjárt az első jeggyel nem rendelkeznek. Ugyanakkor nem minden jegy egyformán fontos abból a szempontból, hogy egy adott példányt mennyire

érzünk a kategória tagjának.

Johnston és Leslie (2019: 187-189) a jegyek eltérő fontosságát egy szemléletes példával, a prototipikus kutya esetével mutatja be. A prototipikus kutya ugat, négy lába van, farka van, nyakörvet visel, stb. Ezek nem szükséges és elégséges feltételek, mivel nem minden kutya rendelkezik ezekkel a jegyekkel, láttunk már három lábú kutyát is, nem beszélve az olyanról, amelyiknek nincs nyakörve. Ha találkozunk egy állattal, és meg akarjuk állapítani, hogy az kutya-e vagy sem, akkor ennek az állatnak a látható jegyeit – illetve azok hiányát – használjuk fel a döntésben, mégpedig úgy, hogy sorra vesszük a prototipikus kutya tulajdonságait, egyeseknek nagyobb, másoknak kisebb fontosságot tulajdonítva. Ezen a ponton felmerül a kérdés, hogy vajon hogyan állapítjuk meg a tulajdonságok fontosságát, azaz a jegyek súlyát?

Johnston és Leslie (2019) szerint egy jegy súlyát két tényező határozza meg. Az egyik a jegy elterjedtsége (*prevalence*), vagyis az, hogy a jelenléte mennyire tűnik általánosnak a kategória tagjai között. Így például az ‘ugat’ és ‘van farka’ súlyozott jegyek, mert a legtöbb kutyánál megjelennek. A másik a jegy érvényessége (*cue validity*), vagyis annak a valószínűsége, hogy az adott jeggyel rendelkező elem a kategória tagja. A kiinduló példában a ‘nyakörvet visel’ a kutyák jelentős részére nem igaz, viszont annak elég nagy a valószínűsége, hogy ha egy állat nyakörvet visel, az kutya. Ezért ez a jegy is nagyobb súlyt kaphat. Minél több súlyozott jegye van egy elemnek, annál tipikusabb tagja egy adott kategóriának.

A prototípus-elmélet nemcsak természetes kategóriák, hanem nyelvi jelenségek vizsgálatára is alkalmazható. A magyar igekötő-állomány meghatározása kapcsán elsőként Forgács (2005) tanulmányában jelenik meg a prototípus-elmülethez közeli gondolat, ő a prágai iskola centrum – periféria modelljét használja az igekötők kategorizálásához. Az igekötőség jegyeit elsősorban szemantikai szempontból, valamint a morfológiai szerkesztettség szempontjából közelíti meg.

Kerekes (2011) már a prototípus-elmüleltre hivatkozva vizsgálja a magyar igekötő-állományt. Vizsgál szemantikai és morfológiai jegyeket – ezekhez Forgács (2005) és Jakab (1976) elemzését veszi alapul –, és bevezet újabb, formai jegyeket is, amelyek az egyes igekötő-jelöltek különböző szerkezetekben való viselkedését írják le. Ezeknek a jegyeknek az értékeit kísérletes módszerrel, anyanyelvi beszélők ítéletei alapján állapítja meg.

Én Kerekes (2011) módszerét követem annyiban, hogy jegyeket határozok meg, és empirikusan vizsgálom ezek meglétét az egyes igekötő-jelöltek – a továbbiakban csak igekötők – esetében.

Eltérek tőle abban, hogy a vizsgálatot nem anyanyelvi beszélők tesztelésével, hanem korpuszmé-  
réssel valósítom meg. Az általam vizsgált jegyek mind morfoszintaktikai jellegűek, szemantikai  
jegy az automatikus mérhetőség hiánya miatt nincsen köztük.

Prototipikus igekötőnek a *meg*-et választottam. Ez messze a leggyakoribb igekötő: ha csak a  
finit igével egybeírt előfordulásait számoljuk, akkor az 1,04 milliárd tokenes, módosított MNSZ2-  
ben 3,27 milliószor fordul elő. Az igekötői státuszát illetően a szakirodalmak egyetértenek, a tör-  
téneti adatoltsága alapján a legősibb eleme a kategóriának, és az egyetlen, amelyik szemantikailag  
teljesen kiüresedett. A kutatásnak ezen a pontján még nem lehettem biztos abban, hogy ezt az ada-  
tok is vissza fogják-e igazolni, azaz hogy a *meg* kiválik-e majd a többi igekötő közül, de a felsorolt  
okok, valamint az intuícióm alapján is a legjobb választásnak tűnt. Az igekötőség mérhető jegyeit  
így elsősorban a *meg* tulajdonságai alapján gyűjtöttem össze.

Úgy döntöttem, hogy bevonok a vizsgálatba minden olyan szót, amely a PREVLEX erőforrás-  
ban – ld. 2.2.2. fejezet – igekötőként szerepel, így a kiinduló adathalmazom összesen 235 tagot  
számlált. Minden szó mellé kigyűjtöttem a vizsgált jegyek előfordulásait az MNSZ2 korpuszból  
(ld. 2.2.1. fejezet).<sup>6</sup> A továbbiakban ezeknek a korpuszméréseknek az eredményét mutatom be.

### 3.3. Az igekötőség mérhető jegyei

Ebben a fejezetben az igekötőségnek olyan lehetséges jegyeit vizsgálom, amelyek automatikus  
módszerrel kimérhetők. Először azokról a jegyekről lesz szó, amelyek a *meg* igekötő esetében is  
jelen vannak, ezért tipikusnak tűnnek az igekötők jellemzése szempontjából. Ezek a morfológiai  
produktivitás<sup>7</sup> (azaz hogy az igekötő mennyire vett és vesz részt új szavak alkotásában), a redupli-  
káció (pl. *meg-megállt*), és a deverbális főnévtől való elválás képessége tagadás esetén (pl. *be nem  
vallás*). Kissé határesetként ide sorolható a *-fele/felé* morféma csatolhatósága is (pl. *szontyolodik  
elfele*) – ez a *meg* esetében semmiképp nem tipikus, de előforduló jelenség.

Ezután két olyan jegyet vizsgálok, amelyek a *meg*-nél nem mutathatók ki, de más, közel-  
prototipikus igekötőknél igen. Ezek a fokozhatóság (pl. *összébbhúz*) és a személyragozhatóság

<sup>6</sup> Az egyes jegyek előfordulási gyakoriságát az olyan tagmondatokban vizsgáltam, ahol az igekötő finit igéhez tarto-  
zott, kivéve értelemszerűen a deverbális főnévtől való elválást tagadás esetén.

<sup>7</sup> Erről külön publikáció született – Kalivoda (2019) –, ahol kissé eltérő számadatok szerepelnek, mint ebben a dolgo-  
zatban. Ennek az oka az, hogy itt a PREVLEX-nek egy újabb, javított verzióját használtam fel.

(pl. *rámnéz*).<sup>8</sup> Végül három általános jegyről lesz szó: a szerkesztettségéről, a szótagszámról és a gyakoriságról. Ezek minden szóra érvényesek, mégis érdemes vizsgálni őket, mert az igekötőknél – ahogy látni fogjuk – jellegzetes értékeik vannak.

### 3.3.1. Morfológiai produktivitás

Egy szóalkotási mintát akkor tekinthetünk produktívnak, ha a minta alapján tetszőleges számú, szemantikailag transzparens szó hozható létre egy adott szemantikai tartományban (Kiefer és Ladányi 2000c: 149). A morfológiai produktivitás esetében is – mint minden nyelvi jelenségnél – célszerű rávilágítani a kompetencia és a performancia közti különbségre. A nyelvi rendszer szintjén a produktivitás egy lehetőség, a minden további nélkül létrehozható szóalakok nem biztos, hogy ténylegesen létrejönnek, és még kevésbé valószínű, hogy benne lesznek egy korpuszban. Emiatt több elméleti morfológus (pl. Dressler 2003) nem tartja helyesnek a kompetenciaszintű lehetőség performanciaszintű valószínűségeen alapuló vizsgálatát. Én mégis az utóbbit célozom meg, mivel ezt a módszert tartom a leginkább objektívnek és reprodukálhatónak, az eredmény pedig tendenciák szintjén érdekes lehet a kompetenciát kutatóknak is.

A morfológiai produktivitás kvantitatív meghatározása Baayen (1989, 2009) nevéhez köthető. Három típust különböztet meg: a megvalósult (*realized*), a terjeszkedő (*expanding*) és a lehetséges (*potential*) produktivitást. A következőkben arról lesz szó, hogy pontosan mik ezek a típusok, és hogyan jellemezzük egy-egy igekötőt.

A **megvalósult produktivitás** ( $P_m$ ) annak a mértéke, hogy egy adott affixum a mérés időpontjáig mennyire vett részt a szóalkotásban, tehát a múltbeli és a jelenlegi szerepe jellemezhető ezáltal. Úgy kapjuk meg, hogy az affixumot (itt: igekötőt) tartalmazó lemmák darabszámát elosztjuk a korpuszban (itt: a PREVLEX-ben) található összes lemma darabszámával.

$$P_m = \frac{\text{az affixumot tartalmazó lemmák (típus)}}{\text{a korpusz összes lemmája (típus)}}$$

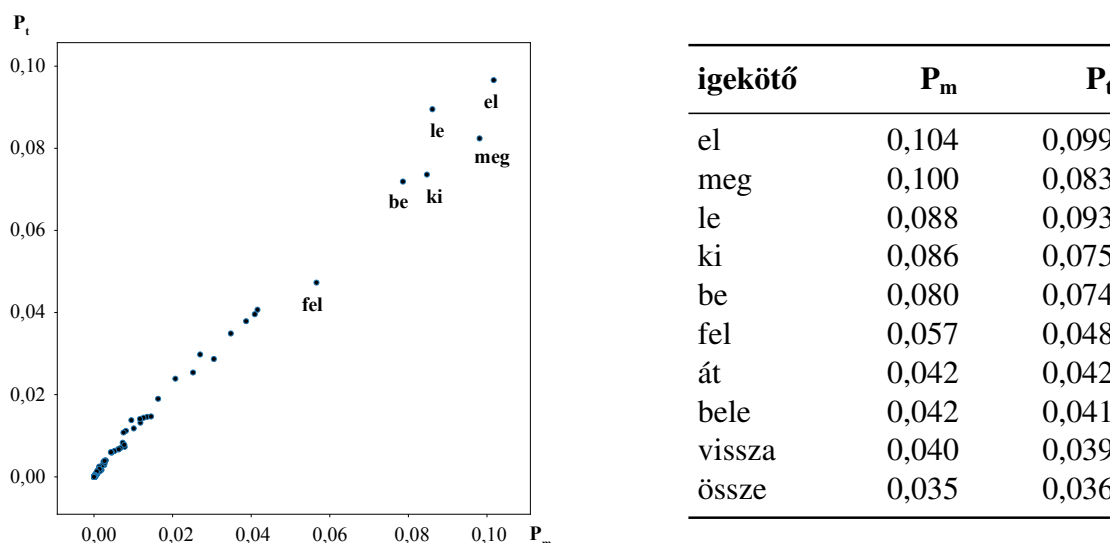
A **terjeszkedő produktivitás** ( $P_t$ ) arról ad jóslatot, hogy az affixumnak várhatóan mekkora szerepe lesz a szóalkotásban a közeljövőben. Ehhez az affixumot tartalmazó hapaxok darabszámát osztjuk el a korpuszban található összes hapax darabszámával. Ez azért is jó mérték, mert a hapaxok

<sup>8</sup> Azért írom ezeket egybe az igével, mert korpuszpéldát idézek. Helyesírási kérdésekre nem térek ki a dolgozatban.

jelentése jellemzően kompozicionális, ezért kevesebb „hamis produktív” találat adódik hozzá az eredményhez, mint a megvalósult produktivitás esetében.

$$P_t = \frac{\text{az affixumot tartalmazó hapaxok}}{\text{a korpusz összes hapaxa}}$$

A PREVLEX igekötőinek megvalósult és terjeszkedő produktivásáról a 2. ábra ad áttekintést. Bal oldalt a két mérték összefüggése síkban ábrázolva látható, az X-tengelyen a  $P_m$ , az Y-tengelyen a  $P_t$  értékeivel. Jobb oldalt a tíz legmagasabb értékű igekötő szerepel.<sup>9</sup>



2. ábra. Az igekötők megvalósult ( $P_m$ ) és terjeszkedő ( $P_t$ ) produktivitása.

Az ősi igekötők (*meg*, *el*, *le*, *ki*, *be*, *fel*) mindkét mérték szerint kiugróan produktívak, bár a *fel* elmarad a többitől. Látható az is, hogy a kétféle produktivitás nagyjából egyenesen arányos. A tendenciától csak a *le* és a *meg* térnek el. A *le* azért is figyelemre méltó, mert a terjeszkedő produktivitása nagyobb, mint a *meg*-é. Ez azt jelenti, hogy várhatóan az igekötős igék alkotásában is egyre nagyobb szerepe lesz.

A Baayen által definiált produktivás-típusok harmadik tagja a **lehetséges produktivás** ( $P_l$ ). Ez az egészen távoli jövőről ad jóslatot: mik azok a most még viszonylag ritkán előforduló affixumok, amelyeknek jó esélye van arra, hogy később sok szó képzésében vegyenek részt? A ter-

<sup>9</sup> Bizonyos igekötők (pl. *fel* – *föl*) egymás alakvariánsaiként is kezelhetők, és a mérés előtt összevonhatók. A kutatás során minden jegy értékét kimértem úgy, hogy ezek az alakvariánsok önálló tételként szerepeltek, és úgy is, hogy összevontam őket (erről bővebben a 3.4. fejezetben lesz szó). Itt csak az összevont verzió eredményeit mutatom be részletesen.

jeszkező produktivitáshoz hasonlóan ez is hapax-alapú mérték, de az eddig látottaktól élesen eltérő eredményt hoz. Úgy kapjuk meg, hogy egy adott affixumhoz tartozó hapaxok tokengyakoriságát elosztjuk az affixumhoz tartozó összes szó tokengyakoriságával.

$$P_i = \frac{\text{az affixumot tartalmazó hapaxok}}{\text{az affixumot tartalmazó összes szó (token)}}$$

A méréshez célszerű gyakorisági küszöböt választani. Minél kisebb tokengyakorisággal osztunk, annál magasabb – és annál kevésbé informatív – lesz a lehetséges produktivitás. A mérést 5-ös és 5000-es küszöbvel (ld. 4. táblázat) végeztem el.

igekötő	token	hapax	P <sub>1</sub>	igekötő	token	hapax	P <sub>1</sub>
égbe	5	3	0,600	agyon	7 588	264	0,035
torkon	9	5	0,556	körbe	11 285	287	0,025
oldalba	18	9	0,500	ide	15 423	296	0,019
porba	8	3	0,380	körül	12 590	226	0,018
fejen	15	5	0,333	hátra	8 256	137	0,017
szarrá	15	5	0,333	utána	8 415	126	0,015
tarkón	10	3	0,300	végig	42 060	619	0,015
földre	25	7	0,280	előre	12 366	165	0,013
ágyba	28	7	0,250	újra	24 230	312	0,013
házhoz	8	2	0,250	egybe	9 831	123	0,013

**4. táblázat.** Lehetséges produktivitás (P<sub>1</sub>), a 10 legmagasabb értékű szó 5-ös küszöb (bal oldalon) és 5000-es küszöb (jobb oldalon) mellett.

A 4. táblázat bal oldalán szereplő szavakat egyetlen nyelvészeti szakirodalom sem sorolná az igekötők közé. Tény viszont, hogy igekötő-szerűbben viselkednek sok más igemódosítónál, és emiatt nem ritka, hogy egybe is íródna az igével. Az előfordulásaik több dokumentumra oszlanak el, tehát nem csak egy ember szóhasználatát látjuk. Többségük egy jól meghatározható szemantikai tartományban mutatja a produktivitás jeleit, például az *oldalba*, *torkon*, *tarkón* támadást kifejező igékkel állnak (pl. *szúr*, *vág*, *rúg*). A *szarrá* (*ázik*, *fagy*, *bombáz*) már absztraktnan értendő – és tulajdonképp a *szét* stilisztikailag jelölt változata.

A 4. táblázat jobb oldalán rangsorolt szavak státusza kevésbé megosztó: a szakirodalomban is felbukkannak igekötőként, bár az egyetértés ezeket illetően sem általános. Könnyen elképzelhető, hogy idővel lényegesen több új szó alkotásában vesznek majd részt. Az itt látható igekötők mind-

egyike kapcsolódhat névszóból képzett igéhez (pl. *végigszambáz, telekommentel, agyonpárnáz, körbekordonoz, idekontárkodik*), ami táptalajt nyújt a kreatív szóalkotásnak.

### 3.3.2. Reduplikáció

A reduplikáció az a morfológiai művelet, amikor a tő egy részének az ismétlése valamilyen nyelvtani vagy szóképzési-szemantikai kategóriát fejez ki (Cser 2017: 12). Az igekötők esetében a reduplikáció teljes, vagyis a tő egészét ismételjük, például *fel-felnézett, vissza-visszatért*.<sup>10</sup>

A legalaposabb munka, amely kizárólag ezzel a jelenséggel foglalkozik, Kiefer (1995) angol nyelvű tanulmánya. A továbbiakban ennek a főbb pontjait foglalom össze, és ahol csak lehet, korpuszmérésekkel támasztom alá vagy kérdőjelezem meg a benne foglaltakat.

Kiefer a bevezetőjében leírja, hogy az igekötő kettőzésének iteratív funkciója van, tehát a cselekvés, történés ismétlődő jellegét fejezi ki. Az iteratív funkció képzése az akcióminőség-képzéshez tartozik, ezért az alapige jelentését nem módosítja. A tanulmány kutatási kérdései közül hármat emelnék ki, amelyek a jelen fejezet szempontjából relevánsak: (1) Mindegyik igekötő reduplikálható? Ha nem, mik ennek a korlátai? (2) A magyarnak van iteratív funkciójú szuffixuma is – ez az -(O)gAt –, felmerül tehát a kérdés, hogy milyen szemantikai különbségek vannak a reduplikált igekötős és a szuffixált igealakok között, már amennyiben vannak különbségek? (3) A reduplikált igekötős igék szintaktikai tulajdonságai eltérnek az „egyszerű” igekötős igékétől? Ha igen, mivel magyarázható ez?

Az első kérdésre – reduplikálható-e minden igekötő? – Kiefer nemleges választ ad. Ezzel egyúttal cáfolja J. Soltész (1959: 176) állítását, miszerint „Önmagával ikerítődve minden igekötő alkothat megismételt igekötőt, akár van még konkrét jelentéstartalma, akár pusztán aspektus-kifejező vagy képzőszerű funkciója van.” Kiefer négy megszorítást fogalmaz meg, amelyek a reduplikáció megjelenését korlátozzák. Az első megszorítás a vizsgált elemek fonológiai súlyára vonatkozik: eszerint két szótagonál hosszabb igekötők nem reduplikálhatók. Ezt az állítását később minden igemódosítóra kiterjeszti. Ez alól egy kivételt találtam, az *előre* elemet. Kiefer ezt nem tekinti igekötőnek, de az igemódosítók közül nem zárja ki. Besorolástól függetlenül tény az, hogy a vizsgált anyagban ez többször is megjelenik reduplikált formában, ld. az (1) példasort. A meta-

<sup>10</sup> Az igekötők reduplikációjára használt egyéb terminusok: megismételt igekötők (J. Soltész 1959), megkettőzött igekötők (Kiefer és Németh 2012), duplikált igekötők (Kerekes 2011).



adatok szerint különböző szerzők használják így. Ettől a kivételtől eltekintve a fonológiai súlyra vonatkozó megszorítás érvényes. A korpuszadatokon az is megfigyelhető, hogy a reduplikáció egy szótagú igeekötők körében gyakoribb jelenség, mint két szótagúaknál.

- (1) a. [...] lovasát a lejtőn **előre-előrerúgta**.  
b. [...] s édesanya is **előre-előrejött** hozzánk, [...]  
c. Kesserű mama **előre-előreszól** a fiatalokhoz, [...]  
d. [...] és minden tanévben egyre-egyre jobban **előre-előrelépnek** ebben is.

Kiefer további megszorításai már nemcsak az igeekötővel kapcsolatosak, hanem a konkrét igétől is függenek. Ilyen a konceptuális megszorítás, amely szerint az intenzitást, „túlzásba vitelt” kifejező igeekötős igeik esetében nem jelenhet meg reduplikáció. Tehát nem működhet például a *\*túl-túlkompenzál* vagy az *\*agyon-agyondicsér*. Az akcióminőséggel, szemantikával kapcsolatos megfigyeléseket nem igazán lehet az MNSZ2-ben ellenőrizni az ehhez szükséges annotáció hiányában. Annyi viszont biztos, hogy a *túl* igeekötő egyáltalán nem jelenik meg reduplikált formában, az *agyon* pedig csak úgy, hogy az igével együtt szó szerinti értelemben (‘megöl’) szerepel, ld. a (2) példát.

- (2) [...] azokat **agyon-agyonverik** bervapumpával, akár rájuk igazolódik valami, akár nem.

Végül Kiefer megfogalmaz két szemantikai megszorítást is. Az egyik azt mondja ki, hogy statikus igeikkel nem lehetséges az igeekötő reduplikációja, nincs tehát olyan alakulat, hogy pl. *\*el-eláll* (az étel), *\*el-elfér* (valami valahol). Kiefer szerint ezek a statikus igeik nem tűrik meg az iteratív értelmű -(O)gAt szuffixumot sem. A második szemantikai megszorítás szerint az olyan állapotigék, amelyek visszafordíthatatlan változást jelentenek, szintén nem engedik az igeekötő reduplikációját. Nem működik tehát az, hogy *\*a szomszéd meg-megöregedett*, mivel ez a történés nem megismételhető egy konkrét alanyra nézve. Ez ugyanakkor nem az igeekötős ige állandó tulajdonsága, hanem a szövegkörnyezettől is függ.

Kiefer a második kutatási kérdése kapcsán – van-e szemantikai különbség a reduplikált igeekötős és az -(O)gAt képzős igealakok között – tisztázza, hogy az -(O)gAt képzőnek két funkciója van, az egyik a cselekvés vagy történés ismétlődő jellegének a kifejezése, a másik a csökkentett intenzitás

kifejezése. Ha az igealakhoz reduplikált igekötő járul, és az -(O)gAt képző is megjelenik rajta, akkor a képző inkább a második funkciójával járul hozzá az ige jelentéséhez, azt érzékeltetve, hogy az esemény „felületesen, csak egy kicsit” történt (pl. *meg-megérintget, fel-felütöget*).

A kiinduló kérdésre azt a választ adja, hogy a két jelenség között van szemantikai különbség, és ez a környezetükben megjelenő időhatározókkal érhető tetten. Megfigyelése szerint a reduplikált igekötős formák nem tűnnek meg olyan időhatározót, amely az esemény rendszeres voltát fejezi ki, állhatnak viszont olyanokkal, amelyek rendszertelen időközökre utalnak. Ezt az állítást megvizsgáltam úgy, hogy a korpuszban a következő mintát kerestem: a *rendszerint, rendszeresen, rendszerességgel, folyamatosan, állandóan, hetente, havonta, naponta* szavak egyike, utána reduplikált igekötős finit igealak. Összesen 45 illeszkedő mondatot találtam – néhányhoz ld. a (3) példasor –, ez a szám egyértelműen cáfolja Kiefer megfigyelését. Nem igaz, hogy a reduplikált igekötős alakok mellett ne állhatna rendszerességre utaló időhatározó, bár ez kétségtelenül nem olyan jellemző, mint az -(O)gAt képző esetében.

- (3) a. [...] az emigrációból is **állandóan bele-beleszólt** az argentin politikába, [...]
- b. Németországban **rendszeresen fel-felröppennek** iparikém-ügyek.
- c. Volt néhány törzshelyem, ahova **rendszeresen vissza-visszatértem** [...]
- d. [...] **folyamatosan fel-felmerül** az ország területi egységének fenntarthatósága.
- e. [...], **naponta bele-belekandikáljon** a sűrűn egymás alá rótt naplósorokba?...

Kiefer utolsó kutatási kérdése a szintaxisához kapcsolódik: Van-e különbség a reduplikált igekötők és az „egyszerű” igekötők mondatbeli eloszlásában? A kérdésre határozott igen a válasz, a két típus szintaktikai viselkedése jelentősen eltér. Preverbálisan a reduplikált igekötő is képes elválni az igétől, de – Kiefer megfigyelése szerint – ekkor csak az *is* szócska ékelődhet az igekötő és az ige közé. A korpuszméréseim ezt annyiban árnyalják, hogy 25 esetből 24-ben valóban az *is* a közbeékelődő elem, de egy alkalommal a *ne* is előfordul, ld. a (4) példasort.

- (4) a. Jó vizsla még **vissza-vissza is sandít** a gazdájára, és vár rá.
- b. Most már **haza-haza is látogatott** Oroszországba.
- c. [...] mert nem tudnám megállni, hogy vadul **fel-fel ne nyihogjak**, [...]

Az ennél is nyilvánvalóbb különbség az, hogy a reduplikált igekötő Kiefer szerint nem lehet hátravetett helyzetben. A „nem lehet”-et a korpuszmérések tükrében úgy szükséges módosítanunk, hogy „nagyon ritkán fordul elő”, az eseteknek mindössze 0,1%-ában. Ez élesen szembenáll azzal, hogy az „egyszerű” igekötők az esetek több mint egyharmadában hátravetettek (erről bővebben ld. a 4.1.2. fejezetet). A posztverbális helyzetet az (5) példaszor szemlélteti.

- (5) a. Fokozódó élvezettel **nézegetek bele-bele** a tévén a közvetített bugyuta színdarabba.  
 b. [...] vajon mifajta szomjúságból **tér vissza-vissza** a „szavalathoz”, vagyis Petőfihez.  
 c. [...] aztán Donáth és Márkus, néha Zimányi Tibor **szólt közbe-közbe**.

A reduplikált igekötők mondatbeli eloszlásához ld. az 5. táblázatot.

pozíció	-2	-1	0	+1
db	25	680	7375	10
%	0,3	8,4	91,2	0,1

**5. táblázat.** A reduplikált igekötő lehetséges pozíciói a finit igéhez képest.

Említésre méltó még az a jelenség, hogy a reduplikált igekötő nemcsak a finit igétől, hanem az infinitívusztól és más igenevektől is elválhat a preverbális tartományban, ld. a (6) példákat.

- (6) a. [...] a magam kurtácska végtelenségét, amely pillanatokra még **el-el** tud **bolondítani**.  
 b. Egymás mellett mentünk, meghitt egyetértésben, s néha még **meg-meg** is **állva**, [...]

Egy későbbi munkájában Kiefer (2003b) a reduplikáció lehetőségét fontos szempontnak tartja az igekötők és a pusztá névszói igemódosítók elkülönítésében. Álláspontja a következő: „Mivel minden háromnál kevesebb szótagú igekötő megkettőzhető (a lexikalizálódott jelentésűeket, valamint az igekötőszerű lexikalizálódott elemeket beleértve), ha valamely igemódosító nem kettőzhető meg (azaz ha vele nem képezhető frekventatív akcióminőség), akkor az nem lehet igekötő.” (Kiefer 2003b: 184). Részben emiatt is kizárja az igekötők közül az *ott*, *itt*, *kint*, *bent*, *fent*, *lent* szavakat – és hozzá kell tennem, hogy a korpuszadatok között valóban nem szerepel reduplikált alakjuk.<sup>11</sup>

<sup>11</sup> Kiefer (2003b: 185) ezeket nemcsak az igekötők, hanem az igemódosítók kategóriájából is kizárja, és lokális jelentésű helyhatározókként elemzi.

Kerekes (2011: 114-116) szintén fontos, kritérium-értékű jegynek tartja a reduplikációt. Bár megjelenhet olyan esetekben is, mint pl. *lassan-lassan végez, egyre-egyre közeledett*, ezek az alakulatok elég távol esnek az igemódosítói kategóriától, attól könnyen elkülöníthetők.

Végül érdekes adat, hogy a korpuszvizsgálat tanúsága szerint mindössze 39 igekötő esetében találunk példát a reduplikációra. Ennek az oka nem annyira az adathiány, mint inkább a már tárgyalt konceptuális és szemantikai megszorítások lehetnek.

### 3.3.3. Elválás deverbális főnévtől tagadás esetén

Igekötős igéből, igenévből képzett főnevek tagadásakor megfigyelhető, hogy a *nem* tagadósó beférkőzhet az igekötő és a képzett szó közé, például *megoldott* → *megoldottság* → *meg nem oldottság*, *visszafizet* → *visszafizetés* → *vissza nem fizetés*. Ezek a deverbális főnevek tehát megőrzik bizonyos igei tulajdonságaikat, és a szerkezetben látható ‘igekötő – tagadósó – igéből képzett főnév’ sorrend megfelel a megszakított tagadás szórendjének (pl. *Ha meg nem teszed, amit mondok, halál fia vagy!*).<sup>12</sup>

Ez a jegy nemcsak az igekötőkre, hanem a legtöbb más igemódosítóra is érvényes (Kerekes 2011: 116). Főleg az olyan szerkezetekre találunk sok példát, amelyekben az igemódosító maga is igei származék, ld. (7a) és (7b), de állandósult szókapcsolatokban sem ritka, ld. (7c) és (7d).<sup>13</sup>

- (7) a. Szerintem a depresszióban az **élni nem akarás** az egy átmeneti állapot.  
b. A vasárnapi nyitva tartás ill. **nyitva nem tartás** mindenkinek jó, [...]  
c. [...] nyereség **rendelkezésre nem állása** bármely okból [...]  
d. [...] mint például: a házasságra lépés, vagy **házasságra nem lépés**.

A korpuszadatok ismeretében mégis úgy látom – Kerekes (2011: 121-122) állításaival egyetértve –, hogy ez egy fontos jegy, mert a szakirodalomban jellemzően igekötőként elismert szavak esetében nagy arányban megjelenik, miközben az olyan szavak, amelyeknél szélsőségesen alacsony értéke van vagy egyáltalán nem jelenik meg, perifériusabb helyzetűnek tűnnek.

<sup>12</sup> Szintaktikai szempontból nagyon érdekes témát jelentenek mind a tagadott deverbális főnevek, mind a tagadó igei szerkezetek. Ezekről részletesebben a 4. fejezetben lesz szó.

<sup>13</sup> A példák a huTenTen12 korpuszból (Jakubíček et al. 2013) származnak.

### 3.3.4. A *-fele/felé* szócska csatolhatósága

Olyan mondatokban, amelyek felszólítók vagy progresszív olvasatúak – azaz épp zajló, folyamatban lévő cselekvést vagy történést fejeznek ki –, az igekötőkhöz a *-fele/felé* morféma kapcsolható (Kerekes 2011: 115), pl. *mennek felfele, gyűródik össze fele*. A *-fele/felé* jelentése ‘valamilyen irányba’, így elsősorban irányjelentésű igekötőkhöz kapcsolódik. Semmiképp nem megkülönböztető jegy, mivel egyéb határozói kifejezésekben is megjelenik, sőt ezekben jellemzőbb. Mégis érdemes foglalkozni vele, mert terjedőben van a prototipikus *meg* igekötő esetében is (pl. *összátok megfele*).

Erről a jelenségről nagyon kevés szó esik a szakirodalomban. Kerekes (2011) említi, mint lehetséges jegyet, de nem tulajdonít neki különösebb jelentőséget. Nádasdy (2005) arról ír, hogy a folyamatos szemlélet kifejezésére már önmagában az igekötő hátravetése is alkalmas, a *-fele/felé* kapcsolása az igekötőhöz csak ráerősít a folyamatosságra. Blogbejegyzések és kommentek formájában viszont több anonim szerzőnél is felbukkan ez a téma.<sup>14</sup> Ezekből az derül ki, hogy a *-fele/felé* kapcsolása az igekötőhöz főleg népnyelvi jelenség – a nyírségi nyelvjárásban a legjellemzőbb –, de a köznyelvben is egyre gyakoribb.

Minden igekötő-jelölt esetében megvizsgáltam ennek a jegynek a jelenlétét a korpuszadatokban, de csak a leggyakoribb esetet vettem figyelembe, a közvetlenül ige utáni pozíciót. Egyedül a *megfele/megfelé* szóalakok esetében döntöttem úgy, hogy egy külön mérés keretében érdemes lehet megvizsgálni a mondatbeli eloszlásukat a finit ígéhez viszonyítva. Ezeknél ugyanis nem merül fel az a lehetőség, hogy „tisztán határozói” szerepük lenne, és igekötői nem. A mérés eredményét a 6. táblázat mutatja be.

pozíció	-1	+1	+2	+3	+4
db	3	57	13	2	1
%	3,9	75	17,1	2,7	1,3

6. táblázat. A *megfele/megfelé* szóalakok pozíciói a finit ígéhez képest.

Láthatjuk, hogy az esetek 96,1%-ában posztverbális az igekötő, vagyis az ige után áll. Ennek az oka, hogy mind a progresszív olvasatú, mind a felszólító mondatok esetében az ‘ige – igekötő’

<sup>14</sup> <https://www.nyest.hu/hirek/a-nap-kommentje>  
[https://itcafe.hu/tema/magyar\\_szavak\\_amiket\\_nem\\_mindenki\\_ismer/hsz\\_1069-1069.html](https://itcafe.hu/tema/magyar_szavak_amiket_nem_mindenki_ismer/hsz_1069-1069.html)

szórend az alapértelmezett. A preverbális, azaz ige előtti helyzet valószínűleg csak azért fordulhat elő, mert a *-fele/felé* szócska már önmagában hordozza a folyamatosságot, ezért a szórend megváltoztatása nem feltétlenül szükséges. A (8) példaszor egy-egy mondattal illusztrál minden pozíciót.

- (8) a. Ami vót a lábán, mikor **megfele halt**, azt a cipőt. -1  
b. **Tisztuljál megfele** szépen. +1  
c. [...], hanem **oldja** azokat **megfelé** legalább generációsan. +2  
d. Rohamosan **tisztul** a közélet **megfele**. +3  
e. **Nézzed** már a dátumot **megfele**, ő írta előbb! +4

Az igekötőhöz kapcsolt *-fele/felé* morféma nemcsak finit igeik, hanem infinitívuszok mellett is megjelenik, ld. a (9) példaszort.

- (9) a. Kezdett már jól **virradni megfele**.  
b. Persze hogy nem tudta **előzni megfele**.  
c. És azt mondják, nem tud **megfele halni**, mer két lelke van.

A vizsgált adatban a *meg*-en kívül két olyan igekötő van, amelyeknek nincsen irányjelentése, mégis „elbírájk” a *-fele/felé* szócskát – ezek a *tönkre* és a *ketté*, ld. a (10) példáit.

- (10) a. [...], az autóban lévő motor meg nyivákol, mert hát ugye ettől **megy tönkrefele**.  
b. Most **hasadjak kettéfele**?

### 3.3.5. Fokozhatóság

Bizonyos, főként határozószói eredetű igekötők középfokban is megjelenhetnek, pl. *összébbhúzó-dik, kívülebb rekedtek, széjjelebb mászik*. Az MNSZ2-ben nem találunk példát arra, hogy a prototipikus *meg* előfordulna középfokú alakban, de tágabb merítést véve, az interneten akad egy-két példa, ld. a (11) példaszort. Az egységes módszertan és az adatok összehasonlíthatósága érdekében ezeket a példákat nem vettem be a méréseim eredményébe, csak érdekességképp tüntettem fel itt.

- (11) a. A lelkem is **megebb nyugodott**, hogy talán tehetek valamit érte.
- b. Tegyenek már valamit, nehogy ez a magyar demokrácia **megdőljön**, mint a pisai ferde torony! Vagy még **megebb**, mint Jerikó falai!

A *meg*-hez hasonlóan ritka ez a jelenség az *el* esetében is, bár erre a vizsgált korpuszadatban valamivel több példát találunk. Kálmán (2013) első sorban ennek az igekötőnek a kapcsán, egy blog-bejegyzésben foglalkozik a témával, amelyet ő fokozás helyett inkább összehasonlításnak nevez, „[...] hiszen arról van szó, hogy két vagy több jelenséget összehasonlítunk abból a szempontból, hogy egy tulajdonsággal milyen fokban rendelkeznek”.

Az *el* igekötő hasonlító alakjai az *elebb*, még ritkábban az *eljebb*. Kálmán azt írja, hogy ezek esetleg tájnyelvi alakok is lehetnek, ezt megerősíteni és cáfolni sem tudjuk, mert nincs olyan felmérés, amelyből kiderülhetne. Fenntartja a lehetőségét annak is, hogy egyszeri szóhasználatról – okkazonalizmusról – van szó, mert korpuszok alapján sem adatolható. Ez a cikk írásának idején még bizonyára így lehetett. Az MNSZ2 itt felhasznált verziójában, amely 2016 augusztus végétől érhető el, már találunk példát az *el* hasonlító alakjaira. Ezek különböző dokumentumokból származnak, ami nem garantálja azt, hogy a dokumentumok szerzői is mind különbözőek, de valószínűsíti. Összesen öt példa van az *elebb/eljebb* alakokra, ld. a (12) példasort. Két esetben a mondat szerzője idézőjelbe is tette az igekötős igét, ami arra utal, hogy ő maga is érzi az alak furcsaságát.

- (12) a. [...], most már ha lesz, akkor se tudok menni -, kicsit **elvadultam**, még **elebb**.
- b. [...], most szakadt el a cérna. Most **szakadt elebb**.
- c. Nem lehet azt mondani, hogy Csucsut „**elebb hanyagoltam**” valamelyest.
- d. [...], jobb, ha Holt Telkek címszó alatt „**elebb temetem**” szegényeket.
- e. [...], sőt a magyar deviza gyengülése számos terméke esetében **eljebb tolta** az árakat.<sup>15</sup>

Kálmán azt is megemlíti, az *át* és a *rá* esetében még kevésbé képzelhető el a fokbeli összehasonlítás, mint az *el*-nél. Ezek nálam sem szerepelnek a korpuszmérés adatai között. Az *át* hasonlító alakjára elvétve lehet internetes példákat találni, ld. a (13) példasort.

<sup>15</sup> A kontextus alapján elképzelhető, hogy itt egyszerűen elgépelésről van szó, *feljebb* → *eljebb*.

- (13) a. [...] mely korlátokon **átlépünk**, természetesen, egyre **átabb**, [...]
- b. **Átalakítottam**, és folyamatosan még **átabb alakítom** a terápiás oldalamat.
- c. [...], mert mi még mindig **átvesszük**. Sőt! Még **átabb**.

Szintén csak érdekes adalékként jegyzem meg, hogy a hasonlító alakok nemcsak finit ige mellett jelenhetnek meg, hanem deverbális főnév, infinitívusz és más igenevek esetében is, ld. (14) példák. Mindössze egy-egy példa van arra, hogy az alábbi igekötők hasonlító alakban állnak, így ezeket még indokolt egyedi szóalkotás eredményének tekinteni.

- (14) a. [...] amihez hozzá veendő még az erdélyiek konzervatívabb volta (a **zokonabb vevés**).
- b. Szóval ez egy nagyon szerteágazó történet, ami még **szertébb fog ágazni**, [...]
- c. Ez nekem persze talány, míg valaki **hozzább értő** nem látja.

Elképzeltetőnek tartom, hogy az igekötők hasonlító alakjainak használata inkább a beszélt nyelvre jellemző. Ott ugyanis gyakoribbnak tűnik az, hogy rövidsége törekszünk, tehát egyszerűen „takarékossági okból” mondjuk például a *még jobban összevesztek* helyett azt, hogy *összébbvesztek*. Mindez csak feltételezés szintjén marad addig, amíg nem vonunk be beszélt nyelvi korpuszt is a vizsgálatba.

### 3.3.6. Személyragozhatóság

Névmási és névutói jellegű igekötőknél figyelhető meg az a jelenség, hogy személyragokat kaphatnak, és sokszor ilyen formájukban is egybeírják őket az igével, különösen akkor, ha az ige rövid, ld. a (15) példasort. Ez a tulajdonság semmiképp nem tipikus jegye az igekötőknek, mivel egyetlen ősi igekötőre sem jellemző. Mégis érdemes foglalkozni vele, mert a korpuszadatok tanúsága szerint a személyragozható igekötők meglehetősen gyakran állnak személyragozott alakokban. Összesen 57 162 olyan eset van, ahol egybeíródnak az igével, és ezek mindegyike UNKNOWN elemzést kap.

- (15) a. Ha **hozzámnyúlsz**, megöllek!
- b. [...] ha MOL kútnál tankolsz, akkor utána mindig **nekedmennek** hátulról.
- c. Verjünk tábort, **ránkfér** a pihenés!
- d. **Utánukvágom**, egyék meg, ha az kell nekik!



Az igekötők személyragozhatósága gyakran tárgyalt téma a szakirodalomban. Élénken vitatott kérdés az, hogy a személyragozott alakok esetében nem történik-e szófajváltás, vagyis igekötő-e még a fenti példákban a *hozzám*, *neked*, *ránk*, *utánuk*, vagy személyes névmásként – az *utánuk* személyragozott névutóként – elemzendő (ld. Elekfi 2001; Rákosi 2014).

A másik, többek által tárgyalt téma az igekötő és a névszói vonzat ragjának rendszeres együttjárása, amely egyes szám harmadik személyű – Rákosi (2014) adatai szerint ritkán többes szám harmadik személyű – vonzat esetén mutatkozik meg (pl. *ránéztem a lányra*, *rájuknéztem a lányokra*). A névmási jellegű igekötőknél ez hangalaki hasonlóságot, egyeztetés-szerű viszonyt jelent a raggal, bővebben ld. Kálmán és Trón (2000). Ilyenek a *bele/belé -bA*, *hozzá -hOz*, *neki -nAk*, *rá/reá -rA*.<sup>16</sup> A névutói jellegű, személyragozható igekötők *-nAk* ragos névszói vonzattal járnak együtt rendszeresen. Ide sorolható az *alá*, *elé*, *főlé*, *közé*, *mellé*, *utána*. Az itt felsorolt igekötők és a névszói vonzatok ragjának rendszeres, együttes előfordulása jól vizsgálható a PREVCONS kutatóeszköz segítségével, amelyről az 5.3. fejezetben lesz szó.

### 3.3.7. Általános jegyek: szerkesztettség, szótagszám és gyakoriság

Ez a fejezet olyan jegyeket mutat be, amelyek minden szónál megfigyelhetők – emiatt úgy tűnhet, hogy túl általánosak, és nem formálható belőlük olyan tipikalitási feltétel, amely segítené az igekötői kategória azonosítását. Valójában viszont fontos lehet az, hogy ezek a jegyek milyen értéket vesznek fel a vizsgált szavak esetében. A *meg* igekötőt alapul véve azt várjuk, hogy az igekötők morfológiaiilag tovább nem bonthatók, rövidek és gyakoriak legyenek.

A **szerkesztettség** azt jelenti, hogy az adott szó morféimákra bontható (Kiefer és Ladányi 2000c: 138). Az ősi igekötőkre – *meg*, *el*, *fel*, *le*, *ki*, *be* – ez nem áll, és a hiányát több szerző kiemelten fontos szempontnak tartja az igekötő-állomány meghatározásában. Az igekötők szerkesztettségét illetően Jakab (1976: 64-66) végezte eddig a legrészletesebb, részben nyelvtörténeti szempontú vizsgálatot. A lehetséges igekötőket két csoportra osztja aszerint, hogy formálisan elemezhetők-e vagy sem, és ezen belül további alcsoportokat határoz meg. A rendszere vázlatosan a következő:

I. leíró nyelvtani szempontból formálisan sem elemezhetők:

1. történetileg morféimavesztéssel redukálódtak (ide tartozik a hat ősi igekötő)

<sup>16</sup> Rákosi (2014) tárgyal még olyan igekötő-szerű elemeket is, amelyek az én adatomból hiányoznak, mert a kiinduló adathalmazban nem szerepeltek az igével egybeírva. Ezek a következők: *érte -ért*, *benne -bAn*, *vele -vAl*.

2. történetileg nem vagy csak fonémavesztéssel redukálódtak, pl. *bele, át, össze, szét*

II. leíró nyelvtani szempontból formálisan elemezhetők, szétválasztható a tő- és a ragmorféma:

1. névszói tő + határozórag, pl. *agy-on, tönk-re, vég-be*
2. névmási tő + határozórag, pl. *ab-ba*
3. határozórag + névszói személyrag, pl. *nek-i*
4. névutó + névszói személyrag, pl. *után-a*
5. határozószó + a közép fok jele, pl. *tová-bb*

Jakab a szerkesztettséget nemcsak morfológiai, hanem szemantikai tulajdonságnak is tekinti. A második csoportba sorolt szavakról ezért azt állapítja meg, hogy alakilag szerkesztettek, jelentés-tanilag viszont nem, mert „[...] már mást jelentenek az egyes szavak a maguk egészében, mint amit a tövek és a ragok jelentésének összegeként jelenteniük kellene. Jelentésük motiváltsága tehát formális.” (Jakab 1976: 66)

Az általam használt módszer nem tesz lehetővé ilyen alapos elemzést. Eszerint szerkesztettnek minősül minden olyan igekötő, amelyet az emMorph automatikus morfológiai elemző (Novák et al. 2016, 2017) szótőre és szuffixumra képes bontani.

A **szótagszám** a szó fonológiai tulajdonsága, és bár kétségtelenül összefügg a szerkesztettséggel – a morfológiailag tovább bontható szavak legnagyobb része több szótagú, a tovább nem bonthatóké egy szótagú –, a két jegy nem ugyanaz. Ezek kapcsolatát a 7. táblázat szemlélteti.

	<b>szerkesztetlen</b>	<b>szerkesztett</b>
<b>egy szótagú</b>	pl. <i>be, meg</i>	pl. <i>jót, részt</i>
<b>több szótagú</b>	pl. <i>össze, vissza</i>	pl. <i>végre, félbe</i>

**7. táblázat.** A morfológiai szerkesztettség és a szótagszám kapcsolata.

Pais (1959: 183) fontosnak, sőt kritérium-értékűnek látja a szerkesztetlenség és az egyszótagúság együttes meglétét. J. Soltész (1959: 15-16) ezt a felfogást túl szigorúnak tartja, és amellett érvel, hogy a hang- és alaktani redukció az igekötők „régiségének” a következménye – tehát grammati-

kalizációs folyamat eredménye –,<sup>17</sup> ezért ne legyen döntő szempont. Egyetértek vele abban, hogy hiba volna például az *agyon*-t kiutasítani az igekötők sorából pusztán azért, mert morfológiailag komplex és két szótagú.

A **gyakoriság** egy szóalak vagy szótípus előfordulásainak száma egy adott korpuszban. A kognitív nyelvészeti irányzatok ennek alapvető jelentőséget tulajdonítanak, így fontos szerepet kap a prototípus-elméletben is. Feltételezhető, hogy egy adott kategóriának a leggyakoribb példányát ítéljük prototipikusnak, és ebből a példányból származtathatóak a tipikalitási feltételek – írja Kiefer (2007: 87) Kálmán László személyes közlését idézve. Én lényegében így jártam el, amikor a *meg* igekötőt jelöltem meg prototípusként.

### 3.4. Igekötő × Jegy mátrixok

Az előző fejezet összesen 11 olyan jegyet mutatott be, amelyek értékei korpuszmérésekkel minden igekötő esetében meghatározhatók voltak. Ezek az értékek elrendezhetők egy mátrixban, ahol a sorokat az igekötők, az oszlopokat pedig a vizsgált jegyek alkotják. Az egyes cellák tartalma az adott jegy értéke az adott igekötő esetében, mintaként ld. a 8. táblázatot.

igekötő	redupl.	-fele	deverb.	fokoz.	szerk.	szótagsz.	gyak.	P <sub>m</sub>	...
...	...	...	...	...	...	...	...	...	...
ki	553	2 067	212	263	0	1	1 225 174	0,083	...
kívül	0	0	1	1	0	2	86	0,001	...
körbe	88	0	0	0	1	2	10 918	0,011	...
körül	3	1	0	0	0	2	11 585	0,008	...
közbe	17	0	1	0	1	2	5 864	0,002	...
...	...	...	...	...	...	...	...	...	...

**8. táblázat.** Részlet az Igekötő × Jegy mátrixból, amely a korpuszban mért jegyértékek abszolút gyakoriságát tartalmazza. A fejlécben szereplő rövidítések feloldása: *redupl.* = reduplikáció, *-fele* = a *-fele/felé* szócska csatolhatósága, *deverb.* = elválás deverbális főnévtől tagadás esetén, *fokoz.* = fokozhatóság, *szerk.* = szerkesztettség, *szótagsz.* = szótagszám, *gyak.* = gyakoriság, *P<sub>m</sub>* = megvalósult produktivítás.

A korpuszadatokból többféle mátrix is létrehozható: a sorok alakulását befolyásolja az, hogy az igekötők alakvariánsairól hogyan döntünk, az oszlopok esetében pedig az a kérdés, hogy biztosan az abszolút értékek mondanak-e el a legtöbbet egy-egy jegyről.

<sup>17</sup> Ennek kapcsán felmerül lehetséges jegyként az igekötő első történeti adatoltsága is, de ilyen korpuszvizsgálatra egyelőre nem vállalkoztam.

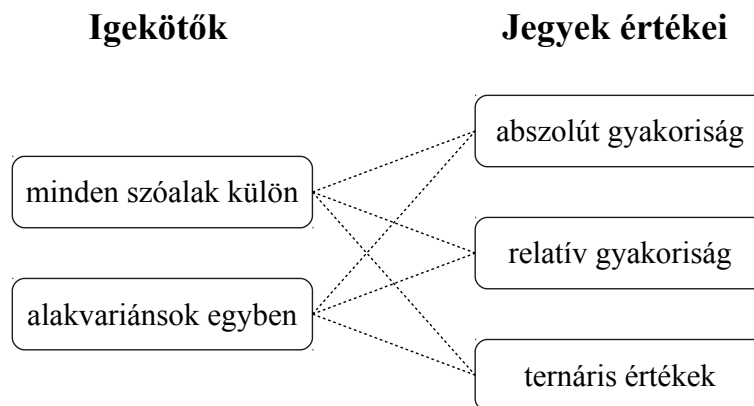
Nézzük először a sorokat: Vegyünk minden szóalakot külön, vagy az alakvariánsokat számoljuk inkább egybe? Alakvariánsnak a következők tekinthetők: *bele – belé, be – bé, fel – föl, felül – fölül, odább – odébb, rá – reá, tele – teli*. Arra jutottam, hogy hasznos lehet ezeket összevonni, mert a párok tagjai között inkább csak stílusbeli, esetleg nyelvjárásbeli eltérés van.

Az oszlopokra vonatkozó kérdés – hogyan érdemes megjeleníteni a jegyek értékét? – már összetettebb. Itt nagyon sok elvi lehetőség adódik, amelyek közül hármat ültettem át a gyakorlatba. Ezek a következők:

1. **Abszolút gyakoriság.** Ez azt mutatja, hogy az adott jegy hány darab szóalaknál valósult meg az adott igekötő esetében, erre hoztam példát a 8. táblázatban. A jegyek értékei ebben az esetben nagyon széles skálán mozognak (a nullától kezdve a milliós nagyságrendig). Ez a legfontosabb abból a szempontból, hogy a legrészletesebb adatot szolgáltatja, és minden más értékadás ebből vezethető le. Ugyanakkor nehezzé teszi az adat elemzését, mert a gyakoriság mindent ural: a milliós nagyságrendben előforduló igekötőknek minden jegyükben esélye van nagyobb értékeket felvenni a csak ezres nagyságrendben előfordulóknál.
2. **Relatív gyakoriság.** Ezt úgy kapjuk meg, hogy a jegynek az adott igekötőnél mért abszolút gyakoriságát elosztjuk a jegy összes előfordulásával. Így minden igekötőnél egy százalékos értéket kapunk, amely azt mutatja, hogy az mekkora részt vesz ki a vizsgált jegy előfordulásaiból. Ez akkor hasznos, ha az adatot nem annyira automatikus módszerekkel, hanem inkább „emberi szemmel” akarjuk vizsgálni, mert minden érték 0-tól 100-ig tartó skálán mozog, ezáltal jól összehasonlíthatók egymással. Automatikus módszerek, különösen a klaszterelemzés esetében – bővebben ld. a 3.6.3. fejezetet – viszont előnytelen ennek a használata, mert az igekötők körében ritkább jegyek (pl. a személyragozhatóság) jellemzően nagyobb értékeket kapnak, mint a sűrűbb jegyek, így súlyozottabban vannak jelen. Ugyanez a probléma akkor is fennállna, ha a jegy adott igekötőnél mért abszolút gyakoriságát az igekötő abszolút gyakoriságával osztanánk le. Ekkor felértékelődnének az olyan esetek, amikor egy nagyon ritka igekötőnél kicsit véletlenszerűen felbukkan egy jegy, például egyszeri performancia-hibából.
3. **Ternáris értékek.** Ekkor az abszolút gyakoriságot mindössze három értékre egyszerűsítjük le: egy jegy nincs (0), minimálisan van jelen (0,5) vagy jelen van (1) az adott igekötő esetében. A 0 érték akkor áll elő, ha a jegy abszolút gyakorisága 0. Ez alól csak maga a

*gyakoriság* jegy a kivétel, ahol Sass (2011: 42) javaslata alapján 5-ös küszöböt alkalmaztam: az igekötőnek legalább ötször kellett előfordulnia ahhoz, hogy ne 0-t kapjon a ternáris mátrixban. 1-es értéket a következő gyakorisági küszöböknél adtam: a produktivitási mértékek esetében 0,0005, a *gyakoriság* jegynél 250 – Sass (2011: 78) alapján –, minden egyéb jegynél 5. A ternáris mátrix 0,5-ös értéke értelemszerűen akkor jelenik meg, ha az adott jegy abszolút gyakorisága nem éri el a szükséges küszöbértéket. Bár a háromértékű jegyek elnagyoltak, leegyszerűsítettek az abszolút és relatív gyakoriságúakhoz képest, mégis nagyon jól használhatónak bizonyulnak akár a kézi, akár a gépi adatelemzésben.

Összegezve tehát hat mátrix hozható létre a sorok és az oszlopok ábrázolásának kombinációiból, ahogy a 3. ábra mutatja.



**3. ábra.** Az igekötők, valamint a jegyek értékeinek ábrázolásai. A vizsgált lehetőségekből összesen hat mátrix állítható elő, a kombinációkat a szaggatott vonalak mutatják.

Ezek az Igekötő  $\times$  Jegy mátrixok nyílt hozzáférésűek, és további kutatások alapjául szolgálhatnak.<sup>18</sup> A soron következő fejezetekben csak azokról írok bővebben, amelyek a korreláció-vizsgálat és a jegyalapú osztályozás szempontjából a leghasznosabbnak bizonyultak.

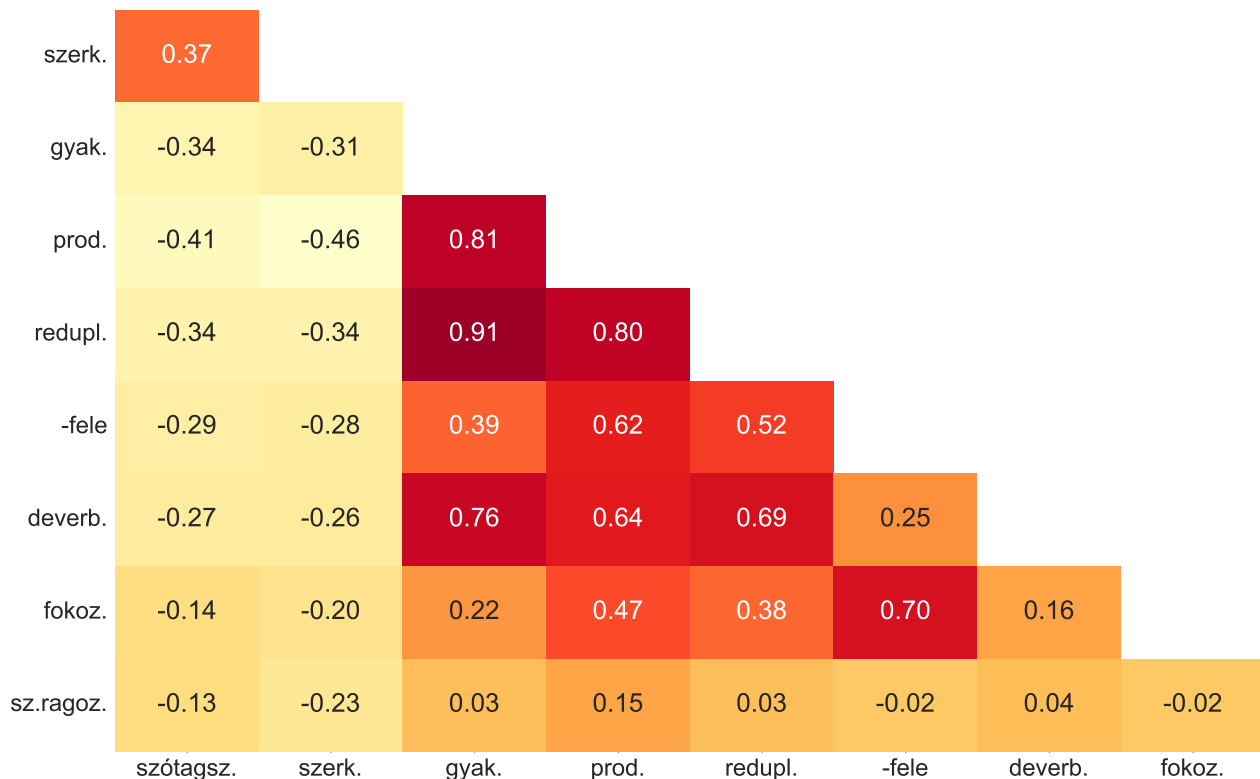
### 3.5. Jegyek közötti korrelációk

Az előző fejezetben ismertetett mátrixok lehetővé teszik azt, hogy minden jegypár közötti korrelációt kiszámoljunk. Ezáltal kimutatható, hogy van-e kapcsolat az egyes jegyek között, és ha igen, ez mennyire szoros. Az eredményül kapott korrelációs együttható ( $r$ ) egy 0 és 1 közötti szám,

<sup>18</sup> <https://github.com/kagnes/prevmatrix>

amely az együttállás szorosságát jelzi: a kapcsolat annál gyengébb, minél közelebb van ez a szám a 0-hoz. Az együtttható előjele a jegyek közti kapcsolat irányát mutatja: negatív előjel esetén minél magasabb az egyik jegy értéke, annál alacsonyabb a másiké, pozitív esetén pedig az egyik jegy magasabb értékeihez általában a másik jegynek is magasabb értékei párosulnak. Ahogy látni fogjuk, a korreláció-számítással statisztikailag igazolhatóvá válik több olyan állítás, amelyet a szakirodalom jellemzően tapasztalati tényként kezelnek, és a bizonyításával nem foglalkoznak.

Elsőként az abszolút gyakoriságok mátrixán végeztem el a vizsgálatot,<sup>19</sup> erről a 4. ábra ad áttekintést. Az eredmények magyarázatát az ábrán látható háromszög felső harmadával kezdem: a szerkesztettség, a szótagszám, a gyakoriság és a produktivitás kapcsolatával.



**4. ábra.** Korrelációk az összevont igekötői alakvariánsokat és abszolút gyakoriságokat tartalmazó mátrixon. A sorokban és az oszlopokban is az igekötőség jegyei láthatók, a cellákban pedig az a korrelációs együtttható, amely az adott sor és adott oszlop jegypárjára érvényes. A cellák színe a korreláció erősségét jelzi: minél nagyobb a korrelációs együtttható, annál vörösebb a cella. A rövidítések feloldása felülről lefelé haladva: *szerk.* = szerkesztettség, *gyak.* = gyakoriság, *prod.* = (megvalósult) produktivitás, *redupl.* = reduplikáció, *-fele* = a *-fele/felé* szócska csatolhatósága, *deverb.* = elválás deverbális főnévtől tagadás esetén, *fokoz.* = fokozhatóság, *sz.ragoz.* = személyragozhatóság, *szótagsz.* = szótagszám.

<sup>19</sup> A relatív gyakoriságokon végzett korreláció-számítás pontosan ugyanezt az eredményt adja, ezért nem tárgyalom külön. Ez ellenőrizhető a <https://github.com/kagnes/prevmatrix> adatai alapján.

A szerkesztettség és a szótagszám között pozitív korreláció figyelhető meg, bár nem túl magas ( $r = 0,37$ ). Ennek okát korábban a 7. táblázatban láthattuk: vannak „kakukktójások”, mint a szerkesztetlen, de több szótagú *össze* vagy a szerkesztett, mégis egy szótagú *részt*, de ezek vannak kisebbségben.

A szótagszám és a szerkesztettség egyaránt negatívan korrelálnak a gyakorisággal (az előbbi esetében  $r = -0,34$ , az utóbbinál  $r = -0,31$ ). A szótagszám és a gyakoriság kapcsolata a Zipf-törvény szerint alakul, amely kimondja, hogy a gyakran használt szavak jellemzően rövidebbek, mint a ritka szavak (Zipf 1932). Ennek az állításnak az általam ismert legszélesebb spektrumú vizsgálata 986 nyelv bevonásával történt, és mindegyik esetében kimutatta a Zipf-törvény érvényességét (Bentz és Ferrer-i Cancho 2016). A szerzők szerint ez olyan robosztus adat, hogy a nyelvek, vagy még általánosabban az információ-feldolgozás és átadás univerzális tulajdonságának mondható.

Az ábrán látható az is, hogy a szerkesztettség és a szótagszám ugyancsak negatívan korrelál a produktivitással ( $r = -0,46$  és  $r = -0,41$ ). Tehát minél produktívabb egy igekötő, annál valószínűbb, hogy rövid és morfológiailag bonthatatlan. Végül nagyon erős pozitív korreláció ( $r = 0,81$ ) figyelhető meg a produktivitás és a gyakoriság között. Bár a két jegy szorosan összefügg, mégsem lehet egymás szinonímájának tekinteni ezeket. Ezt jól példázza az *agyon*, amely a tokengyakoriság szerinti rangsorban a 46., a megvalósult produktivitását tekintve a 20. helyet foglalja el. A gyakoriság és a produktivitás kapcsolata egy 2x2-es mátrixszal írható le, amelyet a 9. táblázat szemléltet.

	produktív	nem produktív
gyakori	pl. <i>meg, el, be</i>	pl. <i>létre, egyet, észre</i>
nem gyakori	pl. <i>pofon, szénné, szarrá</i>	pl. <i>hajba, síkra, csonttá</i>

**9. táblázat.** A produktivitás és a gyakoriság kapcsolata, Pakerys (2017) alapján.

Ezek a korrelációs viszonyok egy jól ismert folyamatot tükröznek, a **grammatikalizációt**. Ennek során egy önálló lexikai elem morfológiai jelenséggé – többnyire affixummá – válik. A folyamatot jellegzetes változások kísérik: a szóalak rövidül és morfémaikat is veszít, mígnem teljesen elveszíti a morfológiai szerkezetét, bonthatatlan lesz. Mindez szemantikai következményekkel is jár: a szó eredeti jelentése gyakran eltűnik vagy más jelentéselemekkel egészül ki. A grammatikalizációról bővebben ld. Kiefer (1996), É. Kiss (2014b) és Cser (2017) munkáit.

Kiefer (1996: 266-267) a szláv nyelvek igekötőit vizsgálva a grammatikalizálódás útját a kö-

vetkezőképp határozza meg: *határozói jelentés* → *határozói és akcióminőségbeli jelentés* → *akcióminőségbeli és aspektuális jelentés* → *aspektuális jelentés*. Hozzáteszi, hogy a magyar igekötők közül szemantikailag csak a *meg* járta be ezt a teljes utat. Morfológiai szempontból egyetlen magyar igekötő sem jutott el a végpontig, mivel elválhatnak az igétől, szintaktikailag önálló szavak.

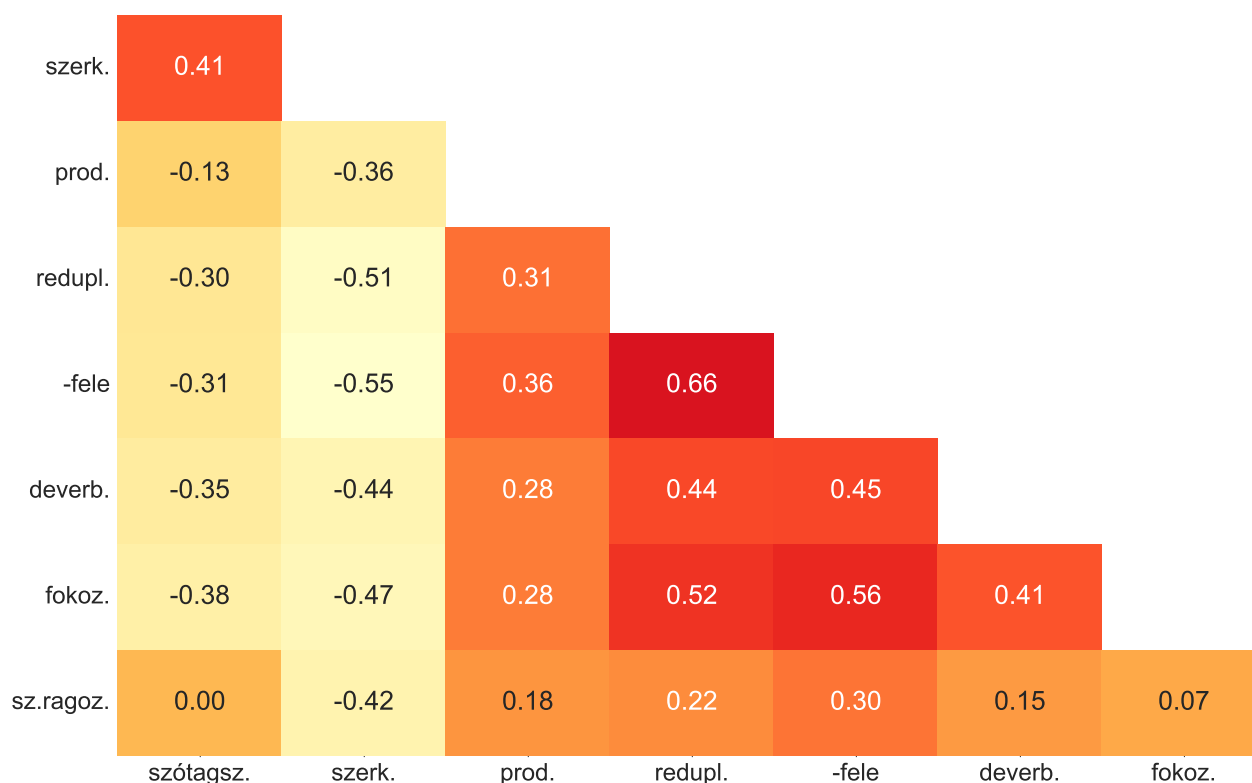
A 4. ábra középső harmadát tekintve a legszembetűnőbb talán a reduplikáció és a gyakoriság közötti kapcsolat. A 0,91 értékű korrelációs együttható olyan közel van az elvileg lehetséges maximumhoz (az 1-hez), hogy kijelenthetjük: a reduplikáció mértéke a gyakoriság függvényében változik. A további összefüggés az, hogy reduplikáció leginkább a grammatikalizálódás folyamatában előrehaladott igekötőkre jellemző – amelyek gyakoriak, produktívak, rövidek és morfológiai szempontból bonthatatlanok.

Az alsó harmadot illetően a fokozhatóság és a *-fele/felé* csatolhatósága közti pozitív korrelációra ( $r = 0,7$ ) érdemes kitérni. Ez leginkább a határozószói jellegből adódik: ennek a két jegynek az értékei az olyan igekötőknél magasak, amelyek irányjelölő funkcióval is bírnak. Ezek a felső harmadban látott jegyekkel lazább kapcsolatban állnak, akár pozitív, akár negatív korrelációkat nézünk. Ez arra utal, hogy a fokozhatóság és a *-fele/felé* csatolhatósága nem lényegi jegyei az igekötőségnek. Különösen igaz ez a személyragozhatóságra, amelynek minden korrelációs együtthatója 0 körül van, vagyis egyetlen más jeggyel sem korrelál.

Az eddigi eredményekkel szemben felmerülhet az a kritika, hogy a bizonyos jegyek között mért erős korreláció valamilyen harmadik változó hatására vezethető vissza, amely mindkét jeggyel szoros kapcsolatban van. Gyanítható, hogy a gyakoriság egy ilyen változó. Ennek a hatását ki tudjuk szűrni úgy, hogy binárisra egyszerűsítjük az adatot: ekkor a *gyakoriság* jegy megszűnik, a többi jegynél pedig csak annyit vizsgálunk, hogy az adott igekötőnél megjelenik-e vagy sem. A kísérlet kedvéért létrehoztam egy ilyen bináris mátrixot, és ennek a jegypárjain is elvégeztem a korreláció-számítást. Az eredményt az 5. ábra szemlélteti.

Azt látjuk, hogy a jegyek közti korrelációk mértéke szinte minden esetben csökkent, de az arányait tekintve nincs lényeges változás. Két nagyobb eltérést látunk a 4. ábrához viszonyítva. Az egyik az, hogy a reduplikáció és a produktivitás közötti korreláció 0,8-ról 0,31-re csökkent. Ennek az az oka, hogy ha binárisra egyszerűsítjük az adatot, akkor az alacsony produktivitás felértékelődik (0,0005-ös értékből 1 lesz), így például az *éhen* elem (*éhenhal*, *éhenveszik*, *éhendöglik*, *éhenpusztul*, *éhengebed*) egyenrangú lesz az olyan igekötőkkel, mint a *meg*, *ki*, *be*, stb. Ugyanez az oka annak is,





**5. ábra.** Korrelációk az összevont igekötői alakvariánsokat és bináris jegyértékeket tartalmazó mátrixon. A sorokban és az oszlopokban is az igekötőség jegyei láthatók, a cellákban pedig az a korrelációs együttható, amely az adott sor és adott oszlop jegypárjára érvényes. A cellák színe a korreláció erősségét jelzi: minél nagyobb a korrelációs együttható, annál vörösebb a cella. A rövidítések feloldása felülről lefelé haladva: *szerk.* = szerkesztettség, *prod.* = (megvalósult) produktivitás, *redupl.* = reduplikáció, *-fele* = a *-fele/felé* szócska csatolhatósága, *deverb.* = elválás deverbális főnévtől tagadás esetén, *fokoz.* = fokozhatóság, *sz.ragoz.* = személyragozhatóság, *szótagsz.* = szótagszám.

hogy a személyragozhatóság és a *-fele/felé* csatolhatósága között enyhe pozitív korreláció jelenik meg ( $r = 0,3$ ). Ebben az esetben az olyan ritka alakok értékelődnek fel, mint a *hozzáfele*, *ráfele*.

Az abszolút gyakoriságokon és a bináris értékeken mért korrelációk összehasonlítása alapján megállapíthatunk két dolgot. Először is azt, hogy a jegyek közti alapvető összefüggések nem változnak attól, ha a gyakoriságra vonatkozó információt elhagyjuk. Ennek viszont nem az az oka, hogy a gyakoriságnak nincs jelentősége. Épp ellenkezőleg: történetileg a gyakoriság olyan szorosan együtt jár más jegyekkel – a grammatikalizációs folyamat révén –, hogy akkor is megjelenik az igekötőség jegyeiben, ha expliciten, külön jegyként nem vesszük figyelembe. Ebből következően megállapíthatjuk azt is, hogy ha az igekötőket a jegyértékeik alapján akarjuk osztályozni, akkor lényegében hasonló eredményt várhatunk az osztályozástól, akár a nagyon részletes abszolút gyakoriságú, akár a végletekig egyszerűsített ternáris vagy bináris adatot választjuk kiindulópontként.

### 3.6. Az igeekötők jegyalapú osztályozása

Mielőtt elkezdenénk az igeekötők jegyalapú osztályozásával foglalkozni, érdemes elgondolkozni egy alapvető kérdésen. A kérdés pedig így szól: Egyáltalán szükség van osztályozásra?

A prototípus-elmélet központi gondolata mellett kitartva úgy gondolom, hogy az osztályozás – vagyis a kategóriahatárok behúzása – nincs összhangban a vizsgált jelenség természetével. Csak gyakorlati szempontból tartom indokoltnak a kategorizálást: jelen esetben azért van rá szükség, hogy megkönnyítse az egyes problémakörök tárgyalását a dolgozatban, különösen a következő, mondattani tárgyú fejezetben.

Látni kell azt is, hogy nem létezik olyan objektív mérce, amelyhez hasonlítva meg tudnánk állapítani az osztályozás eredményének a jóságát. Így ennek a fejezetnek nem lehet célja az objektív, hanem csak a minél kevésbé szubjektív kategorizálás.

Először arról lesz szó, hogy hogyan súlyozhatók a különféle jegyek. Ezután négy osztályozási módszert mutatok be, a legegyszerűbb matematikai módszertől a bonyolultabbak felé haladva. Ezeket végül öt lényeges szempont alapján hasonlítom össze, és választok közülük – hangsúlyozva, hogy a megállapított kategóriák között nincs éles határ.

#### 3.6.1. A jegyek súlyozásának kérdései

A prototípus-elmélet bemutatásánál – ld. 3.2. fejezet – szó volt arról, hogy a kategóriába tartozáshoz nem minden jegy egyformán fontos. Az igeekötőség esetében intuitívan is sejteni lehetett például azt, hogy a morfológiai produktivitás lényegesebb, tipikusabb tulajdonsága lehet az igeekötőknek, mint a személyragozhatóság, és ezt a jegypárok közötti korrelációk vizsgálata is alátámasztotta. Ezért indokoltnak tűnik, hogy az osztályozás során a fontosabb jegyek előnyt élvezzenek másokkal szemben úgy, hogy nagyobb súlyt kapnak.

A jegyek súlyozásának nincs egyezményesen elfogadott módszere, ezért saját módszert kellett kialakítanom. Ezt úgy valósítottam meg, hogy a Johnston és Leslie (2019) által bemutatott, prototipikus kutyát leíró példa mintájára minden jegyet értékeltem aszerint, hogy elterjednek és/vagy érvényesnek tekinthető-e. Az olyan jegyek esetében, amelyek nem jelennek meg minden szónál, működött a párhuzam (ld. a 10. táblázatot), de az általános jegyek – szótagszám, szerkesztettség, gyakoriság – nem tűntek beilleszthetőnek ebbe a rendszerbe. Mivel ezek nemcsak az igeekötőknél,

hanem minden egyes szónál megfigyelhetők és mérhetők, Johnston-ék példája nyomán azt kellene mondanunk, hogy elterjedtek, de nem érvényesek. Ez a megoldás mégsem tűnik helyénvalónak, hiszen a korrelációknál egyértelművé vált, hogy ez a három jegy erősen összefügg egymással, és bár minden szónál megfigyelhetők, az igekötőknél jellegzetes értékeik vannak. Természetes kategóriákkal is párhuzamba vonható ez a gondolat: például mérete minden állatnak van, mégis jellemzően alacsony értéket tulajdonítunk ennek, ha lepkéről van szó, és jellemzően magasat, ha elefántról.

vizsgált jegy	elterjedt	érvényes	indoklás
a háromféle produktivitás	✓	✗	Elterjedt: minden igekötőnél megjelenik. Nem érvényes: más affixum-szerű elemekre is fennáll.
reduplikáció	✓	✓	Elterjedt: megjelenik minden olyan elemnél, amelyet a szakirodalmak többsége igekötőnek címkéz. Érvényes: az igekötőkön kívül csak elszigetelt esetekben jelenik meg (pl. <i>olykor-olykor</i> , <i>már-már</i> , <i>vagy-vagy</i> ).
elválás deverbális főnévtől tagadás esetén	✓	✓	Elterjedt: megjelenik minden olyan elemnél, amelyet a szakirodalmak többsége igekötőnek címkéz. Érvényes: az igekötőkön kívül más lexikai elemekre nem vagy alig jellemző.
-fele/felé	✗	✗	Nem elterjedt: irányjelentéses igekötőknél működik, ezen kívül a <i>meg</i> , <i>ketté</i> , <i>tönkre</i> esetében ritkán. Nem érvényes: tér- és időbeli irányt jelölő határozószóknál általános (pl. <i>mindenfele</i> , <i>másfele</i> , <i>estefele</i> ).
fokozhatóság	✗	✗	Nem elterjedt: leginkább csak irányjelentéses igekötőknél működik. Nem érvényes: melléknevek, határozószók körében sokkal általánosabb.
személyragozhatóság	✗	✗	Nem elterjedt: csak néhány igekötőre jellemző. Nem érvényes: névmások és névutók körében sokkal általánosabb.

**10. táblázat.** A jegyek súlyozása aszerint, hogy elterjedtek és/vagy érvényesek-e (ezekről bővebben ld. a 3.2. fejezetet). A ✓ igenlő, a ✗ nemleges választ jelent.

A 10. táblázatban feltüntetett jegyek súlyozása úgy valósult meg, hogy ahol egy ✓ szerepel, ott az adott jegy oszlopát megdupláztam a mátrixban, ahol kettő, ott megtripláztam. Az általános jegyek esetében – azok jellegzetes értékei, korrelációja és a szakirodalmak többségében is jelzett szerepe alapján – a jegy háromszoros súlyt kapott, tehát két ✓-nak megfelelő.

### 3.6.2. Egyszerű matematikai műveletek

A legegyszerűbb osztályozó módszer a következő lépésekből áll: (1) minden igekötő esetében súlyozzuk a jegyeket – az előző fejezetben bemutatott eljárással –, (2) összeadjuk a jegyértékeket, (3) az igekötőket csökkenő sorrendbe rendezzük a jegyértékeik összege alapján. A várakozásunk az, hogy a rangsor elejére kerülnek a legtipikusabb igekötők, a végére pedig a legkevésbé tipikusak. A módszer eredményeként tehát automatikusan kialakul egy rangsor, kategóriák viszont nem. A kategóriahatárokat úgy lehet megállapítani, hogy ahol két egymást követő igekötő jegyérték-összege között ugrásszerű az eltérés – nagyobb az átlagnál –, ott kategóriahatárt feltételezhetünk.

Ez a módszer a ternáris mátrixon adja a legjobb eredményt.<sup>20</sup> Abszolút és relatív gyakoriságok esetében ugyanis az osztályozás eredménye nem sokban különbözik attól, mintha egyszerűen a gyakoriságuk alapján rendeznénk sorba az igekötőket. Ez utóbbi nem volna szerencsés, mert úgy tűnne, hogy például az *egyet* szó igekötő-szerűbb, mint a *zokon*, pedig mindkettő csak egy-egy igével áll (*egyetért*, *zokonvesz*), és mindössze a gyakoriságukban van komoly eltérés az *egyet* javára. A ternáris mátrixon végzett osztályozás eredményét a 11. táblázat szemlélteti.

legmagasabb összeg		közepes összeg (~11,5)		legalacsonyabb összeg	
be	23,0	agyon	12,0	hadat	1,5
fel	23,0	egybe	12,0	hatályba	1,5
ki	23,0	helyre	12,0	küszöbön	1,5
le	23,0	kívül	12,0	kútba	1,5
el	22,5	közre	12,0	részre	1,5
alá	22,0	szembe	12,0	sárba	1,5
meg	22,0	újra	12,0	semmibe	1,5
át	21,5	alább	11,5	valóra	1,5
oda	21,5	kinn	11,5	vérig	1,5
össze	21,5	viszont	11,5	érvénybe	0,0

**11. táblázat.** Első módszer: Osztályozás a jegyértékek összege alapján. Az első oszlopban az a tíz igekötő látható, amely a legmagasabb összeget érte el, a második oszlopban a középmezőnyből szerepel tíz igekötő, az utolsó oszlopba a legalacsonyabb összegekkel bíró tíz elem került.

A 11. táblázat azt mutatja, hogy a módszer eredménye lényegében megfelel a várakozásnak: a leg-

<sup>20</sup> A ternáris mátrix esetében először szükséges az, hogy a más jegyekkel negatívan korreláló jegynek az inverzét vegyük. Tehát szótagszám esetén az 1 szótagú igekötő kap 1, a 2 szótagú 0,5, a 3 szótagú 0 értéket – mivel a rövidség a tipikus igekötői tulajdonság. Ugyanígy a szerkesztettség esetében akkor kell 1-es érték, ha az igekötő szerkesztetlen, és akkor kell 0-s, ha szerkesztett.

magasabb összegű szavakat a szakirodalmak kivétel nélkül igekötőnek tekintik, a legalacsonyabb összegűeket pedig soha – ezek csak az automatikusan előállított igei vonzatkerettárakban vannak igekötőként felcímkezve. Az én intuícióm szerint még fontos lenne, hogy a hat ősi igekötő – *meg, el, ki, be, fel, le* – szorosan együtt álljanak, ez viszont nem teljesül, mert az *alá* beférkőzik közéjük. Ez nem a módszer számlájára írandó, sokkal inkább a jegyek súlyozásával függ össze. Amennyiben negatív súlyt adunk a se nem elterjedt, se nem érvényes jegyeknek – nem pedig érintetlenül hagyjuk őket, mint ahogy én jártam el –, akkor az *alá* hátrébb sorolódik, és az ősi igekötők egysége megmarad.

Második lehetőségként továbbgondoltam azt a módszert, amelyet Kerekes (2011) is alkalmazott. Ennek az a lényege, hogy minden jegyet halmazként értelmezünk, és ha egy igekötő rendelkezik egy adott jeggyel, akkor megjelenik az adott halmaz elemeként. Minél több halmaz metszetében van egy igekötő, annál tipikusabb.

legtöbb halmazban		a halmazok kb. felében		egy halmazban sem	
be <sub>1</sub>	10	fölé <sub>5</sub>	5	hadat	-
fel <sub>1</sub>	10	agyon <sub>6</sub>	4	hatályba	-
ki <sub>1</sub>	10	egybe <sub>6</sub>	4	küszöbön	-
le <sub>1</sub>	10	félbe <sub>6</sub>	4	kútba	-
el <sub>2</sub>	9	helyre <sub>6</sub>	4	részre	-
meg <sub>2</sub>	9	ketté <sub>6</sub>	4	semmibe	-
oda <sub>3</sub>	9	közre <sub>6</sub>	4	sárba	-
vissza <sub>3</sub>	9	szembe <sub>6</sub>	4	valóra	-
össze <sub>3</sub>	9	tönkre <sub>6</sub>	4	vérig	-
alá <sub>4</sub>	9	újra <sub>6</sub>	4	érvénybe	-

**12. táblázat.** Második módszer: Osztályozás halmazok alapján. Az első oszlopban az a tíz igekötő látható, amelyek a legtöbb halmaznak elemei – mellettük látható, hogy pontosan hány halmazról van szó –, a második oszlopban a középmezőnyből szerepel tíz igekötő, az utolsó oszlopba tíz szó került azok közül, amelyek nem elemei egyik halmaznak sem. Ha az igekötők mellett azonos index szerepel, akkor azok az igekötők ugyanabban a metszetben vannak.

A halmazba tartozás igen–nem típusú kérdés, ezért bináris adatot használtam hozzá. Ezt úgy állítottam elő, hogy a ternáris mátrix 0,5-ös értékeit 0-ra fordítottam át, tehát ‘minimálisan van jelen’-ből ‘nincs’ lett, hogy a nagyon ritka előfordulások ne torzítsák az eredményt. Ennél a módszernél is kialakul egy rangsor aszerint, hogy az adott igekötő hány halmaznak az eleme, valamint automatikus kategóriák is előállnak: az egy metszeten belüli elemek értelemszerűen egy kategóriába tartoznak.

Ebben az esetben összesen 44 metszetet kapunk.

A 12. táblázat azt mutatja, hogy az így kapott eredmény alapjaiban hasonlít az előzőre. A legkevésbé tipikus elemek ugyanazok, mint az összegző módszernél. A középmezőnyben már vannak eltérések – a szakirodalmakban is jellemzően ezeknél a szavaknál kezdődik a bizonytalanság. A legtipikusabbnak besorolt elemeknél az a legszembetűnőbb eltérés az összegző módszerhez képest, hogy itt a hat ősi igekötő egysége megmaradt, sőt az *el* és a *meg* külön csoportot alkotnak.

### 3.6.3. Klaszteranalízis

Ebben a fejezetben olyan módszerekre térek át, amelyek jóval bonyolultabb matematikára épülnek, mint az eddigiek: ezek a klaszterező eljárások. A klaszterezés egy gyűjtőfogalom: olyan eljárás, amely adatpontokat csoportosít kizárólag olyan információ alapján, amely ezekből kinyerhető. A célja az, hogy az egy csoportba – klaszterbe – kerülő adatpontok egymáshoz hasonlóak legyenek, és más klaszterek adatpontjaitól különbözzenek (Tan et al. 2018: 528). A jegyeket itt dimenzióknak nevezzük, és az adatpontok hasonlósága vagy különbsége ezeknek a dimenzióknak a mentén határozható meg.

A két egyszerű matematikai módszerrel ellentétben a klaszteranalízisnél használhatóvá válnak az abszolút gyakoriságokat tartalmazó mátrixok. Viszont ahogy láttuk, ezek az adatok Zipf-eloszlást mutatnak, a klaszterező eljárások pedig eredményesebbek, ha normál-eloszlású adat a bemenetük. Ekkor az adatpontok értéke kisebb skálán mozog, nincsenek nagyon kiugró értékek. Számos olyan skálázó eljárás készült, amellyel normál-eloszlásúvá alakítható az adat.<sup>21</sup> Az Igeköti  $\times$  Jegy mátrixok esetében a Yeo–Johnson transzformáció bizonyult a legjobbnak, az eljárás matematikai hátterét ld. Yeo és Johnson (2000) tanulmányában. Először tehát így transzformáltam az abszolút gyakoriságú adatot, és csak ezután következett a klaszterezés.<sup>22</sup> A ternáris adaton is lefuttattam a klaszterezéseket (itt kicsi skálán mozognak az értékek, így nincs szükség transzformációra). Mindkét adattípusból kapott eredményről beszámolok majd.

Elsőként a k-közép (k-means) klaszterezést próbáltam ki (Pedregosa et al. 2011). Ennek az az előnye, hogy minden elemet csoportosít, ami a klaszterező eljárások esetében nem evidens. A hátránya viszont az, hogy előre meg kell határozni a klaszterek számát, bár a legoptimálisabb

<sup>21</sup> Ezek jól áttekinthetők a `scikit-learn` (Pedregosa et al. 2011) weboldalán:

[https://scikit-learn.org/stable/auto\\_examples/preprocessing/plot\\_all\\_scaling.html](https://scikit-learn.org/stable/auto_examples/preprocessing/plot_all_scaling.html)

<sup>22</sup> A transzformált mátrixok is elérhetők a <https://github.com/kagnes/prevmatrix> linken.

klaszterszám megállapítására vannak jól bevált módszerek.<sup>23</sup>

A klaszterezés esetében csoportosítás történik, olyan sorbarendezés viszont nem, mint amit az egyszerű módszereknél láttunk. Ezért az eredményről úgy tudok mintát közölni, hogy választok egy-egy elemet abból a három mezőnyből, amely a két egyszerű módszernél elkülönült (legnagyobb – közepes – legalacsonyabb értékűek), és megmutatom, hogy a választott elem melyik más elemekkel került egy klaszterbe. A k-közép klaszterezés eredményét a 13. táblázat szemlélteti.

<b>ternáris értékek</b>	
<b>meg</b>	alá be bele el elő fel fenn haza hozzá ide ki le meg neki oda rá szét vissza át össze
<b>agyon</b>	agyon egybe előre fejbe félbe félre helyre hátra keresztbe keresztül ketté kölcsön körbe közbe közre szembe tovább tönkre végig újjá újra
<b>valóra</b>	[...] semmibe sorban szemben szemre számon számot szárba színre szóba szót szóvá sárba síkra testet testre tetten tisztán tollba trónra táncra törbe utol valóra végbe véghez végre vérig zokon
<b>normalizált abszolút gyakoriság</b>	
<b>meg</b>	alá be bele el elő előre fel haza hozzá ide ki le meg oda rá szét túl vissza végig át össze
<b>agyon</b>	agyon alul alább benn elé felül félre fölé hanyatt helyre hátra kinn körbe körül külön mellé neki nyilván odébb ott rajta szerte széjjel tele tova utol utána által
<b>valóra</b>	[...] semmibe sorba sorban sorra szemben szemet szemre számba számon számot szárba színre szóba szót szóvá szörnyet sárba síkra talpra testet testre tisztán tollba trónra táncra tőkön tönkre törbe valóra véget véghez vérig zokon zsebre

**13. táblázat.** Harmadik módszer: Osztályozás k-közép klaszterezéssel. A felső sávban a ternáris, az alsó sávban a normalizált abszolút gyakoriságú jegyértékeken végzett klaszterezésből látható egy kis részlet. A klaszteren belül az elemek betűrendben követik egymást, tehát a sorrendnek elvi jelentősége nincs. A *valóra*-féle klaszternél a felsorolás nem teljes.

A k-közép klaszterezés esetében azt látjuk, hogy a csoportosítás eléggé elnagyolt: a vizsgált elemek összesen öt kategóriába sorolódnak, és a hat ősi igekötő nem különül el olyan világosan, mint például a halmazos módszernél. A ternáris értékek és a normalizált abszolút gyakoriság között az *agyon*-féle klaszterben látszanak leginkább különbségek, vagyis az egyszerű módszerek szerinti skála középső sávjában.

<sup>23</sup> Én a „könyökmódszert” használtam, erről bővebben ld.: <https://www.geeksforgeeks.org/elbow-method-for-optimal-value-of-k-in-kmeans/>

Ezután a HDBScan klaszterezéssel kísérleteztem (McInnes et al. 2017). Ennek nagy előnye a k-középhez képest az, hogy az adat alapján magától alakítja ki a klaszterek számát. A hátránya viszont az, hogy nem klaszterez minden elemet: ha egy adatpontnak egy vagy több dimenzióban kiugró értékei vannak (outlier), akkor azt kiszórja. Ez történik például az *agyon*-nal, ld. a 14. táblázatot. A kiszórt elemek nem alkothatnak egy külön csoportot – csak annyi a közös bennük, hogy valamely jegyeik alapján nagyon eltérnek az átlagtól.

<b>ternáris értékek</b>	
<b>meg</b>	alá be bele el fel hozzá ide ki le meg oda rá szét vissza át össze
<b>agyon</b>	agyon helyre körbe
<b>valóra</b>	hatályba küszöbön semmibe valóra
<b>normalizált abszolút gyakoriság</b>	
<b>meg</b>	be el fel ki le meg szét át
<b>agyon</b>	-
<b>valóra</b>	életben életre érvénybe hatályba igazat keresztbe kétségbe küszöbön lényegre oldalba piacra semmibe témába tükörbe valóra világgá világra zavarba

**14. táblázat.** Negyedik módszer: Osztályozás HDBScan klaszterezéssel. A felső sávban a ternáris, az alsó sávban a normalizált abszolút gyakoriságú jegyértékeken végzett klaszterezésből látható egy kis részlet. A klaszteren belül az elemek betűrendben követik egymást, tehát a sorrendnek elvi jelentősége nincs.

### 3.6.4. A módszerek közötti választás

Az előző két fejezetben négy módszert láttunk arra, hogy hogyan határozhatjuk meg az igekötők állományát az Igekötő  $\times$  Jegy mátrixok alapján (és ez természetesen nem merítette ki az összes lehetőséget). Zárásképp érdemes sorra venni néhány olyan szempontot, amelyek mentén összehasonlíthatók ezek a módszerek. A szempontokat a 15. táblázatban tekintem át.

Az összehasonlításból világossá válik, hogy a négy módszer egyike sem felel meg minden szempontnak. Ötből két szempont szerint gyengén vagy gyengébben teljesítenek. Innentől kezdve csak azon múlik a választás, hogy melyik szempontokat tartjuk feláldozhatóbbnak a többinél. Ha például ragaszkodunk ahhoz, hogy a normalizált abszolút gyakoriságot használjuk fel a ternárisal vagy a binárisal szemben, akkor valamelyik klaszteranalízist célszerű választani. Ha inkább azt



szempont	összegző	halmazos	k-közép	HDBScan
minden elem kategorizált	✓	✓	✓	✗
a kategóriák száma automatikusan alakul ki	✗	✓	✗	✓
a tipikalitási skála automatikusan alakul ki	✓	✓	✗	✗
a jegyek súlyozhatók	✓	✗	✓	✓
normalizált abszolút gyakoriság használható	✗	✗	✓	✓

**15. táblázat.** A négy vizsgált módszer áttekintése öt szempont alapján. A ✓ azt jelenti, hogy a módszer megfelel az adott szempontnak, a ✗ pedig azt, hogy nem felel meg, vagy csak nagyon erőltetetten feleltethető meg neki.

tartjuk fontosnak, hogy tipikalitási skálát kapjunk, tehát ne csak csoportok látszódnak, hanem a csoportok rangsora is, akkor az egyszerű matematikai műveleteket érdemes mérlegelnünk. Végül, mivel nincs mihez mérni az eredményül kapott osztályozás jóságát, a döntésre a saját intuíciónk is hatással lesz.

A kutatásom célja szempontjából elsődlegesnek tartottam a tipikalitási skálát – ez a prototípus-elmélet egyik legfontosabb fogalma –, ezen belül lényegesnek tartottam a hat ősi igekötő egységét. Így tehát a halmazos módszer mellett döntöttem. A 44 metszetet – tkp. kategóriát – 4 nagyobb kategóriává egyszerűsítettem. Ezek a következők:<sup>24</sup>

**I. Prototipikus:** *be fel ki le el meg*

**II. Centrális:** *oda vissza össze alá szét át rá bele hozzá ide haza neki elé hátra elő túl széjjel mellé utána félre előre*

**III. Félperiférikus:** *alul felül körül külön odébb szerte tele tova által körbe közbe végig tovább benn ott fölé agyon egybe fejbe félbe helyre keresztbe keresztül ketté közre szembe tönkre újjá újra*

**IV. Periférikus:** *fönn közé bent fenn alább helyt kinn együtt viszont rajta helyben kölcsön rendbe szörnyet véget éhen hátra hanyatt kívül jól ellen nyilván utol jól részt egyet végre ajtót arcul békén csúcsra csődbe célba falra fejen férjhez földhöz földre fülön füstbe főbe hadba harc-ba hasba hasra hegyet helybe házhoz igazat irányt kezét kézre lábra lángba lázba lényegre mellbe mennybe nagyot nagyra nyakon oldalba padlót partra piacra pofon pofán porba rendre ringbe rosszul seggbe seggre sorba sorra szarrá szemem szemet szemmel szájon számba*

<sup>24</sup> A kategóriákon belül is megőriztem az elemek halmazos módszerrel előállított sorrendjét.

*szárnyra szénné talpon talpra tarkón torkon témába térdre tökön törvényt tükörbe tűzbe utat  
világgá világra vízre zavarba zsebre ágyba égbe életben életre észhez útba útra abba cserben  
ellent jóvá karban kétségbe közben közzé létre szemre szóba végbe véghez észre zokon szót*

Összesen 50 olyan lexikai elem van, amely az igekötőségnek egyetlen vizsgált jegyével sem rendelkezik, vagyis egy halmaznak sem volt eleme. Ezekkel a dolgozat további részében nem foglalkozom, további korpuszmérésekben nem vettem őket figyelembe.<sup>25</sup>

A korpuszadatok alapján a prototipikus igekötők főbb jellemzőit így határozhatjuk meg: (1) Nagyon gyakoriak: az 1,04 milliárd tokenes, módosított MNSZ2-ben a *meg* adja a tokenek 3%-át, a *le* – amely a legkevésbé gyakori a prototipikusak közül – a 0,7%-át, és ebbe még csak a finit igével egybeírt eseteiket számoltuk bele. (2) Szerkesztetlenek, azaz morfológiailag tovább nem bonthatók. (3) A fonológiai kopás következtében egyszótagúak. (4) A morfológiai produktivitásuk kimagasló: a terjeszkedő produktivitásuk 0,099 és 0,048 közötti. (5) Reduplikálhatók. (6) A deverbális főnév tagadásakor elválhatnak – ezt a megfigyelést a 4. fejezet adatai alapján kiterjeszthetjük majd számos más igei származékra is a deverbális főneveken túl.

Ezek tehát a prototipikus igekötők olyan jellemzői, amelyek mérhető jegyek alapján meghatározhatók. A *meg* kitüntetett helyzete pusztán ezeket látva nem egyértelmű – úgy tűnik, egy a prototipikusak közül –, de egyéb, főként szemantikai szempontok alapján belátható. A *meg* nem irányjelölő, az ige jelentését olyan szempontból módosítja, ami grammatikai kategóriává vált, például aspektus és akcióminőség szempontjából, ezáltal az egyetlen magyar igekötő, amely jelentésánál a grammatikalizáció teljes útját bejárta. Nem volt tehát hibás az előfeltevés, hogy az igekötőség mértékét a *meg* jegyértékeihez viszonyítva érdemes megállapítani.

### 3.7. Összegzés

Ez a fejezet azzal a kérdéssel foglalkozott, hogy mely szavak tekinthetők igekötőnek, és milyen szempontok alapján. A kiindulási pontom az volt, hogy a kérdés nem dönthető el feketén-fehéren. Az igekötők és más igemódosítók közé nem húzható éles határ, hanem fokozatos az átmenet. Az igemódosítói szerepű lexikai elemek „igekötőségének a mértéke” aszerint határozható meg, hogy

<sup>25</sup> Az 50 kiejtett igemódosító a következő: állást bosszút bérbe csonttá élen érvénybe földet görcsbe gúzsba hadat hajba hatályba kupán kárba készen kézben kézen kézhez kútba küszöbön latba lángra lóvá nyomon nyomra nyélbe pórul rabul romba részre semmibe sorban számon számot szárba színre szóvá sárba síkra testet testre tetten tisztán tollba trónra táncra törbe valóra vérig

mennyire hasonlítanak egy prototipikus elemre, amelyre egyezményesen azt tudjuk mondani, hogy ez egy igekötő. Prototipikus igekötőnek a *meg*-et választottam a kimagasló gyakorisága, a történeti adatoltsága, a szakirodalomban egységes státusza és a saját intuícióm alapján.

A fejezet első részében a kutatás elméleti és módszertani háttéréről számoltam be. Röviden összefoglaltam az igekötő-fogalom meghatározásával kapcsolatos problémákat és nézetkülönbségeket, és kitértem arra, hogy az igekötő-állományt érintő bizonytalanság meglátszik a korpuszok meglehetősen következtelen annotációján is. Ezután bemutattam a prototípus-elméletet, amely lehetővé teszi az elmosódott kategóriahatárok kezelését, és megfogalmazható benne olyan állítás, hogy egy lexikai elem igekötő-szerűbb, mint más elemek. A prototípus-elméletet úgy kíséreltem meg átültetni a gyakorlatba, hogy meghatároztam és kimértem az igekötőségnek olyan morfoszintaktikai jegyeit, amelyek az MNSZ2 korpuszban jól kereshetők voltak. Az igekötőség számos más, lehetséges jegyét – például szemantikai jellegűeket – az automatikus mérhetőség hiánya miatt nem tudtam figyelembe venni.

A második részben, egyúttal a kutatás adatgyűjtési szakaszában az igekötőség mérhető jegyeivel foglalkoztam. Ezek a következők voltak:

- morfológiai produktivitás, amelynek Baayen (1989, 2009) nyomán három típusát különítettem el, ezek a megvalósult, a terjeszkedő és a lehetséges produktivitás
- a reduplikáció lehetősége (pl. *be-beoson*)
- elválás deverbális főnévtől tagadás esetén (pl. *össze nem férés*)
- a *-fele/felé* szócska csatolhatósága (pl. *nézzed megfele*)
- fokozhatóság (pl. *összébbhúz*)
- személyragozhatóság (pl. *rámnéz*)
- három általános jegy, amely minden szónál mérhető, de a prototipikus igekötő esetében jellegzetes értékei vannak: szerkesztettség, szótagszám és gyakoriság

Az adatgyűjtés eredményéből Igekötő × Jegy mátrixokat hoztam létre, amelyek elsősorban a jegyértékek ábrázolásában különböznek. A legfontosabb a korpuszon mért abszolút gyakoriságokat

tartalmazó mátrix, ebből vezettem le a relatív gyakoriságokat, valamint a ternáris (0 – 0,5 – 1 értékekre egyszerűsített) mátrixokat. Ezek nyílt hozzáférésű adatok, és további kutatásokhoz is felhasználhatók. Mivel a fejezet mottója örök érvényű, szeretném jelezni, hogy aki nem ért egyet az általam meghatározott igekötő-állománnyal, az ettől függetlenül felhasználhatja az Igekötő  $\times$  Jegy mátrixokat a saját munkájához, mivel ezek egyszerű mérési eredmények – teljesen objektív és elmélet-semleges adatok.

A fejezet harmadik, adatelemző szakaszában kiszámoltam minden jegypár korrelációját, először az abszolút gyakoriságokat tartalmazó mátrixon, majd a gyakorisági tényező teljes kivonásával egy bináris mátrixon (itt tehát csak ‘van’ vagy ‘nincs’ értéke lehetett a jegyeknek). A korrelációszámítás legfontosabb eredménye a következő: Világosan kimutathatóvá vált a grammatikalizáció folyamata az igekötők szerkesztettségének, szótagszámának, gyakoriságának és produktivitásának az összefüggései alapján. A megvalósult és a terjeszkedő produktivitás erős pozitív korrelációt mutat a gyakorisággal, míg a szótagszám és a szerkesztettség ezekkel negatív korrelációban áll – a gyakori és produktív igekötők jellemzően rövidek és morfológiai szempontból bonthatatlanok.

A bináris adaton végzett korreláció-számításból az derült ki, hogy a jegyek közti alapvető összefüggések nem változnak attól, ha a gyakoriságra vonatkozó információt elhagyjuk. Ennek az az oka, hogy a gyakoriság történetileg olyan szorosan együtt jár más jegyekkel – éppen a grammatikalizációs folyamat révén –, hogy akkor is érezhető a hatása az igekötőség jegyeiben, ha külön jegyként nem vesszük figyelembe.

Az adatelemző szakasz utolsó lépéseként arra tettem kísérletet, hogy a vizsgált lexikai elemeket a jegyértékeik alapján rangsoroljam egy tipikalitási skála mentén, és végül csoportosítsam ezeket. Ahogy jeleztem, ez nincs összhangban az igekötőség természetével, ezért mindössze a tárgyalást megkönnyítő kategóriák megnevezéséről van szó. A jegyalapú osztályozás céljából négy módszert próbáltam ki, amelyek röviden a következők:

- összegző módszer (minden igekötőnél vesszük a jegyértékeik összegét, és az összeg szerint csökkenő sorrendbe rendezzük az elemeket)
- halmazos módszer (minden jegyet halmazként értelmezünk, az adott jeggyel bíró igekötőket pedig az adott halmaz elemeiként – ezután aszerint sorrendezzük az igekötőket, hogy hány halmaz metszetében jelennek meg)

- k-közép klaszterelemzés
- HDBScan klaszterelemzés

A négy módszert ezután összehasonlítottam öt lényeges szempont szerint, amelyek a következők:

- Lehet-e bennük minden elemet kategorizálni?
- Automatikusan alakítják-e ki a kategóriák számát?
- Kiadnak-e tipikalitási skálát, vagy egymás mellé helyezik a csoportokat?
- Megoldható-e bennük a jegyek súlyozása?
- Használható-e bennük a (normalizált) abszolút gyakoriság, vagy egyszerűsített jegyértékekre van szükség?

Minden vizsgált módszerre igaz, hogy ebből az öt szempontból kettőnek nem felel meg. Így végül aszerint választottam közülük, hogy a kutatás célját tekintve melyik szempontokról lehet könnyebben lemondani, és az így fennmaradó lehetőségek közül melyik hozza az intuíciómhoz közelibb eredményt. A halmazos módszert választottam, és ennek alapján négy kategóriát neveztem meg: a prototipikus, a centrális, a félperiférikus és a periférikus igekötőket. A következő fejezetben látni fogjuk, hogy ezeknek a szintaktikai viselkedésében is eltérő tendenciák mutathatók ki.

## 4. Az igekötős szerkezetek típusai és szintaktikai jellemzői

„Ha elképzelni el tudja,  
miért nem tudja megalkotni?”

(korpuszpélda)

Az igekötők mondattani viselkedésük alapján önálló szavak, szemben például a szláv nyelvek prefixumaival, amelyek sosem válhatnak el az igétől. A magyar nyelv igekötői nem pusztán elválni képesek, hanem több szónyi távolságba is kerülhetnek (pl. *meg ugyan nem vettem, adtak már itt el*). Ez a tulajdonságuk csökken, de nem szűnik meg azzal, ha az igéből igenevet vagy főnevet képzünk (pl. *nem tévesztendő össze, vissza nem fizetés*). Még arra is találunk példát, hogy az igekötő az igenevekből képzett határozószók, melléknévek esetében sem tapad szorosan az igei származékhoz (pl. *meg sem születetten, a „be nem indulós Trabant”*). Ilyen és ehhez hasonló korpuszpéldák vizsgálatával a céloom két kérdés megválaszolása: Milyen mintázattípusokat mutatnak az igekötős szerkezetek? Mikor és mennyire távolodhat el az igekötő az igétől, igenévtől, igei származéktól?

Az igekötők mondatbeli helyének és szintaktikai önállóságának jelentős a szakirodalma. Azt viszont, hogy elválás esetén milyen távolságba kerülhet az igekötő, korábban nem vizsgálták, pedig egy ilyen kutatásnak gyakorlati és elméleti jelentősége is van. Az előbbire példa a nyelvtechnológiai felhasználás: gondoljunk egy szintaktikai elemző programra, amelynek az a feladata, hogy összekösse a mondatban az igekötőt azzal a szóval, amelytől elvált. Az utóbbi pedig abban nyilvánul meg, hogy az ‘igekötő – ige/igenév’ szórendet és az igekötő eltávolodását tekintve is kirajzolódnak olyan tendenciák, amelyeket csak adatvezérelt módszerrel lehet kimutatni.

A fejezet két nagy részből áll: egy szinkrón és egy diakrón korpuszvizsgálatból. A szinkrón korpuszvizsgálatban sorra veszek minden igekötős szerkezetet, és a mérési eredményeket az ige, igenév vagy igei származék típusa szerint tárgyalom. A diakrón esetében egy szűkebb jelenségkört vizsgálok, a finit ige és a hozzájuk tartozó prototipikus igekötők viszonyát, de ezt nagyobb adathalmazzal végzem, az ómagyar kortól napjainkig terjedő szöveganyag bevonásával. A két kutatáshoz külön-külön módszertani alfejezetek tartoznak, mert ezek eltérnek a felhasznált korpuszokat, valamint az adatgyűjtés és adatfeldolgozás részleteit tekintve is. Végül összefoglalom a két vizsgálat legfontosabb eredményeit.

## 4.1. Szinkrón vizsgálat: a mai magyar nyelvállapot

Ebben a fejezetben 20-21. századi szövegek vizsgálatán keresztül ismertetek minden olyan ige-kötős szerkezetet, amelyben az igekötő elválhat az igétől, igenévtől, vagy ritkábban egyéb igei származéktól. Ezek a következők:

- finit ige (pl. *megnéz, megnézek, ...*)
- infinitívusz (pl. *megnézni, megnéznem, ...*)
- határozói igenév (pl. *megnézve, megnézőn*)
- melléknévi igenév (pl. *megnézhető, megnézendő, megnéző, megnézett*) – ezzel együtt tárgyalom még a melléknévi igenévből képzett mellékneveket (pl. *megnézős*) és határozószókat is (pl. *megnézetten*)
- deverbális főnevek (pl. *megnézés*)

A felsorolt szerkezeteket elsősorban az igekötőik eloszlási mintái alapján jellemzem. Ahol csak lehet, kitérek az igekötő pozícióját befolyásoló tényezőkre, valamint a szakirodalomban kevésbé tárgyalt szintaktikai alakulatokra is.

### 4.1.1. A kutatás módszere

A szinkrón korpuszvizsgálathoz az MNSZ 2.0.4 módosított verzióját használtam (erről bővebben ld. a 2.2.1. fejezetet). Az adatgyűjtés első lépéseként lekértem a korpuszból minden olyan mondatot, amely valószínűleg tartalmaz igekötős igét vagy igenevet. Abból indultam ki, hogy az ilyen szerkezetek túlnyomó részében az igekötő egy tagmondatban van azzal a szóval, amelyhez tartozik. Ez alól vannak kivételek – pl. a *ki kell, hogy mondjam* típusú szerkezet –, ezek kigyűjtésére utólag végeztem célzott kereséseket. Szintén utólag gyűjtöttem össze az olyan eseteket, ahol az igekötő deverbális főnévtől vagy egyéb igei származéktól válik el (pl. *el nem engedős, szoros női barátság*). Ezután a lekért mondatokból lokálisan külön alkorpuszokat hoztam létre aszerint, hogy a vizsgált szerkezetben milyen típusú ige(név) szerepel. Így egy mondat több alkorpusznak is a része lehetett, és egy alkorpuszon belül többször is megjelenhetett:

[...], tehát minden egyes figyelembe vehető pont *el* kell legyen cseszve.

→ -hAtÓ végződésű melléknévi igenevek korpusza: *el+vehető*

→ finit igék korpusza: *el+kell*

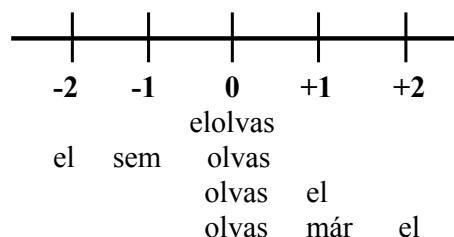
→ finit igék korpusza: *el+legyen*

→ határozói igenevek korpusza: *el+cseszve*

Az adatfeldolgozásnak egy jóval későbbi szakaszában döntött el, hogy a fenti példa négy lehetőségéből csak az utolsó érvényes, ezen a szinten még minden kombinációt szükséges volt feltételezni.

Az így létrejött összes alkorpusz minden mondatát szűrtem aszerint, hogy a feltételezett ‘igekötő + igei/igenévi lemma’ kombináció szerepel-e a PREVLEX táblázatban (erről bővebben ld. a 2.2.2. fejezetet). Ez magával hozza jó találatok elvesztését is (pl. *elé+gördít*, *belé+gebed*), egyszerűen azért, mert ezek nincsenek feltüntetve a PREVLEX-ben. Sokkal nagyobb mértékű viszont a rossz kombinációk kiszűrése (pl. *ki+tud*, *bele+visel*).

Minden mondat esetében meghatároztam az igekötő pozícióját. Ez is automatikus módszerrel történt úgy, hogy a mondatokat egy számegyenesre képeztem le, ahol az origó az az ige(név) volt, amelyhez az igekötő feltételezhetően tartozik.<sup>26</sup> A negatív tartományba kerültek a preverbális, a pozitív tartományba a posztverbális igekötők, ld. a 6. ábrát.



**6. ábra.** Az igekötők pozíciójának meghatározása az igehez vagy igenévhez képest.

A -1 pozíciót csak az olyan igekötőknél vettem figyelembe, amelyek gyakrabban jelennek meg az igétől különírva akkor is, ha közvetlenül az ige előtti pozícióban vannak (tehát amelyeknek ez a „hivatalos írásmódja”). Itt főleg a periférikus csoportba sorolt igekötőkre kell gondolni, pl. *szénné*, *pofon*. A -1 pozíciót ilyenkor 0-nak tekintettem, mert nem akartam pusztán helyesírási tényezők miatt élesebben elválasztani a két esetet. Ugyanakkor felmerül egy olyan helyzet, amikor a -1

<sup>26</sup> A finit igéket érintő korpuszméréseknél a ‘finit ige + *volna*’ szerkezetet egy szónak tekintettem, mivel a *volna* is időjeles ige, és szoros egységet alkot az őt megelőző, tartalmas igével. Így tehát az *olvasott volna el* példa esetében az *el* igekötő +1 pozícióba került, nem pedig +2-be.



pozíció nem pusztán helyesírási kérdés, hanem tényleg van önálló létjogosultsága. Ez akkor fordul elő, amikor az igekötő kontrasztív topik szerepű (és különösen akkor, ha a mondat egzisztenciális, vagyis egyszer már megtörtént eseményt hangsúlyoz), pl. *Be mentek már oda, de ki nem jöttek soha*. Ezek az itt bemutatott mérésekből kimaradtak, de tervezem megvizsgálni azt, hogy az eredeti -1 pozíciós adatok között mennyi ilyen szerkezet van.

Ezután következett a munka leghosszabb szakasza: minden alkorpusz minden pozícióját tanulmányozni kellett, majd külön szűrőszabályokat írni rájuk. A szabályok általában szófaji címkesorozatokra illeszkedő reguláris kifejezések. Egy gyakran használt szűrőszabályom például az, hogy ha 'finit ige + igekötő + finit ige' sorozat szerepel a találatban, akkor az kiszűrendő, mert hibás annotációról van szó (pl. *erre ő csak állt meg bámult*, ahol a *meg* kötőszó tévesen igekötőnek elemzett). Ezeket a szabályokat aztán egyesével kellett futtatni az adott pozíció mondatain, és ellenőrizni, hogy valóban a hibás találatokat szűrik-e, és nem illeszkednek-e sok helyes találatra.

A fentiekből következik, hogy automatikus szűrésekkel nem lehet hibátlan kiinduló adatot előállítani, viszont lehet arra törekedni, hogy az adat minél tisztább legyen.<sup>27</sup> Ekkora korpuszméret esetén a kézi ellenőrzés reménytelen vállalkozás. Fontos tehát hangsúlyozni, hogy a továbbiakban bemutatásra kerülő eloszlási minták fenntartással kezelendők. A megadott darabszámok hibás találatokat is magukban foglalnak, de a vizsgált értékek egymáshoz való viszonya ezek ellenére is jól tükröződik.

#### 4.1.2. Finit ige

Az igekötős szerkezetek közül messze a leggyakoribbak a finit igei szerkezetek, ami egyáltalán nem meglepő, mivel a finit ige – amelynek ideje, száma, személye van – általában a mondat főigéje. Mielőtt ismertetném az igekötők disztribúciójára vonatkozó korpuszmérések eredményeit, érdemes tisztázni néhány fogalmat, amely szükséges az adatok megértéséhez.

Az első ilyen fogalompár a semleges és a nem-semleges mondat, mert a két típus különbsége az igekötő-ige szórendben is megmutatkozik. Alberti (2006: 37-38) három nyelvészeti terület irányából közelíti meg a semleges mondat fogalmát. Szemantikai szempontból azt mondhatjuk, hogy nincs benne szerkezeti többletjelentés, vagy csak minimális. Hangtanilag semleges intoná-

<sup>27</sup> Ezzel a problémakörrel a szakdolgozatomban és két publikációmban is foglalkozom, ld. Kalivoda (2016, 2017, 2018). Mivel ez inkább technikai jellegű téma, nem tartom igazán ebbe a dolgozatba illőnek – az említett munkákban lehet erről bővebben tájékozódni.

ció jellemzi, vagyis a mondatban a tartalmas szavak az alapértelmezett kezdőszótagi hangsúlyukat viselik, a grammatikai funkciószavak hangsúlytalanok. Pragmatikai szempontból a semleges mondatban bizonyos szavak előfeltevést hordoznak, de a szórendi és intonációs sajátosságok nem hoznak további előfeltevéseket. A nem-semleges mondatban a hangsúly eloszlása nem egyenletes, és speciális jelentés-többletet hordoz, amely szerkezetileg is megjelenhet (Alberti 2006: 54). Az igekötő semleges mondatban jellemzően preverbális,<sup>28</sup> nem-semleges mondatban lehet pre- és posztverbális, a mondat típusától és a szándékolt jelentéstől függően. A 16. táblázat áttekintést nyújt a kombinációs lehetőségekről.

Mondattípus	Preverbális igekötő	Posztverbális igekötő
Kijelentő (semleges)	Lemegyek a közértbe.	-
Kijelentő (progresszív)	-	Verik le a képet.
Kijelentő (egzisztenciális)	-	Veszítettem már el könyvet.
Felszólító	Elhallgass végre!	Menj el hozzájuk!
Felkiáltó	Milyen korán felkelt!	Milyen későn jött meg!
Tagadó	El sem hiszem.	Nem hiszem én már el.
Tiltó	Meg ne kérdezd!	Ne mondd el!
Óhajtó	Bárcsak elhinné végre.	Csak szokna le Valenciáról!
Kérdő (eldöntendő)	Eljössz a találkozóra?	-
Kérdő (kiegészítendő)	-	Miért mennek el máris?
Szerkezeti fókuszos	-	Így került végül meg.

**16. táblázat.** Mondattípusok, amelyekben preverbálisan, posztverbálisan vagy mindkét módon megjelenik az igekötő, egy-egy példával.

A 16. táblázatban felsorolt mondattípusok közül a progresszív és egzisztenciális kijelentő mondatok igényelnek magyarázatot. Az előbbieket esetében azon van a hangsúly, hogy éppen történik az adott esemény, az utóbbiaknál pedig azt emeljük ki, hogy az esemény már legalább egyszer előfordult.

Az igekötők finit igéhez viszonyított, pozíció szerinti eloszlásának ismertetését<sup>29</sup> **a preverbális igekötőkkel** kezdem. Az igekötők kanonikus helye a közvetlenül preverbális – „az igével egybeírt” – pozíció. Az igétől elvált preverbális igekötőket intuitívan is jóval ritkébbnek vélhetjük, mert a magyar mondatok bal perifériáján az összetevők elrendeződése sokkal megszorítottabb, mint az ige

<sup>28</sup> Létezik ige–igekötő sorrend semleges mondatban is, ld. Kádár (2007) és Peredy (2011) csángó nyelvjárási adatait. A magyar köznyelvre ez nem jellemző.

<sup>29</sup> Ilyen méréseket korábbi munkáimban is bemutattam (Kalivoda 2016, 2018). Az itt közölt adatok számszerűen kissé eltérnek a korábbiaktól, aminek az az oka, hogy a publikációk óta a korpuszmérést nagyobb szöveganyagon és több igekötő bevonásával végeztem el.

utáni mondatszakaszban. A korpuszadatokban szélsőértéknek az igekötő -6-os pozíciója bizonyult, vagyis az igekötők és a finit ige között legfeljebb öt másik szó állt.

<b>IK-csoport</b>	<b>-6</b>	<b>-5</b>	<b>-4</b>	<b>-3</b>	<b>-2</b>	<b>0</b>
prototipikus	-	2	1	1 638	251 354	9 989 626
centrális	5	7	324	403	53 174	1 796 921
félperiférikus	-	2	9	463	11 326	267 078
periférikus	-	8	15	195	15 011	428 040
<b>összes db</b>	<b>5</b>	<b>19</b>	<b>349</b>	<b>2 699</b>	<b>330 865</b>	<b>12 481 665</b>
prototipikus	-	-	-	0,02	2,45	97,53
centrális	-	-	0,02	0,02	2,87	97,09
félperiférikus	-	-	-	0,17	4,06	95,77
periférikus	-	-	-	0,04	3,39	96,56
<b>összes %</b>	<b>-</b>	<b>-</b>	<b>0</b>	<b>0,02</b>	<b>2,58</b>	<b>97,39</b>

**17. táblázat.** Lehetséges igekötő pozíciók a finit ige bal oldalán. Az adatok előbb a négy igekötő-csoport szerinti bontásban láthatók, majd – vastagon szedve – csoportra bontás nélkül. A fejléc alatti öt sor abszolút, a legalsó öt sor relatív gyakoriságokat tartalmaz.

A 17. táblázatban látható, hogy a közvetlenül preverbális helyzetű igekötők aránya kiemelkedően magas: közel 12,5 millió találattal a 0 pozíció fedi le a preverbális esetek 97,39%-át. A következő olyan pozíció, ahol a finit igehez tartozó igekötő állhat, a -2. Ide leggyakrabban *is*-szerkezet (16a) vagy annak tagadása (16b, 16c), tagadó felszólítás, tiltás (16d) vagy tagadás (16e) esetén kerül az igekötő. Az igekötőt az igétől elválasztó szavak tételesen felsorolhatók: *is, sem, se, nem, ne, sose, sosem, sohase, sohasem*.

- (16) a. 10-12 óra között **le** is **zárják** a teret.
- b. Pedig még **ki** sem **próbáltam**.
- c. **Vissza** se **gyere**!
- d. Aztán **fel** ne **fázz**!
- e. Egy 15 milliós terepjárával **be** nem **mennék** a fák közé.

Az említetteken kívül azokat a tagmondatokat is a -2 pozícióhoz számoltam, ahol az igekötőhöz az -e simulószó járul, és ez választja el a finit igétől. Erre a sztenderd nyelvhasználatban szokatlan szerkezetre összesen 52 találat volt a korpuszban, amelyek közül hármat mutat be a (17) példásor.

- (17) a. És **meg-e találják** vajon a hetediket?
- b. [...] néztek engem és azon gondolkoztak, hogy **fel-e keltsenek**.
- c. A magyar területen **át-e jöttek**, azt én nem láttam.

Amíg -2-ben több mint 330 ezer igekötőt lehetett találni, -3 pozícióban mindössze 2 699 állt. A példák túlnyomó többsége tagadó szerkezet, és az igekötőt a leggyakrabban a következő szósorok választják el az igétől: *nem is, már nem, még nem, soha nem, én nem, senki sem, ugyan nem*, ld. a (18) példákat.

- (18) a. Azt feltettem, **le** már nem **szedem** potyára.
- b. **Le** sem nagyon **ütötte** a labdát szegény, csak csinálta, amit az edző kért tőle.
- c. **Meg** ugyan nem **vettem**, de elég alaposan beleolvastam.
- d. Ami viszont biztos, hogy **el** semmiképpen nem **vesz** a szórakozásból, [...]
- e. [...] egy hálátlan kutya, aki **meg** sem igen **érdemli** azt a nagy szerencsét, ami érte.

A -4 pozíciós igekötők már elhanyagolhatóan kevesen vannak, ezekre mindössze 349 találat volt. Ezek leginkább centrális igekötők – különösen a *haza, vissza, ide, oda* –, ritkább esetben (fél)periférikusak, ld. a (19) példákat.

- (19) a. Még **haza** se nagyon szívesen **eresztik**.
- b. **Vissza** ezt már úgysem **szívják**...
- c. [...], hogy **oda** még véletlenül sem **tévedhet** a gyerek.
- d. De **tönkre** ebbe én nem **mentem**.
- e. Ha **egyét** gyakran nem is **értek** vele, de ezt azért becsülöm.

Elvértve még -5 és -6 pozícióban is találunk igekötőt, mindig kontrasztív topikként. Itt is a centrális és (fél)periférikus igekötők jellemzőek, ld. a (20) példasort.

- (20) a. [...] hét hónapot voltam otthon, **vissza** már nem az eredeti munkakörömbe **jöttem**, [...]
- b. [...], ketten mentünk, ám **haza** már hárman egy potyautassal **érkeztünk**.
- c. [...] mit **egybe** csak a súly maltera **köt**.

A preverbális igekötőkkel kapcsolatban érdemes kitérnünk egy olyan szerkezetre, amellyel É. Kiss (2003, 2007) részletesen foglalkozik. A szerkezet sajátossága az, hogy az igekötőt időjeles mellékmondatból – kötőmódban álló finit ige mellől – fölérendelt tagmondatba emeljük ki (pl. *el kell, hogy engedj*). Ha a fölérendelt tagmondatban a *kell* ige szerepel, az igekötő kiemelése kötelező (vö. *meg kell, hogy tudd*  $\leftrightarrow$  *\*kell, hogy megtudd*). A *szabad* esetében a kiemelés opcionálissá válik (vö. *el szabad, hogy vezessen* és *szabad, hogy elvezessen*). Mindez a pusztá névszói igemódosítókra is érvényes. É. Kiss (2007) megállapítja a következőket: (1) Ha a *hogy*-kötőszós tagmondat tagadó vagy fókuszos, nem lehet kiemelni az igemódosítót. (2) Egy-két nem operátor-szerepű kifejezésen az igemódosító átemelhető. (3) A kiemelés feltétele a *kell/szabad* és a *hogy* közvetlen szomszédossága.

A korpuszvizsgálat során összesen 3 537-szer fordult elő ez a szerkezet, minden esetben a *kell* igével (amely néha *kellene, kéne* alakban áll). A kiemelt igekötő és a *kell* között csak az *is* és a *sem* klitikumok állhatnak, az időjeles mellékmondatban viszont több és többféle szó is megjelenhet a kötőmódú finit ige előtt, ld. a (21) példasort.

- (21) a. A konkurencia miatt **meg** kell, hogy majd **változzon** a MATÁV üzletpolitikája.  
 b. [...], akkor azért **hozzá** kell, hogy azt is **tegyük**, [...]  
 c. **El** kell, hogy az anyja **engedje**.  
 d. Nyilván mindenki számára és azért azt **be** kell, hogy én magam is **lássam**, [...]

Az igekötők eloszlási mintáinak vizsgálatát a hátravetett, más szóval **posztverbális igekötőkkel** folytatom. A hátravetett igekötőt tartalmazó tagmondatok – a 28-as lábjegyzetben említett, nagyon ritka példák kivételével – nem-*semlegesek*. Érdekes adat, hogy +1 pozíciójú igekötőt közel 8 millió mondatban találunk az MNSZ2-ben. Ez a 0 pozíció gyakoriságával összevetve azt jelenti, hogy az igekötők az esetek több mint egyharmadában posztverbálisak.

A 18. táblázat mutatja azt a tendenciát, hogy az igekötők jóval nagyobb intervallumon helyezkedhetnek el a finit ige utáni mondatszakaszban, mint az ige előtt, de hátravetéskor 97,15%-ban mégis közvetlenül az ige mögé kerülnek. A relatív gyakorisági adatok alapján azt is megállapíthatjuk, hogy a prototipikus igekötők távolodnak el legkevésbé az igétől (98%-ban közvetlenül követik), míg a másik három csoport tagjainál a +1 pozíció relatív gyakorisága 92,5 és 95,7% kö-

IK-csoport	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10	+11
prototipikus	6 018 141	108 978	10 865	4 061	2 095	55	3	-	1	-	-
centrális	1 311 130	48 091	7 455	2 296	922	324	21	6	-	-	1
félperiférikus	203 998	11 798	3 027	1 194	524	40	6	1	-	1	-
periférikus	391 254	18 606	5 891	4 417	1 719	21	8	2	-	-	-
<b>összes db</b>	<b>7 924 523</b>	<b>187 473</b>	<b>27 238</b>	<b>11 968</b>	<b>5 260</b>	<b>440</b>	<b>38</b>	<b>9</b>	<b>1</b>	<b>1</b>	<b>1</b>
prototipikus	97,95	1,77	0,18	0,07	0,03	-	-	-	-	-	-
centrális	95,69	3,51	0,54	0,17	0,07	0,02	-	-	-	-	-
félperiférikus	92,48	5,35	1,37	0,54	0,24	0,02	-	-	-	-	-
periférikus	92,73	4,41	1,40	1,05	0,41	-	-	-	-	-	-
<b>összes %</b>	<b>97,15</b>	<b>2,30</b>	<b>0,33</b>	<b>0,15</b>	<b>0,07</b>	<b>0</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>-</b>

**18. táblázat.** Lehetséges igekötő pozíciók a finit ige jobb oldalán. Az adatok előbb a négy igekötő-csoport szerinti bontásban láthatók, majd – vastagon szedve – csoportra bontás nélkül. A fejléc alatti öt sor abszolút, a legalsó öt sor relatív gyakoriságokat tartalmaz.

zött alakul. Megfogalmazhatunk tehát egy olyan általánosítást, hogy minél igekötő-szerűbb egy elem, annál valószínűbb, hogy posztverbális helyzetben közel marad a finit ígéhez.<sup>30</sup>

A (22) példaszor egy-egy mondattal illusztrál minden pozíciót:

- (22) a. Ennek nyomán 17 megyében **nyújtottak be** felülvizsgálati kérelmet [...] +1
- b. – **Heverj** csak **el** te is nyugodtan – mondta Pilátus. +2
- c. Akkor hogy **függenek** itt most **össze** a szándékok és a bekövetkező események? +3
- d. Én mindenesetre Dancsó úrral **értek** ebben a dologban **egyedül**. +4
- e. Háromszor **veri** ezt kenden Lúdas Matyi **vissza**! +5
- f. Azért **mentem** egy kicsit a popzene felé **el**, [...] +6
- g. Tehát ezért **függenek** ezek a folyamatok egymással nagyon szorosan **össze** [...] +7
- h. Ezen leánnyal **utazott** ön azután egyszer később a héven estefelé **haza**? +8
- i. Azt mondják, száz napból tizenöt-húszban **sodor** a szél dögszagot a sárvári ATEV üzem felől **ki** az országútra. +9
- j. [...], hogy miért **teszi** megannyi egyébként oly sikeresnek látszó ember saját magát szánt-

<sup>30</sup> Ebből az is következik, hogy a finit igtől való eltávolodás mértéke az igekötőség morfoszintaktikai jegye lehet, és a jegyalapú osztályozásban – ld. 3.6. fejezet – is felhasználható.

szándékkal **tönkre**. +10

- k. 27 gyereket **vitt** egy feltehetően részeg buszsofőr Szentesen még csütörtökön egy sportrendezvény után **viszsa** az iskolába. +11

A (22i) és (22k) példákban egy helyhatározó áll közvetlenül az igekötő előtt illetve után, így felmerülhet az a probléma, hogy az igekötő sokkal inkább a helyhatározóhoz, mintsem a finit igéhez kötődik. További példákhoz ld. Alberti (2006: 50).

A hátravetett igekötők eloszlási mintáinak ismeretében érdemes megvizsgálni **az ige–igekötő távolság lehetséges okait**. A 18. táblázat tanúsága szerint a posztverbális igekötők 99,5%-ban közvetlenül vagy egy szónyi távolságban követik a finit igét. Ennek a pszicholingvisztikai magyarázata elsősorban a távolsági hatásban keresendő, miszerint „minél nagyobb a távolság a szöveg alkotóelemei között, annál nagyobb a hiba lehetősége és annál kapacitásigényesebb az összekapcsolás” (Németh 2001: 94). Hasonló gondolatot fogalmaz meg Alberti (2006: 82) is: „Fontos forrása azonban a jelentés-kalkulációnak a lexikonból hozott jelentés is, amit viszont nehezebben felismerhetővé tesz például az, ha az igekötő(szerűen viselkedő elem) eltávolodik az igétől.”

Az igekötők szerepet játszanak az ige irányította mondatfeldolgozásban, argumentumok keresésében (Pléh 1998: 133-135), ezért intuitívan is számítani lehetett rá, hogy az ige elhangzása után minél előbb következnie kell az igekötőnek, amely módosítja az ige jelentését. Különösen igaz ez a prototipikus igekötőkre, amelyek nagyon sok igével kombinálhatók. A testesebb, részben idiomatikus igekötők (pl. *tönkre*, *ketté*) kevesebb igével állhatnak, az elhangzott kontextusból rendszerint kiszámítható, hogy ezek az igekötők fognak még elhangzani. Ezek távolabb kerülhetnek az igétől: a *tönkre*, ahogy a (22j) példában láttuk, +10 pozícióban is előfordul.

Fontos lenne magyarázatot adni az adatok 0,5%-ára is, tehát azokra, amelyekben az ige és az igekötő között kettő vagy több szó áll. Két szempontból<sup>31</sup> elemzem a korpusz adatait: (1) A metaadatok, főként írott és beszélt nyelvi oppozíció, valamint a stílusregiszter alapján. (2) Fonológiai szempontból, amelyet a szavak hosszának kimérésével lehetett megvalósítani.

Az első szempont, tehát az írott és beszélt nyelvi stílus kapcsán abból indultam ki, hogy a szóbeli megnyilatkozások alapvetően tervezetleneek, spontának az írásbeliekhez képest, így szintakti-

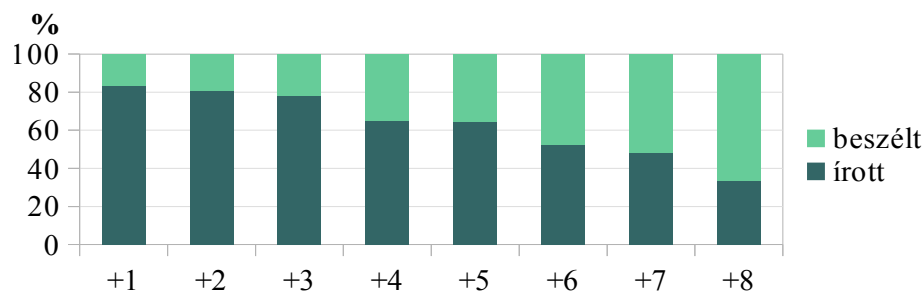
<sup>31</sup> Érdemes volna szemantikai szempontból is megvizsgálni a posztverbális igekötős igéket, mert az is szerepet játszhat az igekötő eltávolodásának mértékében, hogy szemantikailag mennyire transzparens az adott igekötős ige. Ennek megvalósítása szisztematikusan, számítógépes módszerekkel egyelőre túl nagy vállalkozásnak tűnt.

kialakulására kevésbé szabatosak (Goody 1987: 258-289). Ezért arra számítottunk, hogy a szélsőbb pozíciók ígékötőt tartalmazó mondatok többsége eredetileg szóban hangzott el. Az MNSZ2 meta-adatai segítségével viszonylag egyszerűen lehetett automatikus vizsgálatot végezni azzal kapcsolatban, hogy melyik ígékötői pozícióra milyen szövegtípus jellemző. A szövegtípusra vonatkozó adatokat a következő kategóriákba rendeztem:

(1) *Írott nyelvi*: sajtó, személyes (ezek jellemzően internetes blogok és fórumhozzászólások), tudományos, szépirodalom (külön a próza és a vers), hivatalos (többnyire jogszabályok).

(2) *Beszélt nyelvi*: rádióban elhangzott szöveg, felszólalás.

A 7. ábra azt mutatja be, hogy az egyes pozícióknál milyen arányban reprezentáltak az írott illetve beszélt nyelvi szövegek.



**7. ábra.** Írott és beszélt nyelvi szövegek százalékos aránya az egyes pozíciókon. A távolabbi pozíciókon már csak egy-egy találat volt, ezért ezeket nem vettem figyelembe.

Az adatok értelmezéséhez tudni kell, hogy a teljes MNSZ2-ben a beszélt nyelvi adat nagyon kevés (Oravecz et al. 2014): a rádiós szövegek a teljes szövegállomány 5,4%-át alkotják, és ha a hivatalos szövegek felét szóban elhangzottnak tekintjük, a végeredmény akkor sem több 9,4%-nál. A +1 pozíció szövegtípus szerinti megoszlása még nem tér el nagyban a teljes MNSZ2-től: lényegesen több írott szöveg a forrása (83,51%), mint beszélt nyelvi (16,49%). A beszélt nyelvi szövegek aránya lassacskán növekszik: minél távolabbi ígékötő-pozíciót vizsgálunk, annál valószínűbb, hogy szóban hangzott el a forrásmondat. A +7 pozíciónál már a mondatok több mint fele szóbeli szövegből származik. A (23) példaszor első három tagmondata országgyűlési felszólalásban, az utolsó három pedig rádióműsorban hangzott el.

- (23) a. [...] **tette** a minisztérium ezt a pályázati felhívást **közzé** [...]  
 b. [...] **került** talán ebben a jogharmonizációs folyamatban **elő** [...]  
 c. [...] **támasztom** ezt a tőlem talán szokatlan megfogalmazást **alá** [...]



- d. [...] **dobja** valakinek az egyéniségét a hullámhegyre **fel** [...]
- e. [...] **merül** ez a kérdés ilyen sarkallatosan **föl** [...]
- f. [...] **jámulhatott** Ön szerint ehhez a szörnyű tethez **hozzá** [...]

Azt mondhatjuk tehát, hogy a szövegtípus és a posztverbális igekötő lehetséges eltávolodása összefüggenek: szerkesztetlen – főleg szóban elhangzott – szövegben nagyobb annak a valószínűsége, hogy az igekötő távol kerül a finit igétől, mint gondozott, szerkesztett szöveg esetében.

Térjünk rá a második szempontra, a fonológiai tényezőkre. Otto Behaghel német nyelvész a *Deutsche Syntax* negyedik kötetében öt törvényt – pontosabban tendenciát – fogalmazott meg, amelyekkel magyarázatot kívánt adni a szavak és frázisok mondatbeli elhelyezkedésére (Behaghel 1932: 4-8). Ezek közül a negyedik azt mondja ki, hogy a rövidebb összetevő megelőzi a hosszabbat, ha ezt szintaktikai szabály nem gátolja.<sup>32</sup>

É. Kiss (2007) amellet érvel, hogy ez a tendencia – a növekvő összetevők törvénye – a magyarban a finit ige utáni mondatszakaszra érvényes: az ige utáni, alapvetően szabad szórendet befolyásolja az, ha az ígét követő összetevők fonológiaiilag különböző súlyúak. Ezt többek között a (24) példával szemlélteti. Az ideális sorrend a (24)-ban finit ige után az ‘igekötő – határozó – tárgy’, egyéb kombinációk legalábbis szokatlanul hangzanának.

(24) János sértette meg nagyon Éva nagyapját.

A bemutatott tendencia jól látszik a posztverbális igekötőt tartalmazó mondatok esetében. A mérést a következőképp végeztem el: Az olyan tagmondatokban, amelyek +1, +2 és +3 pozícióban tartalmaztak igekötőt, kimértem az ige után álló három szó hosszát (karakterzámban), és átlagoltam az egyes pozíciók hossz-értékeit. Az eredményeket a 19. táblázat foglalja össze.

A nagyon gyakori (97,15%-os), +1 pozíciójú igekötőt tartalmazó mondatokban az igekötő átlagosan 2,9 karakter hosszú, az ezt követő szavak 4,1, majd 4,7 hosszúak. A növekvő összetevők törvénye a +2 pozíciós igekötők esetében is érvényesnek látszik: az átlagosan 3,4 hosszú igekötőt annál valamivel rövidebb szó (3,1) előzi meg, és lényegesen hosszabb követi (4,5). A tendencia +3 pozíciós igekötők esetén már érvényét veszti: az átlagosan 4,2 hosszúságú igekötőt nála hosszabb

<sup>32</sup> Ezt a jelenséget Pāṇini, a 4. században élt szanszkrit grammatikus is felismerte, a szakirodalomban Pāṇini törvényeként is hivatkoznak rá, ld. O’Connor (1978: 97-99) és Cooper és Ross (1975: 71). Ő a szanszkrit szóösszetételek vonatkozásában figyelte meg azt a tendenciát, hogy két összetételi tag közül jellemzően a rövidebb kerül előrébb.

alkorpusz	a finit ige utáni 3 szó átlagos hossza			%	a finit ige és az igekötő között álló leggyakoribb szavak
+1	2,9	4,1	4,7	97,15	-
+2	3,1	3,4	4,5	2,30	<i>majd, már, csak, -e, még, is, ezt, itt, most, ez, ...</i>
+3	3,5	5,4	4,2	0,33	<i>egy kicsit, én is, -e már, már csak, -e majd, ...</i>

**19. táblázat.** A növekvő összetevők törvénye – számokban. Az *alkorpusz* azt mutatja, hogy melyik igekötő-pozícióhoz tartozó tagmondatokról van szó, a 3. % oszlop pedig ezek aránya a posztverbális igekötők korpuszában.

szó előzi meg. Fontos észrevennünk, hogy ez a helyzet csak nagyon kevés mondatban áll fenn, a +3 pozíciós igekötők a posztverbális eseteknek mindössze 0,33 százalékát alkotják.

Látni kell azt is, hogy ez a kísérlet nem a mondatösszetevők, hanem a szavak fonológiai súlyát vizsgálja, mert az MNSZ2 korpusz annotációja ezt teszi lehetővé. Ezzel tehát a növekvő összetevők törvénye még nem teljes körűen bizonyított, viszont az eredmények alapján nagyon valószínűnek tűnik, hogy a magyarban az ige utáni szabad szórend esetén a fonológia hat a szintaxisra, és az egyes mondatösszetevők sorrendjét befolyásolja a fonológiai súlyuk.

#### 4.1.3. Infinitívusz

A hagyományos grammatikák az infinitívuszt főnévi igenévnek hívják. Ez az elnevezés arra utal, hogy az infinitívusznak számos főnévi tulajdonsága van: azonos jelentésű az -Ás képzős deverbális főnevekkel (vö. *írni fárasztó* → *az írás fárasztó*), felvehet birtokos személyragokat (pl. *írnom, írnod*), és megjelenhet minden olyan mondatrészi szerepben, amelyben a főnevek, főnévi kifejezések is. Az itt tárgyalt téma szempontjából viszont lényegesebbek az igei tulajdonságai,<sup>33</sup> és ezeket jobban megragadja az ‘infinitívusz’ elnevezés, amely arra utal, hogy ezek az igeenes kifejezések értelmezhetők időjel nélküli tagmondatként is (vö. *kitörni készült* → *arra készült, hogy kitörjön*).

Az infinitívusz képzője a *-ni*, bár Kiefer (2003a: 175) megjegyzi, hogy ez a képzés korlátlanul termékeny, és a létrejövő szó jelentése kompozicionális, ezért a *-ni* inkább az inflexiók toldalékokhoz hasonlít. Az infinitívusz felvehet személyragokat, ekkor a szerkezetben megjelenhet olyan részesesetű névszói kifejezés, amely az infinitívusz személyragjával számban és személyben egyezik, jelentésánál pedig az alanyaként funkcionál (É. Kiss 2003: 77), pl. *(nekem) olvasnom kell*.

<sup>33</sup> Az infinitívusz igei tulajdonsága például az, hogy lehet tárgya, viszont főnévi bővítményei – jelző, névelő – nem lehetnek.

Az igekötők eloszlási mintáinak vizsgálatakor szükségesnek tartottam két szerkezettypus elkülönítését. Az elsőnél nem férkőzik be más ige a szerkezetbe, a másodiknál igen.<sup>34</sup> Mindkét esetben lehetséges az igekötő elválása az infinitívustól a preverbális és a posztverbális tartományban is, de a mozgásterét tekintve nagy különbségek fedezhetők fel.

A tárgyalást azzal a szerkezettypussal kezdem, amelynél nincs beférkőző ige vagy igenév, ehhez ld. a 20. táblázatot.

<b>IK-csoport</b>	<b>-4</b>	<b>-3</b>	<b>-2</b>	<b>0</b>	<b>+1</b>	<b>+2</b>
prototipikus	5	47	4 788	2 078 633	247	9
centrális	2	27	1 619	350 269	301	11
félperiférikus	3	8	1 036	60 831	330	17
periférikus	-	1	595	82 608	106	3
<b>összes db</b>	<b>10</b>	<b>83</b>	<b>8 038</b>	<b>2 572 341</b>	<b>984</b>	<b>40</b>
prototipikus	-	-	0,23	99,76	0,01	-
centrális	-	0,01	0,46	99,44	0,09	-
félperiférikus	-	0,01	1,66	97,76	0,53	0,03
periférikus	-	-	0,71	99,15	0,13	-
<b>összes %</b>	<b>-</b>	<b>0</b>	<b>0,31</b>	<b>99,65</b>	<b>0,04</b>	<b>0</b>

**20. táblázat.** Lehetséges igekötő pozíciók az infinitívusznál abban az esetben, ha a szerkezetbe nem férkőzik be finit ige vagy határozói igenév. Az adatok előbb a négy igekötő-csoport szerinti bontásban láthatók, majd – vastagon szedve – csoportra bontás nélkül. A fejléc alatti öt sor abszolút, a legelső öt sor relatív gyakoriságokat tartalmaz.

A (25) példasor illusztrál minden pozíciót.

- (25) a. [...], hogy felsorolásukba **bele** sem is érdemes **fogni**. -4
- b. Pláne, ha módosult tudatállapotában **föl** sem képes **fogni**, mi történik vele. -3
- c. Keveseknek lenne lelkierije **meg** is **tenni** ezt egy ilyen helyzetben! -2
- d. Nagyon nehéz **előjönni** a sötét mélységből. 0
- e. [...], de egy vidéki városban, faluban nem feltétlenül **találni meg** eme szerszámat. +1
- f. A műfordítás kétségtelenül úgy jó, ha nem **venni** rajta **észre**, hogy műfordítás. +2

<sup>34</sup> A vizsgált korpuszban a leggyakoribb beférkőző igék a következők: *kell, tud, lehet, fog, akar, kíván, szeret(ne), szokik* (múlt időben), *mer, tetszik* (mint udvariassági formula), *kezd, bír, próbál, szabad, talál, szándékozik, sikerül, merészel, óhajt, igyekszik*. Ezeket a továbbiakban segédige-szerű elemeknek hívom.

Látjuk, hogy a hátravetett igekötős esetek ritkák: akkor fordulnak elő, ha az infinitívuszt tagadjuk vagy saját fókusza van. A -4 és -3 pozíció is nagyon ritka, itt az az érdekes, hogy szinte mindig az *érdemes*, *képes*, *szokás* szavak egyike veszi át a finit ige helyét állítmányi szerepben, és ez elé kerülhetnek az igemódosítók. Bár ezek szófajilag nem igék, a beférkőzés képessége szempontjából a segédige-szerű elemek közé sorolhatók. Kálmán C. et al. (1989) ezek közül a *szokás*-t be is sorolják a segédigék csoportjába, és jelzik, hogy funkcionális szempontból más hasonló szavakra is tekinthetünk így.

Térjünk rá a második szerkezet típusra, amelyben finit ige vagy határozói igenév férkőzik be az infinitívuszi szerkezetbe. Az igekötők disztribúcióját ld. a 21. táblázatban.

IK-csoport	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	+2
prototipikus	4	7	29	96	223	344	3 865	9 268	61 733	652 289	153
centrális	1	-	3	10	35	106	933	2 148	11 205	128 470	409
félperiférikus	-	-	-	3	8	45	573	1 084	3 658	39 046	380
periférikus	-	-	-	2	4	11	248	548	3 596	37 861	21
<b>összes db</b>	<b>5</b>	<b>7</b>	<b>32</b>	<b>111</b>	<b>270</b>	<b>506</b>	<b>5 619</b>	<b>13 048</b>	<b>80 192</b>	<b>857 666</b>	<b>942</b>
prototipikus	-	-	-	0,01	0,03	0,05	0,53	1,27	8,48	89,60	0,02
centrális	-	-	-	0,01	0,02	0,07	0,65	1,50	7,82	89,64	0,29
félperiférikus	-	-	-	0,01	0,02	0,10	1,28	2,42	8,17	87,16	0,85
periférikus	-	-	-	-	0,01	0,03	0,59	1,30	8,51	89,57	0,05
<b>összes %</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>0,01</b>	<b>0,03</b>	<b>0,05</b>	<b>0,59</b>	<b>1,36</b>	<b>8,37</b>	<b>89,49</b>	<b>0,10</b>

**21. táblázat.** Lehetséges igekötő pozíciók az infinitívusznál, beférkőző szerkezet esetén. Az adatok előbb a négy igekötő-csoport szerinti bontásban láthatók, majd – vastagon szedve – csoportra bontás nélkül. A fejléc alatti öt sor abszolút, a legelső öt sor relatív gyakoriságokat tartalmaz.

A (26) példasor illusztrálja az egyes pozíciókat. Két megfigyelést biztosan tehetünk: (1) Bár az igekötő preverbálisan nagyon messzire eltávolodhat az infinitívustól, mindig közel van a finit igehez, az ahhoz tartozó igemódosítói pozíciót foglalja el. Közte és a finit ige között csak klitikum-szerű szavak állhatnak (*is*, *se*, *sem*, *ne*, *nem*). (2) Segédige-szerű elem akkor is be tud férkőzni a szerkezetbe, ha az infinitívusz igekötője hátravetett, ld. (26k) példa.

- (26) a. **Meg** tudják az emberek ezt egy tulajdonosi közösségi gyűlés közben is **beszélni**. -11  
b. [...] **el** tudnánk magunknak a saját szánk íze szerinti tökéletes édességet **készíteni**. -10  
c. [...] **fel** kell tudniuk az olívaolajtermelésre szánt támogatás egy részét **használni** [...] -9

- d. [...], természetesen **hozzá** lehet majd az ön ügyrendi indítványához is **szólni**. -8
- e. [...], ha **össze** lehetne őket többé-kevésbé objektív módon is **mérni**, [...] -7
- f. **Át** kellett ezt a régi épületet **építeni**, azóta is folynak még a munkálatok. -6
- g. Hát **fel** lehet ezt józan ésszel **fogni**? -5
- h. [...] innen a cégtől már kifizelték, és **elő** kéne már otthon **fizetnem**, [...] -4
- i. Nekem nagyon kellemetlen, hogy állandóan **meg** kell önöket **szakítanom**. -3
- j. **Tönkre** akarja **tenni** a bútoraimat? -2
- k. Közben már **húzni** kezdi **be** a zsalut, aztán beteszi az ablakot is. +2

Az első esetet már többen leírták és magyarázták, legelőször Kálmán C. et al. (1989), akik rámutatnak, hogy bár az igemódosító az infinitívussal alkot szótári egységet, szórendi és prozódiai szempontból szorosabban kapcsolódik a segédige-szerű elemhez. Az én méréseim alátámasztják ezt a megfigyelést, és kiegészítik azzal, hogy ha az 'igekötő + finit ige' egység elhangzott, az infinitívuszt akár az összes bővítmény is megelőzheti a további mondatszakaszban – minthogy a magyarban a finit igét követő tartományban a szórend kötetlen (É. Kiss 2007).

A második eset kapcsán azt figyelhetjük meg, hogy ez mindig nem-semleges mondat típust, azon belül pedig progresszív mondatokban jelenik meg, tehát amelyek éppen zajló eseményt jelölnek. A beférkőző ige szinte kizárólag a *kezd*.

A beférkőzés kapcsán még arra szeretném felhívni a figyelmet, hogy a segédige-szerű elem határozói igenév is lehet. Erre összesen 183 példát találtam a korpuszban, néhányat a (27) példasor szemléltet. Érdekes az is, hogy ezek legnagyobb része (166 db) -vÁn képzős igenév, pedig ez sokkal ritkább, mint a -vA képzős párja. Megfigyelhetjük, hogy ilyen esetekben az igekötő a határozói igenév előtti pozíciót foglalja el, de az infinitívusztól összességében nem kerül olyan messze, mint finit igék beférkőzésekor. Ez talán azzal függhet össze, hogy a határozói igenév nem időjeles.

- (27)
- a. [...] a látszatát is **el** akarván **kerülni** annak, hogy beavatkozik amerikai belügyekbe.
  - b. [...], **el** kívánván **rettenteni** Milosevics jugoszláv elnököt attól, hogy [...]
  - c. Aki gyámságról beszél, az a kritikát utasítja el kapásból, **meg** sem akarva **hallani** azt.
  - d. [...], **rá** próbálva **ruházni** a felelősséget a legkülönbözőbb nehézségekért .

Végül két olyan szerkezetet mutatok be, amelyek az eddig tárgyalt méréseknek nem képezték részét, de mindenképp említést érdemelnek. Az első esetben azt látjuk, hogy az igekötő infinitívusz mellől is kiemelhető fölérendelt tagmondatba. Ez a következő mintázattípussal írható le: ‘igekötő + *kell/kéne/kellene, hogy tud* (kötőmódban) + infinitívusz’. Összesen 25 példa van erre, néhányhoz ld. a (28) példasort.

- (28) a. [...] csinálhat amit akar, te akkor is **át** kell, hogy tudd **ütni** a védelmét.  
b. Egyébként is, mindent **meg** kell, hogy tudjon **oldani**, hiszen orvos.  
c. [...] ez egy infrastruktúra, amihez a mai világban **hozzá** kell, hogy tudjunk **férni**.  
d. Valóban, hát **el** kéne, hogy tudjunk odáig **jutni**, hogy [...]

A másik esetben egy elliptikus szerkezetet látunk a következő mintázattípussal: ‘topik-szerepű infinitívusz + az infinitívusz igekötője + finit ige’. Tehát az infinitívuszt másodjára már nem ejtjük ki, csak az igekötőjét. Az MNSZ2-ben összesen 106 példát találunk erre a szerkezetre. Ebből 50-ben szorosan követi egymást a három elem, 56-ban pedig áll egy szó a topik-szerepű infinitívusz után, ld. a (29) példasort.

- (29) a. De ez nem keseríti meg az életemet, mert hát **megbocsájtani** mindig **meg lehet**.  
b. [...], és az a válasz, hogy **kimenni** csak **ki tudnánk**, de hogy megyünk vissza?  
c. Leírni nem, vagy csak nehezen, de **eljátszani** valószínű **el tudnám**.  
d. Megsimogatni nem lehet, de **megcsodálni** mindenféleképpen **meg lehet**.

#### 4.1.4. Határozói igenév

A határozói igenév képzője a -vA és a -vÁn. Az utóbbi az archaikusabb, és a mai nyelvhasználatban már lényegesen ritkább (a találatoknak mindössze 1,81%-át teszi ki). A két képzőnek nem pusztán a gyakorisága tér el, hanem a használati körük is, bővebben ld. Sárík (1998).

Akárcsak az infinitívusznál, úgy a határozói igenévnél is indokoltnak tartom két szerkezettípus elkülönítését. Az egyik esetben állapot- vagy módhatározóként áll önmagában (pl. *kihajolva*

*integetett*), a másikban a létigével együtt passzív szerkezetet alkot (pl. *meg van oldva*). Látni fogjuk, hogy a határozói igenév igekötője is hátravethető, de csak akkor, ha az igenév állapot- vagy módhatározói szerepet tölt be. A passzív konstrukcióban az igekötő mindig preverbális.

A két típus elkülönítését indokolja az is, hogy a -vÁn képzős igenevek passzív szerkezetben nem állnak. Ezzel ellentmondani látszó példából mindössze kettő volt a korpuszban, de ezek figyelmen kívül hagyhatók, ugyanis 1500-as és 1600-as évekből származó mondatok idézései mai magyar szövegen belül.

A 22. táblázat azt mutatja be, hogy hogyan alakul az igekötők mondatbeli eloszlása az első esetben, tehát amikor állapot- vagy módhatározói szerepű igenévhez kapcsolódnak.

<b>IK-csoport</b>	<b>-3</b>	<b>-2</b>	<b>0</b>	<b>+1</b>	<b>+2</b>	<b>+3</b>
prototipikus	33	1 470	338 797	7 434	45	1
centrális	9	701	141 824	2 201	38	2
félperiférikus	2	50	16 252	348	7	1
periférikus	1	96	17 913	581	18	-
<b>összes db</b>	<b>45</b>	<b>2 317</b>	<b>514 786</b>	<b>10 564</b>	<b>108</b>	<b>4</b>
prototipikus	0,01	0,42	97,41	2,14	0,01	-
centrális	0,01	0,48	97,96	1,52	0,03	-
félperiférikus	0,01	0,30	97,55	2,09	0,04	0,01
periférikus	0,01	0,52	96,26	3,12	0,10	-
<b>összes %</b>	<b>0,01</b>	<b>0,44</b>	<b>97,53</b>	<b>2</b>	<b>0,02</b>	<b>0</b>

**22. táblázat.** Lehetséges igekötő pozíciók az állapot- vagy módhatározói szerepű igenév esetében. Az adatok előbb a négy igekötő-csoport szerinti bontásban láthatók, majd – vastagon szedve – csoportra bontás nélkül. A fejléc alatti öt sor abszolút, a legelső öt sor relatív gyakoriságokat tartalmaz.

A (30) példaszor egy-egy mondattal illusztrál minden pozíciót.

- (30) a. Ez utóbbi törekvése **ki** ugyan nem **mondva**, sőt el is leplezve, de [...] -3
- b. Mindezt otthonról **ki** sem **mozdulva**, a számítógépünkön végezzük. -2
- c. Utólag **visszatekintve** ezt a középiskolások is tudják. 0
- d. Mese egy nőről, aki ablakvilágba zárta magát, onnan **szórva szét** üzeneteit [...] +1
- e. [...] gondozta titokban a sírt, állandó virágtakaróval **fedve** azt **be**. +2
- f. [...] ezzel **szoktatva** szépen lassan **át** az embereket az új rendszerre. +3

Ennél ritkább eset az, amikor a határozói igenév a létigével együtt passzív szerkezetben áll. A jelenség okát valószínűleg nyelvművelői törekvésekben kell keresnünk. Honti és H. Varga (2006), valamint H. Varga (2015) arról számolnak be, hogy a ‘létige + -vA képzős igenév’ típusú passzív szerkezet a 19. század második felétől kezdve „üldözött”, mert a kor nyelvművelői tévesen germanizmusnak ítélték, a közoktatásban tiltották, és irodalmi körökben is igyekeztek visszaszorítani. Az igekötők disztribúcióját a 23. táblázat szemlélteti.

<b>IK-csoport</b>	<b>-9</b>	<b>-8</b>	<b>-7</b>	<b>-6</b>	<b>-5</b>	<b>-4</b>	<b>-3</b>	<b>-2</b>	<b>0</b>
prototipikus	1	1	7	43	140	612	5 865	72 429	94 722
centrális	-	2	2	8	42	415	740	12 329	13 188
félperiférikus	1	-	1	-	-	31	94	1 720	1 367
periférikus	-	-	-	-	-	12	64	891	1 696
<b>összes db</b>	<b>2</b>	<b>3</b>	<b>10</b>	<b>51</b>	<b>182</b>	<b>1 070</b>	<b>6 763</b>	<b>87 369</b>	<b>110 973</b>
prototipikus	-	-	-	0,02	0,08	0,35	3,37	41,67	54,49
centrális	-	0,01	0,01	0,03	0,16	1,55	2,77	46,13	49,35
félperiférikus	0,03	-	0,03	-	-	0,96	2,92	53,52	42,53
periférikus	-	-	-	-	-	0,45	2,4	33,46	63,69
<b>összes %</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0,02</b>	<b>0,09</b>	<b>0,52</b>	<b>3,28</b>	<b>42,33</b>	<b>53,76</b>

**23. táblázat.** Lehetséges igekötő pozíciók a passzív szerkezetben megjelenő igenév esetében. Az adatok előbb a négy igekötő-csoport szerinti bontásban láthatók, majd – vastagon szedve – csoportra bontás nélkül. A fejléc alatti öt sor abszolút, a legalsó öt sor relatív gyakoriságokat tartalmaz.

Itt a finit igékhez viszonyítva szinte ellentétes tendenciát látunk az igekötők eltávolodásában: a periférikus igekötők közel maradnak az igenévhez, és a félperiférikusak között is csak elvétve van egy-egy eltávolodó példa, vannak lyukas pozíciók. Leginkább a centrális igekötőkre jellemző az eltávolodás, de prototipikusaknál sem ritka a jelenség. Talán arról lehet szó, hogy a (fél)periférikus igekötők kevés igével állnak, és ezek is olyan igék, amelyek szemantikai okokból nem vagy ritkán kerülnek passzív szerkezetbe. A (31) példasor egy-egy mondattal illusztrál minden pozíciót.

- (31) a. Ez azért dühítő, mert az USÁ-ban **tele** van a közoktatás a bölcsitől kezdve olyan szakemberekkel **szórva**, akik [...] -9
- b. [...], hát teljesen **össze** van a keze a mai napig is **nyomorodva**, [...] -8
- c. [...] kirámoltak ugyan, csakumpak, de ezzel **el** is van a gond egy életre **intézve**! -7
- d. **Ki** vannak ezek a marketinges dolgok **találva**. -6



- e. **Meg** van az a hapsi **buggyanva**. -5
- f. **Meg** vannak ezek egyáltalán **fogalmazva**? -4
- g. Amerikában ezzel **el** is van **intézve** a dolog, Európában kicsit más a helyzet. -3
- h. A kocka **el** van **vetve**. -2
- i. Miért nincs minden **megoldva**? 0

A passzív szerkezetben álló határozói igenév szintaktikai viselkedése az infinitívussal mutat rokonságot. Itt is megfigyelhető az, hogy az igekötő a finit ige előtti igemódosítói pozícióba törekszik, és a finit igével alkotott prozódiai egysége fontosabb, mint az igenévvel alkotott szótári egysége. Minden nem közvetlenül preverbális pozícióra igaz, hogy a létige beférkőzik a szerkezetbe, néha egy másik, segédige-szerű elemmel együtt (pl. *el kell lennie küldve*).

A passzív határozói igenév esetében is lehetséges az a szerkezet, amelynél az igenév igekötőjét időjeles mellékmondatból fölérendelt tagmondatba emeljük ki, bár ez elég ritka, mindössze 14 példa van rá. A mintázattípus így egyszerűsíthető le: 'igekötő + *kell/kellene/kéne*, *hogy legyen* + határozói igenév', ld. a (32) példasort. Az utolsó két példa tanúsága szerint az igenevet egy rövid szó is megelőzheti a szerkezetben.

- (32) a. De persze ehhez **ki** kell, hogy legyen **alakítva** egy megfelelő munkakultúra.
- b. Feltétlenül **össze** kell, hogy legyen **kötve**, ez egy hatalmas fegyelmező erő.
- c. [...] legyen azonosító adata, **fel** kell, hogy legyen rajta **tüntetve** [...]
- d. [...], továbbá az összekötő telefon vonal is **ki** kell, hogy legyen már **építve**, [...]

#### 4.1.5. Melléknévi igenév

A különféle igeneves szerkezetek közül a melléknévi igenevek azok, amelyek kapcsán a legnagyobb nézetkülönbségeket találjuk a szakirodalomban. Ezek jelentős része szófaji vita, amely abból a kérdésből indul ki, hogy minek tekintsük a melléknévi igenevet: igéből képzett melléknévnek, non-finit igének vagy önálló szófajnak? Ha az utóbbi két lehetőség egyikét fogadjuk el, akkor szükséges feltételeznünk, hogy a melléknévi igenév melléknévvé tud válni. Ez vezet el egy

újabb szófaji kérdéshez: Hogyan állapítható meg, hogy egy adott mondatban a kérdéses szó melléknévi igenév vagy melléknév? A problémakör részletes tárgyalásához ld. Kenesei (2000) és Komlósy (1992) munkáit.

A szófaj-problémához kötődnek a melléknévi igenevek végződéseit érintő viták. Ha a melléknévi igenevet önálló szófajnak vagy melléknévnek tekintjük, akkor a toldaléka csak képző lehet, míg ha non-finit igeiként elemezzük, a toldalék lehet inflexió is, bővebben ld. Kiefer (2003a). Ebben a kérdésben külön figyelmet kap a -hAtÓ státusza. Kiefer (2003a) amellett érvel, hogy ez melléknévképző, és a -hAtÓ toldalékos szavak az MNSZ2 korpuszban is melléknévként annotáltak. Ahogy a későbbiekben látni fogjuk, ez a megoldás gyakorlati szempontból előnytelen, ugyanis a -hAtÓ végződésű igei származékok a szintaktikai viselkedésük tekintetében nagyon hasonlítanak a finit igeikre. Az igeekötőjük hátravetett is lehet, és messzire eltávolodhat az igei származéktól. Mindezt nehéz megmagyarázni – valamint a korpuszban megtalálni – akkor, ha ezeket a szavakat melléknévként címkézzük fel.

Bőséges szakirodalma van a melléknévi igenevek alcsoportjait érintő kérdéseknek is. Ezek a rendezőelv-viták arra mutatnak rá, hogy a melléknévi igenevek hagyományos felosztása – folyamatos, befejezett, beálló – túláltalánosít, bővebben ld. Kiefer (1992) és Laczkó (2000). Én úgy kerülöm meg ezt a problémát, hogy a különféle igenevekre nem a hagyományos terminusokkal, hanem egyszerűen a végződéseikkel hivatkozok.

Bár a dolgozatban az igeekötők mondattani viselkedésével foglalkozok részletesen, az eredményeim az igenevekkel kapcsolatos, fentebb említett kérdések tisztázásához is hozzájárulhatnak.

A fejezet további részében eltekintek az elméleti vitáktól, és rátérek a korpuszadatok ismertetésére. Ezek alapján meglehetősen unortodox módon fogom tárgyalni a melléknévi igenevek témáját. Először a -hAtÓ, majd az -AndÓ végződésű igenevekről lesz szó. Azért indokolt ezt a két típust egymáshoz közel bemutatni, mert szintaktikai szempontból nagyon hasonlóan viselkednek, ugyanis az igeekötőjük hátravetett is lehet, ha állítmányi helyzetbe kerülnek (pl. *aligha oldható meg, nem tévesztendő össze*). Ezután az -Ó és -(O)(t) végződésű igeneveket tárgyalom – ezeknek az igeekötője csak preverbálisan válhat el (pl. *el nem jövő korszak, vissza sem fizetett tartozás*). Végül olyan szerkezetekről lesz szó, amelyeknél ezek az igenevek egyértelműen melléknévi, illetve határozószói képzőt kapnak, és az igeekötő mégis elválhat tőlük.

Az első vizsgált típus, a **-hAtÓ végződésű igenév** tehát állhat és gyakran is áll hátravetett igeekötő-

vel, amikor állítmányi helyzetbe kerül. A 24. táblázat tanúsága szerint az igekötő egészen messzire kerülhet az igenévtől. A (33)-ban minden igekötő-pozícióhoz láthatunk egy példát.

IK-csoport	-4	-3	-2	0	+1	+2	+3	+4	+5
prototipikus	-	27	5 387	428 358	94 442	1 102	22	2	1
centrális	3	20	2 862	52 804	15 446	232	8	1	1
félperiférikus	2	132	588	5 597	1 449	40	4	1	-
periférikus	-	7	120	20 117	2 597	41	2	-	-
<b>összes db</b>	<b>5</b>	<b>186</b>	<b>8 957</b>	<b>506 876</b>	<b>113 934</b>	<b>1 415</b>	<b>36</b>	<b>4</b>	<b>2</b>
prototipikus	-	0,01	1,02	80,92	17,84	0,21	-	-	-
centrális	-	0,03	4,01	73,98	21,64	0,33	0,01	-	-
félperiférikus	0,03	1,69	7,53	71,64	18,55	0,51	0,05	0,01	-
periférikus	-	0,03	0,52	87,91	11,35	0,18	0,01	-	-
<b>összes %</b>	<b>0</b>	<b>0,03</b>	<b>1,42</b>	<b>80,28</b>	<b>18,04</b>	<b>0,22</b>	<b>0,01</b>	<b>0</b>	<b>0</b>

**24. táblázat.** Lehetséges igekötő pozíciók a -hAtÓ végződésű melléknévi igenév esetében. Az adatok előbb a négy igekötő-csoport szerinti bontásban láthatók, majd – vastagon szedve – csoportra bontás nélkül. A fejléc alatti öt sor abszolút, a legalsó öt sor relatív gyakoriságokat tartalmaz.

- (33) a. [...], keseregnek még ma is sokan a **vissza** már soha nem **fordítható** jelenségen. -4
- b. Így érthető, ha **el** nem is **fogadható**, hogy [...] -3
- c. Informatikával foglalkozni érdekes, soha **meg** nem **unható** foglalatosság. -2
- d. Ha a vizsgálat jellege szükségessé teszi, több szakértő is **kirendelhető**. 0
- e. A hidegneutronokkal sokféle elemzés **végezhető el**. +1
- f. Mennyi idő alatt **valósítható** ez **meg**? +2
- g. [...] ha van is, az mindig akkor **vonható** már csak **le**, amikor semmi értelme. +3
- h. [...], hogy gyors és nem **keverhető** már olyan könnyen **le** egy sporteseményen. +4
- i. Egyetértek még az eljárásban is, két szinten **indítható** szerintem ez a munka **el**: [...] +5

Ha az igekötő preverbálisan elvált, akkor rendszerint tagadó szerkezetet látunk, ritkábban *is*-est. Posztverbális igekötő esetében többnyire rövid szavak állnak az igenév és az igekötő között. Figyelemre méltó az is, hogy a -hAtÓ végződésű igenév nem csak az igekötők disztribúcióját, hanem a szerkezetbe gyakran beférkőző szavakat tekintve is hasonlít a finit ígére (ld. a 25. táblázatot) –

hátravetett igekötője viszont csak akkor lehet, ha nincs finit ige a tagmondatban, így a -hAtÓ ennek a helyét foglalja el (Nádasdy 2001).

pozíció	köztes szavak
-3	már nem, nem is, még nem
-2	nem, is, sem
+2	majd, csak, ez, -e, még, az, már, itt, is, leginkább, jól, most, úgy, ezek, vele, ugyan, egyértelműen, csupán, talán, mindez, pontosan, újra

**25. táblázat.** Az igekötő és a -hAtÓ végződésű melléknévi igenév közti szavak, pozíciók szerinti bontásban. A táblázat csak a legalább ötször előforduló találatokat tartalmazza.

Az **-AndÓ végződésű igeneveket** a hagyományos nyelvtanok beálló melléknévi igenévnek nevezik. Komlósy (1992: 363) nem tartja szerencsésnek ezt az elnevezést, mert ezek elsősorban szükségességet fejeznek ki, nem annyira jövőt vagy utóidejűséget. A szakirodalmak rendszerint csak a -hAtÓ végű igeneveknél említik azt, hogy hátravetett igekötők lehetnek, az -AndÓ végűeknél ez nem merül fel. Laczkó (2000: 441) ki is mondja, hogy ezek az igenevek csak jelzői szerepűek lehetnek, ezért nem állhatnak hátravetett igekötővel. A korpuszadatok ezt egyértelműen cáfolják, ld. a 26. táblázatot. Az kétségtelenül igaz, hogy az -AndÓ esetében a posztverbális igekötők nem olyan gyakoriak, mint a -hAtÓ esetében, de nem is elhanyagolhatóan ritkák.

IK-csoport	-3	-2	0	+1	+2
prototipikus	-	87	72 633	774	7
centrális	-	2 980	5 923	699	1
félperiférikus	1	5	645	7	1
periférikus	-	9	4 471	30	-
<b>összes db</b>	<b>1</b>	<b>3 081</b>	<b>83 672</b>	<b>1 510</b>	<b>9</b>
prototipikus	-	0,12	98,82	1,05	0,01
centrális	-	31,03	61,68	7,28	0,01
félperiférikus	0,15	0,76	97,88	1,06	0,15
periférikus	-	0,20	99,14	0,67	-
<b>összes %</b>	<b>0</b>	<b>3,49</b>	<b>94,79</b>	<b>1,71</b>	<b>0,01</b>

**26. táblázat.** Lehetséges igekötő pozíciók az -AndÓ végződésű melléknévi igenév esetében. Az adatok előbb a négy igekötő-csoport szerinti bontásban láthatók, majd – vastagon szedve – csoportra bontás nélkül. A fejléc alatti öt sor abszolút, a legalsó öt sor relatív gyakoriságokat tartalmaz.

A centrális igekötők esetében első ránézésre nagyon furcsa számokat látunk: az előfordulásaik közel kétharmadában (31,03%-ban) -2 pozícióban állnak, ami komoly eltérés a másik három csoport szabályosnak mondható, egymásra nagyjából hasonlító mintázataitól. Erre az a magyarázat, hogy a centrális *vissza* igekötő nagyon sokszor megjelenik a következő kifejezésben: *vissza nem térítendő* (támogatás, hitel, kölcsön, stb.). Tehát tulajdonképpen ennek az egy igekötőnek a kollokációs viszonyai húzzák el az adott csoport egészét. Ugyanez történik a +1 pozíció esetében is, csak itt az *össze* igekötővel: a *nem tévesztendő össze* és a *nem keverendő össze* kifejezések gyakorisága eredményezi a centrális igekötőknél látható 7,28%-os értéket.

A (34) példaszor egy-egy mondattal illusztrál minden pozíciót. Egyértelműen látszik, hogy a tárgyalt igenév állítmányként is szerepelhet a mondatban, hátravetett igekötővel.

- (34) a. [...] **tovább** már nem **bontandó** elemi műveletek vagy elemi utasítások [...] -3  
 b. Ehhez ad az állam **vissza** nem **térítendő** támogatást. -2  
 c. Ennyi potenciális, **leküzdendő** ellenség nincs. 0  
 d. [...] egy vélemény nem **keverendő össze** az igazsággal a frankótutival. +1  
 e. [...] a muzeális tárgyak védelmével foglalkozó rész **emelendő** még **ki**, [...] +2

A szerkezetben az igekötőt és az igenevet elválasztó szavak mindannyian egy szótagúak, többségük klitikum-szerű. A -3 pozícióban a *már nem*, -2-ben *nem*, *is*, *sem*, +2 pozícióban pedig az *-e*, *még*, *csak* szavak fordulnak elő beférkőző elemként.

Az **-Ó végződésű melléknévi igenévi kifejezésben** az igekötő elválhat az igenévtől, de csak preverbálisan. Még fókusz vagy tagadás esetén sem lehet hátravetett. Ennek az az oka, hogy az -Ó végződésű igenevek valóban csak jelzői helyzetben állnak, és ekkor a szerkezet fejező, tehát utolsó tagja az igenév. Az igekötők disztribúcióját a 27. táblázat szemlélteti.

Amint látjuk, ez egy nagyon megkötő szerkezet, az esetek 98,34%-ában az igekötő nem válik el az igenévtől. A centrális igekötőknél figyelhető meg a legnagyobb hajlandóság az elválásra, de valószínűleg ennek is egy-egy kollokációs viszony az oka. A -2 pozícióhoz tartozó tagmondatokat vizsgálva az derül ki, hogy a leggyakoribbak között vannak a következő, centrális igekötőt tartalmazó szerkezetek: *oda nem illő*, *vissza nem térő*, *össze nem függő*, *össze nem illő*. Ezek szinte frazéma-szerűek. A (35) példaszor egy-egy mondattal illusztrálja az igekötő három pozícióját.

IK-csoport	-3	-2	0
prototipikus	-	4 375	1 402 104
centrális	3	3 130	289 834
félperiférikus	-	114	36 185
periférikus	1	579	88 131
<b>összes db</b>	<b>4</b>	<b>8 198</b>	<b>1 816 254</b>
prototipikus	-	0,31	99,69
centrális	-	1,07	98,93
félperiférikus	-	0,31	99,69
periférikus	-	0,65	99,34
<b>összes %</b>	<b>0</b>	<b>1,66</b>	<b>98,34</b>

**27. táblázat.** Lehetséges igekötő pozíciók az -Ó végződésű melléknévi igenév esetében. Az adatok előbb a négy igekötő-csoport szerinti bontásban láthatók, majd – vastagon szedve – csoportra bontás nélkül. A fejléc alatti öt sor abszolút, a legalsó öt sor relatív gyakoriságokat tartalmaz.

- (35) a. [...] a Franciaországban élő, de **haza** soha nem **térő** Albert Dezsővel [...] -3
- b. Ha pedig egymással **össze** nem **függő** különféle jogszabályokat gyúrunk egybe, [...] -2
- c. A madarak beszéde éjszaka, hihetetlen, **megindító**, szinte emberi beszéd. 0

A beférkőző szavak -2 pozíciónál a *nem, is, sem*, ebben a gyakorisági sorrendben. A -3 pozíciónál pedig a *soha sem, talán nem, ugyan nem, soha nem*.

Az **-(O)(t)t végződésű igenevek** szintaktikailag ugyanúgy viselkednek, mint az -Ó végűek. Az igekötő csak preverbálisan válhat el, de az esetek 99,29%-ában közvetlenül az igenév előtt marad. Tehát ilyen szempontból az **-(O)(t)t** képzős melléknévi igeneveknek már jóval gyengébbek az igei, és erősebbek a melléknévi tulajdonságai.

Az igekötők disztribúcióját a 28. táblázat mutatja be. Ezek a gyakorisági adatok fenntartással kezelendők, mert az **-(O)(t)t** végződésű igenév sokszor keveredik a múlt idejű finit igével az automatikus morfológiai elemzés és szófaji egyértelműsítés során. Sok esetben csak a kontextust ismerve lehet elkülöníteni a kettőt, amire ekkora, milliós nagyságrendű adat esetében nem lehet vállalkozni.

A 28. táblázat adatai alapján olyan tendencia rajzolódik ki, hogy minél prototipikusabb az igekötő, annál könnyebben el tud válni az igenévtől. Valójában nem egészen erről van szó. A -2 pozíciós találatokat alaposan megvizsgálva inkább azt mondhatjuk, hogy minél prototipikusabb egy

<b>IK-csoport</b>	<b>-2</b>	<b>0</b>
prototipikus	19 927	2 614 069
centrális	1 551	304 835
félperiférikus	153	50 030
periférikus	178	95 755
<b>összes db</b>	<b>21 809</b>	<b>3 064 689</b>
prototipikus	0,76	99,24
centrális	0,51	99,49
félperiférikus	0,30	99,70
periférikus	0,19	99,81
<b>összes %</b>	<b>0,71</b>	<b>99,29</b>

**28. táblázat.** Lehetséges igekötő pozíciók az -(O)(t)t végződésű melléknévi igenév esetében. Az adatok előbb a négy igekötő-csoport szerinti bontásban láthatók, majd – vastagon szedve – csoportra bontás nélkül. A fejléc alatti öt sor abszolút, a legalsó öt sor relatív gyakoriságokat tartalmaz.

igekötő, annál több és gyakoribb kollokációs viszonya van, amely megszakított tagadó szórenddel rögzült. Például kiugróan gyakori a *meg nem nevezett* (forrás, személy), *meg nem erősített* (hír, információ), (18. életévét) *be nem töltött* (személy), *fel nem tett* (kérdés), *be nem tartott* (ígéret).

A (36) példasor egy-egy mondattal illusztrálja az igekötő két lehetséges pozícióját. Köztes szóként csak a *nem*, *is*, *sem* fordulnak elő, ebben a gyakorisági sorrendben.

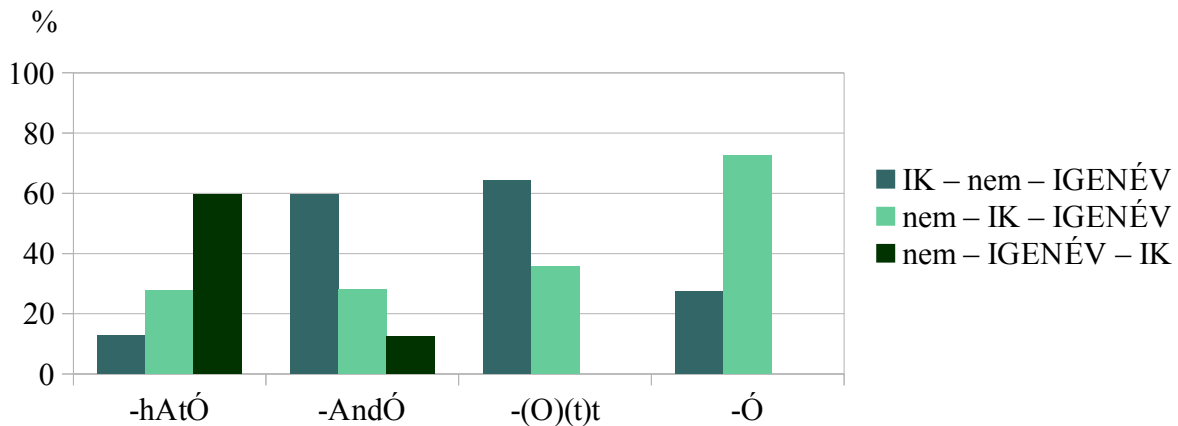
- (36) a. Az előző kormány által előírt és a kormányváltásig **be is tartott** menetrend szerint. -2
- b. Elég egy **elhamarkodott**, egy hirtelen gesztus, és mindenki félreérti. 0

Érdemes még kitérni arra a kérdésre, hogy **tagadáskor milyen szórendi mintázatokat látunk** az igekötős melléknévi igenevek esetében. Arról már esett szó korábban, hogy hátravetett igekötője csak a -hAtÓ és -AndÓ végződésű igeneveknek lehet, és csakis akkor, ha ezek állítmányi helyzetben állnak – állítmányoknál mind a három szórend előfordul, ld. a (37) példákat.

- (37) a. Ez az indoklás **el nem fogadható**.
- b. Ez az indoklás *nem elfogadható*.
- c. Ez az indoklás *nem fogadható el*.

Jelzői helyzetben minden melléknévi igenév-típus megjelenik, és ekkor is ingadozó szórendet figyelhetünk meg: a tagadószó beférkőzhet a szerkezetbe, de ez nem szükségszerű. Kemény (2018)

arról számol be, hogy a megszakított szórend (tehát ‘igekötő – tagadószó – igenév’) a régebbi – és a nyelvművelők körében elfogadottabb –, az új tendencia az, hogy az igekötő nem válik el az igenévtől. Ez azzal függhet össze, hogy minél inkább melléknévnek érzi a beszélő az adott szót, annál kevésbé választja le róla az igekötőt.



8. ábra. A lehetséges szórendek arányai a négy melléknévi igenév-típus tagadásakor.

A 8. ábra azt mutatja, hogy igenév-típusonként jelentős eltérések vannak a különféle tagadó szórendek arányaiban. Ennek elsősorban nem szintaktikai okai vannak, hanem – ahogy korábban is láttuk – vannak gyakori szókapcsolatként rögzült szerkezetek, amelyeknek a tagadó szórendje is rögzült (pl. a *be nem tartott ígéret* bevett szófordulat, szemben a *nem betartott ígéret*-tel).

Az eddig tárgyalt melléknévi igenevek bármelyik típusából képezhetünk **határozószót** az -An képzővel. Az igekötő ezekben a szerkezetekben is el tud válni az igei alaptagtól, tipikusan megszakított szórendű tagadással. Erre 261 db példát találtam (249 *nem*, 12 *sem*). Ezen kívül két esetben az *is* szócska férkőzött a szerkezetbe. A (38)-ban látható egy-egy példa a három beférkőző elemre.

- (38) a. Az egyház **meg** nem **szűnően** vallja ezt mindennapi liturgiája során, [...]
- b. [...] a bogárdi nagymama nevét vitte volna idő előtt, még **meg** sem **születetten** sírba?
- c. GPRS, EDGE és UMTS is van, egy csúcskategóriás mobiltól **el** is **várhatóan**.

Az -Ó végződésű melléknévi igenevekből -s képzővel **melléknév** képezhető (pl. *okoskodós*). Ezeknél is megfigyelhető a megszakított szórend tagadás esetén, bár ritka jelenség: 9 db példát találtam rá, és ebből 4 db azonos forrásból – valószínűleg egy szerzőtől – származik, ld. a (39) példasort.



- (39) a. [...], egymásra borulás, **el** nem **szakadós**, szoros női barátság.  
 b. És e miatt az eset miatt alakult ki nála ez az **el** nem **engedős** ragaszkodás?  
 c. Aztán jött a „**be** nem **indulós** Trabant”, február 9-én pedig az erzsébeti Trabant.

A megszakított szórend az -(A)tlan képző használatakor is érvényre tud jutni. Erre 8 db példát találtam, ezek közül 2-ben az *is*, 6-ban a *nem* férkőzik be a szerkezetbe. Az előbbi, *is*-es esetek elég szokatlanul hangzanak, de jól átfordíthatóak gyakoribb szerkezetekbe, pl. *meg is valósíthatatlan* → *meg sem valósítható*, ld. a (40) példasort.

- (40) a. [...] és talán **meg is valósíthatatlan** társadalmi újításoktól várja minden üdvét, [...]  
 b. [...] elég komoly elmaradásban vagyunk, és valahol **el is fogadhatatlan**, hogy [...]

A következő, *nem*-es példákban viszont hibázásokat látunk – (41) példasor –, egyszerűen nem a szándékolt jelentés jön ki a szerkezetből. Arról van szó, hogy a két tagadás kioltja egymást: *le nem záratlan* → *lezárt*. Viszont a szerző nyilván arra gondolt, hogy *lezáratlan* vagy *le nem zárt*, ezek már ugyanazt fejezik ki.

- (41) a. [...] ebben a bizonyos **le nem záratlan** peres ügyünkben is [...]  
 b. Elmosolyodtam, mikor felismertem a **le nem tagadhatatlan** hasonlóságait.  
 c. [...] akkor szinte **be nem hozhatatlan** versenyelőnyre tett volna szert [...]

#### 4.1.6. Deverbális főnév

A deverbális főneveket igéből vagy igenévből képezzük. Produktív képzőjük az -*As*, amely igékhez kapcsolható (pl. *megbecsül* → *megbecsülés*), valamint a -*sÁg*, amely igenévhez járul (pl. *felkészült* → *felkészültség*). Az így létrejött főnevek több szempontból igei tulajdonságokat őriznek, ezek közé tartozik az is, hogy az igekötő el tud válni tőlük.<sup>35</sup> A továbbiakban csak olyan példákkal foglalkozom, amelyekben az igekötő elvált.

<sup>35</sup> Szabolcsi és Laczkó (1992) az -*As* képzős főneveket két csoportra osztja a szintaktikai viselkedésük alapján. Az egyik csoport a „tény, cselekvés” jelentésű, komplex eseményt leíró főneveké, amelyek az ige teljes vonzatkeretét megtartják, a másik az „eredmény” jelentésűeké, amelyeknek nincsenek kötelező vonzataik, így kevésbé igei jellegűek. Gyanítható, hogy az igekötő elválása inkább az igei tulajdonságokat jobban őrző, első csoport esetében jelenik meg. Ennek a bizonyítása megkívánná a korpuszadat szemantikai elemzését is, amelyre a nagy elemzendő adatmennyiség és az automatikus módszer hiánya miatt nem vállalkoztam.

Az igekötő nem lehet hátravetett, és preverbálisan is csak megszakított tagadó szórend esetén válhat el. Az utóbbi jelenség igekötő-csoportok közötti megoszlását a 29. táblázat szemlélteti. Szembetűnő, hogy a félperiférikus igekötőkkel bezárólag csökkenni látszik a szerkezet előfordulási gyakorisága, de a periférikusaknál jelentősen megugrik – holott intuitívan inkább folyamatos csökkenést várnánk. Ennek akkor találjuk meg a magyarázatát, ha igekötők szerinti bontásban is vizsgáljuk az adatot. Így az látszik, hogy a periférikus igekötők közül csak néhánynak a kollokációs viszonyai miatt magas itt a tagadásakor elvált igekötők aránya. Lexikalizált eseteket látunk, egy-egy igével, pl. *egyét nem értés, együtt nem működés, viszont nem látás, végre nem hajtás*. Ezzel szemben a prototipikus igekötők csoportja sokféle igével alkot ilyen szerkezetet, tehát ezek az esetek valóban produktívak.

IK-csoport	db	%
prototipikus	4 207	78,64
centrális	897	16,77
félperiférikus	18	0,33
periférikus	228	4,26
<b>összesen</b>	<b>5 350</b>	<b>100</b>

**29. táblázat.** Deverbális főnév tagadásakor elvált igekötők abszolút (*db*) és relatív (%) gyakorisága.

A szerkezetbe férköző szavak csak a *nem* és a *sem* – ezt előrevetítette a 3.3.3. fejezet –, ld. a (42) példasort.

- (42) a. Hol jelentkezik ez a **meg nem értettség**?
- b. Az állami adósság **ki nem fizetése** gerjeszti a folyamatot!
- c. A többiektől elnézést a **meg sem említésért**, [...]
- d. Kivonulás, **be sem vonulás** - de azért lesz kiutazás.

A korpuszadat tanúsága szerint a megszakított tagadó szórend -tAl képzős deverbális főnevek esetében is megjelenik. Ez a képző már nem produktív, az adatban a következő eseteit látjuk: *vesz* → *vétel*, *tesz* → *tétel*, *hoz* → *hozatal*, *visz* → *vitel*. 81 olyan példa van, ahol az igekötő ezektől elválik. Mindegyik hivatalos, többnyire jogi- vagy sajtószövegekből származik, ld. a (43) példasort.

- (43) a. [...] vagyis a felszólítás **át nem vételével** nem lehet elszabotálni a fizetést [...]
- b. [...] ügyészi perbeli cselekmények megtétele vagy **meg nem tétele** miatt [...]
- c. [...] sérelmezi a pócsi kegykép Bécsbe szállítását és **vissza nem hozatalát**.
- d. [...] elfogadott mennyiséget meghaladó cukor **ki nem vitelének** eseteiről volt szó.

Megfigyelhető az is, hogy bizonyos szintaktikai környezetekben nem lehetséges a megszakított tagadó szórend, ekkor a *nem* a teljes főnévi kifejezés elé kerül. Ez a helyzet áll fenn fókusztagadásnál, ld. a (44) példasort. Ekkor ‘nem – hanem’ típusú szembeállítás van a mondatban, akár kimondva, akár csak odaértve.

- (44) a. [...] amiért valaki megdolgozik, az **nem ellátás**, azért dolgozni kell.
- b. Ott kiadatás volt, és **nem átadás**.
- c. A tragédiában a burleszk keserves feleút, **nem megérkezés**.
- d. [...], hogy **nem megérkezés** van, hanem út van.

Amikor a deverbális főnévi kifejezést úgy tagadjuk, hogy nincs fókuszban, kétféle szórend lehetséges. A tagadószó beférkőzhet a szerkezetbe – erre példák a (42) és a (43) –, vagy állhat a szerkezet előtt, ld. a (45) példákat.

- (45) a. [...] azt a kamatot, amit a tb **nem befizetése** után zsebre tehetett volna.
- b. [...], amely egyelőre a jegyzőkönyv **nem aláírását** eredményezte.
- c. [...] olyan mélységben soha nem volt a cigányságnak az elfogadottsága, vagy a **nem elfogadottsága**, mint most [...]
- d. Tehát e hiedelem betartásának, illetve **nem betartásának** következményeinél [...]

## 4.2. Diakrón vizsgálat: az ómagyar kortól napjainkig

Az igekötőzés, mint az aspektusjelölés alternatív módszere a 12. század végétől adatolható jelenség. Az ómagyar kor végére, tehát a 16. század elejére már stabilan működő rendszernek tekinthető (É. Kiss 2014a: 69-70). Az igekötők állományát ekkor még kevés elem alkotja, lényegében a hat

prototipikus igekötő: *meg, el, fel, ki, be, le*. Ezeknek a szintaktikai viselkedését mutatom be ebben a fejezetben az ó-, közép- és újmagyar kor korpuszai alapján. A kutatás a felsorolt igekötők és a hozzájuk tartozó finit igék egymáshoz viszonyított helyének, távolságának szisztematikus mérésén alapszik. Az eloszlási minták mellett azt is vizsgálom, hogy a tagadó mondatokra milyen arányban jellemző az ‘igekötő – tagadószó – ige’, valamint a ‘tagadószó – ige – igekötő’ szórend az egyes nyelvtörténeti korokban. Ez azért is érdekes kérdés, mert a szinkrón vizsgálat során láttuk, hogy a mai magyar nyelvállapotban sok igei származéktól csak tagadás esetén válhat el az igekötő, mégpedig az elsőként említett, archaikus szórendet őrizve.

Először ismertetem a felhasznált korpuszokat, valamint az adatgyűjtés és adatfeldolgozás módszerét. Ezután következnek a korpuszmérések eredményei, és végül egy olyan vizsgálat, amely a kétféle tagadó szórend arányának változását mutatja be.

#### **4.2.1. A kutatás módszere**

A kutatás során négy korpuszt használtam, amelyek anyaga több mint 800 évet fed le. A kapott eredmények értelmezésekor fontos szem előtt tartani ezeknek a korpuszoknak a jelentős méretbeli különbségeit is.

Az ómagyar kori adatok (896–1526) forrása az Ómagyar Korpusz – a továbbiakban OMK –, amely tartalmazza az összes fennmaradt ómagyar kori és néhány középmagyar kori szövegemléket (Simon és Sass 2012). Mérete 2,2 millió szövegszó. Jelen kutatás nem használja fel a teljes anyagot, csak azt az öt forrást, amelyhez morfológiai elemzés is tartozik. Ezek a következők: *Jókai-kódex* (1440 körülről), *Müncheni kódex* (1466), *Festetics-kódex* (1492–94), *Guary-kódex* (1508 előtről), *Könyvecse az szent apostoloknak méltóságokról* (1521). A kutatás szempontjából releváns korpuszrész mérete 132 ezer szövegszó.

A középmagyar kori (1526–1772) szövegek vizsgálatához a Történeti Magánéleti Korpuszt – TMK – használtam (Dömötör 2009). Itt az élőnyelvhez közel álló írásos anyagok érhetők el: főként boszorkányperek során készült jegyzőkönyvek, valamint misszilisek, elküldött levelek (főúri családi levelezések, jobbagylevelek). 4,34 millió szövegszó alkotja a korpuszt, amely teljes egészében szótövesített és morfológiailag elemzett.

Az újmagyar kori (1772-től napjainkig tartó) szöveganyaghoz egyrészt a Magyar Történeti Szöveggyűjtő – MTSZ – használtam (Ittész 2009), másrészt a teljes disszertáció alapjául szolgáló MNSZ

2.0.4 módosított verzióját (erről bővebben ld. a 2.2.1 fejezetet). Az utóbbi 20-21. századi szövegeket tartalmaz, míg az MTSZ a teljes újmagyar kort igyekszik reprezentálni. Ez a korpusz 30 millió szövegszóból áll. Részlegesen feldolgozott, csak a szóalakok érhetők el benne. A morfológiai elemzés megvalósítása után a teljes korpusz részét képezhette a kutatásnak.

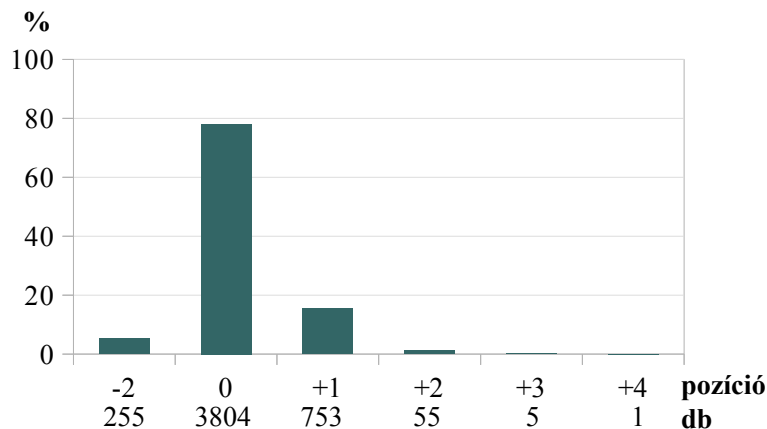
Az adatgyűjtésben a legfontosabb szempont az volt, hogy minden olyan mondat bekerüljön a vizsgált anyagba, amely egy tagmondaton belül tartalmaz egy finit igét és egy hozzá tartozó prototipikus igekötőt, bármilyen alakváltozatban (*meg~mög, fel~föl, be~bé*). Az OMK és TMK esetében viszonylag kevés anyaggal kellett számolni, így ezekben csak a prototipikus igekötőkre kerestem rá, és később saját scriptek segítségével szűrtem ki a finit igét is tartalmazó találatokat. Az MTSZ esetében – ahol csak a nyers szöveg állt rendelkezésre –, le kellett kérni minden olyan szövegrészt, amelyben a vizsgált igekötőkkel azonos írásképi szó szerepel (tehát ezen a szinten pl. a *mögötte* is érvényes találat volt). Csak a kapott találatok automatikus elemzése után lehetett tovább szűrni a mondatokat. Az MNSZ-hez kapcsolódó munkafolyamat a 4.1.1. fejezetben olvasható, itt a továbbiakban csak az OMK-t, TMK-t és MTSZ-t érintő adatfeldolgozási lépéseket ismertetem.

A lekérések után egységes formátumúvá kellett alakítani a három korpusz kimeneteit, hogy később mindet ugyanúgy lehessen kezelni. Az elemzett korpuszok esetében a gyűjtött anyagot olyan XML-be konvertáltam, amelyben egy sor egy mondatnak felel meg, és minden mondatnál elkülönülnek a meta-adatok is, ha elérhetők (pl. a pontos forrás, stílusregiszter). Az MTSZ-nél szükséges volt a betűhű szöveg normalizálása. Ez egyrészt a mai sztenderdtől eltérő karakterek átalakítását jelenti, másrészt a kötőjellel egybeírt szósorok elválasztását indokolt esetben (pl. *meg-betstelenül* → *megbecstelenül*, *meg-sem-is-hallgat* → *meg sem is hallgat*). Ezután lehetett elvégezni a szófaji egyértelműsítést és morfológiai elemzést a magyarlanc 3.0 programmal (Zsibrita et al. 2013).

Minden korpusz esetében automatikus módszerrel szűrtem a tagmondatok halmazát. Csak olyan találatokat őriztem meg, ahol egy tagmondaton belül volt az igekötő és a finit ige. Azt is ellenőrizni kellett, hogy az igekötő valóban a finit igéhez tartozik-e, nem pedig valamilyen igenévhez. Ezek a szűrések már tagmondatok szintjén történtek, de a teljes, eredeti mondatok elérhetőek maradtak az azonosítóik segítségével. A tagmondatokban az igekötő és a finit ige távolságát a 4.1.1. fejezetben bemutatott módszerrel mértem ki.

#### 4.2.2. A prototipikus igekötők pozíció szerinti eloszlása

Az **ómagyar adatok** a már említett öt forrásból származnak. Ez összesen közel 8700 mondat, amelyből 4900 tartalmaz igekötős finit igét. A legnagyobb terjedelmű forrás a Müncheni-kódex, amely a Huszita Biblia – az első magyar nyelvű bibliafordítás – részleteit tartalmazza. Németi György másolta a kódexet a moldvai Tatros településen 1466-ban (Nyíri et al. 1971). Ez a mű azért is érdemel külön említést, mert sok olyan nyelvi konstrukciót tartalmaz, amely a másik négy forrásban nem jelenik meg. A 9. ábra szemlélteti az ómagyar anyagon végzett mérés eredményét. Látható, hogy az igekötők döntő többsége a preverbális tartományban van (83,3%), ennek fő oka az archaikus tagadó szerkezet fölénye az újabb, fordított szórendű tagadással szemben.



9. ábra. Prototipikus igekötők mondatbeli eloszlása az OMK öt elemzett forrásában.

A legszélső preverbális igekötői pozíció a -2, jellemzően tagadó mondatokkal, ld. a (46) példasort. Mindössze három olyan példa van, ahol a beférkőző szó a *sem*, és egy olyan, ahol az *is*.

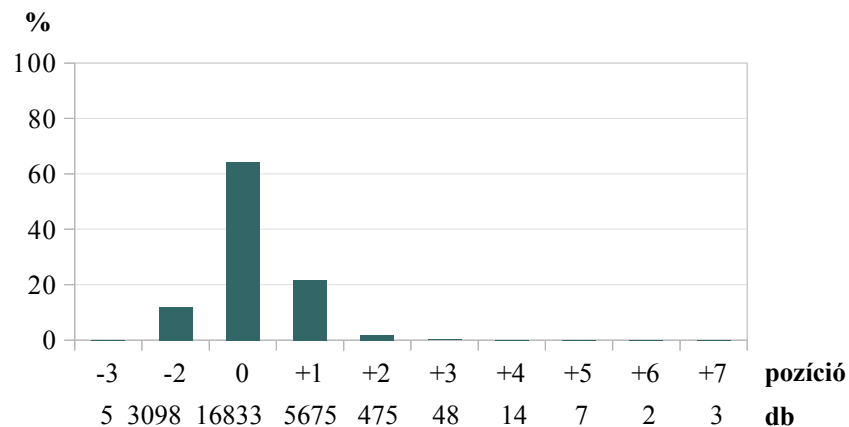
- (46) a. [...] én lelkemet **el** nem **hagytad** sok veszedelmekben. (*Festetics-kódex*, 1492–94)  
b. [...] tehát önmagát **fel** nem **akasztja** vala. (*Guary-kódex*, 1508 előtt)

A posztverbális igekötőt tartalmazó mondatok – ld. (47) példák – jellemzően felszólítások, vagy szerkezeti fókuszot tartalmaznak. Valószínűleg a kódexek stílusából adódóan kevés a kérdő mondat. Az új tagadó szerkezet itt még elenyészően van jelen.

- (47) a. **Jelentette** vala az Úristen **meg** asszony Jacobának, [...] (*Jókai-kódex*, 1440 körül)  
b. Immár **bizonyíttassék** az is **meg**, [...] (*Guary-kódex*, 1508 előtt)

A találatok nagy része +1 pozícióban tartalmaz igekötőt, 55 esetben +2-ben. Ide kerültek a ‘finit ige – *vala/volna* – finit ige igekötője’ konstrukciók (mind nem-semleges mondat), főleg ezek alkotják a találatokat. A leggyakoribb beférkőző szavak: *azt*, *-e*.

A **középmagyar adatokban** (ld. 10. ábra) nő a posztverbális igekötők aránya (16,7%-ról 23,8%-ra). A beszélt nyelvi jelleg miatt sok a felszólító mondat. A legszélső preverbális pozíció a -3, amely minden esetben tagadó mondatot jelöl. A beférkőző szavak egy tagadószó (*nem*, *sem*, *ne*) és az *is*, nem feltétlenül ebben a sorrendben, ld. (48).



**10. ábra.** Prototipikus igekötők mondatbeli eloszlása a TMK-ban.

(48) a. [...] hogy a szemeim héjai **ki** is ne **forduljanak** miatta. (*Kanizsai Orsolya, 1559*)

b. [...] az ő kenése és segedelme nélkül **meg** sem is **csillapodtak**. (*boszorkányper, 1750*)

A TMK korpusz érdekessége, hogy viszonylag sok olyan mondatot tartalmaz, amelyben a hátravetett igekötő távol került a finit igétől. A (49) példákban látható, hogy ez nem csak egy-két forráson belüli elszigetelt jelenség, hanem a teljes középmagyar korra jellemző, és különböző személyektől származó adat (ahol tulajdonnév van feltüntetve, ott az adott levelet diktáló személyről van szó). Figyelmet érdemel az is, hogy a példákban az igekötők után tagmondathatár van, tehát ezek nem elemezhetők valamilyen határozói frázis részeként (mint pl. *be a szobába*):

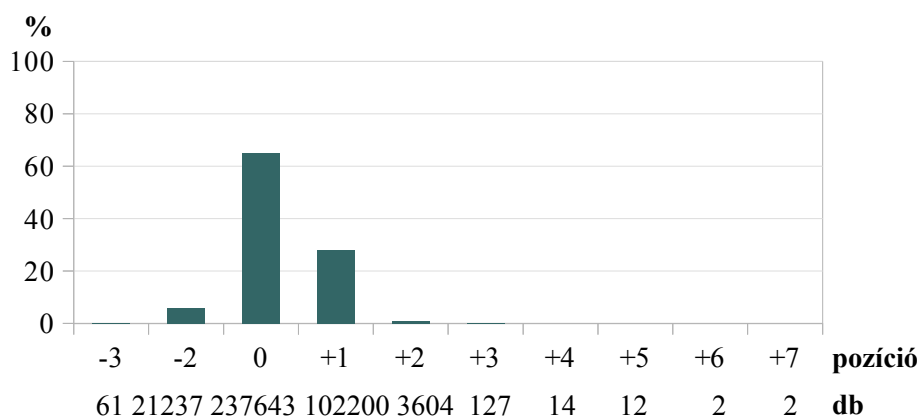
(49) a. [...] gyakran **égeti** a kása a szánkat **meg**, [...] (*Kanizsai Orsolya, 1550*)

b. [...] immár **adnáják** csak a hideg tetemet **meg**, [...] (*Nádasdy Anna, 1551*)

c. [...] **menne** együtt az én [szol]gámmal **be**, [...] (*Telegdy János, 1593*)

- d. Úgy **küldém** Szokoly Albertet mindjára mintegy posta módon **el**, [...]  
(*Telegdy Kata, 1600*)
- e. [...] **tartson** az Isten jó egészségben **meg**! (*Poppel Éva, 1629*)
- f. [...] **jött** a száján olyan szózat **ki**, [...] (*boszorkányper, 1700*)
- g. [...] hogy soha bizony sem **gyógyul** ő abból a nyavalyából **meg**. (*boszorkányper, 1714*)
- h. [...] **jött** is egy kis munka németül róla **ki**, [...] (*Szilágyi Sámuel, 1730*)

Az **újmagyar adatok** esetében két korpusz is rendelkezésre állt, az MTSZ és az MNSZ. Az előbbi 1772 és 2010 között keletkezett forrásokat tartalmaz, ezek adatai láthatók a 11. ábrán. A megszakított szórend itt is a tagadó mondatokra jellemző – ld. az (50) példákat –, de jól látszik, hogy a preverbális igekötők aránya tovább csökkent (70,9%).



**11. ábra.** Prototipikus igekötők mondatbeli eloszlása az MTSZ-ben.

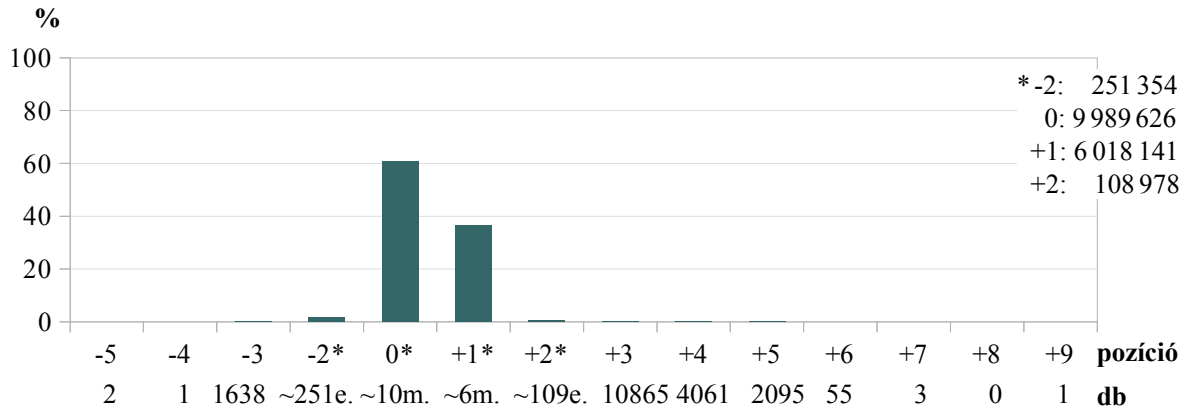
- (50) a. Kin kívül más **meg** is nem **oltalmazhatja**. (*próza, 1781*)
- b. Azért bizonyosan **meg** sem is **mondhatom**, [...] (*próza, 1781*)

A posztverbális igekötők esetében most is észlelhető néhány nagyobb eltávolodás a finit igétől – ld. az (51) példákat –, de ez a korpusz méretéhez viszonyítva nem olyan jelentős, mint a TMK esetében.

- (51) a. **Keljen** bár az egész világ **fel** ellened! (*próza, 1849*)
- b. [...], nem **üledpedhetik** a víz a kávéval annyira **meg** a testben, [...] (*orvosi szöveg, 1791*)

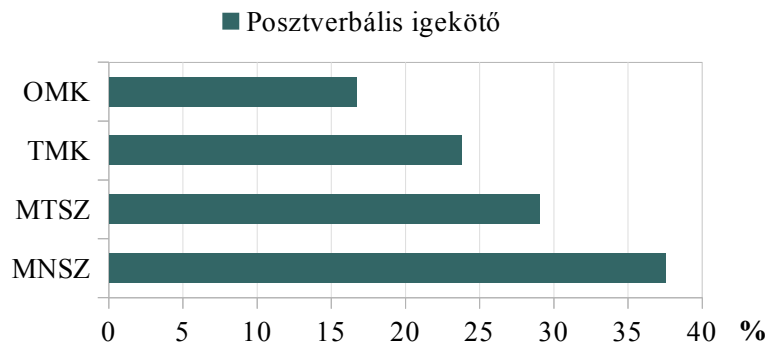


Az MNSZ korpusz 20-21. századi szövegeiben a preverbális igekötők aránya csak 62,5% (ld. 12. ábra). Ekkorra már teret nyert a fordított szórendű tagadó szerkezet. Az MNSZ-en végzett korpuszmérést részletesen tárgyalja a korábbi 4.1.2. fejezet.



**12. ábra.** Prototipikus igekötők mondatbeli eloszlása az MNSZ-ben.

Az igekötők eloszlásában megfigyelhető az a tendencia, hogy az ómagyar kortól napjainkig monoton növekszik a fókuszált összetevőt tartalmazó mondatok aránya. Az Ómagyar Korpuszban még csak 16,7%-os a jelenlétük, míg az MNSZ-ben 37,5%, ld. a 13. ábrát. Ez összefüggésbe hozható a fordított szórendű tagadás térnyerésével, amelyről a következő fejezetben lesz szó.

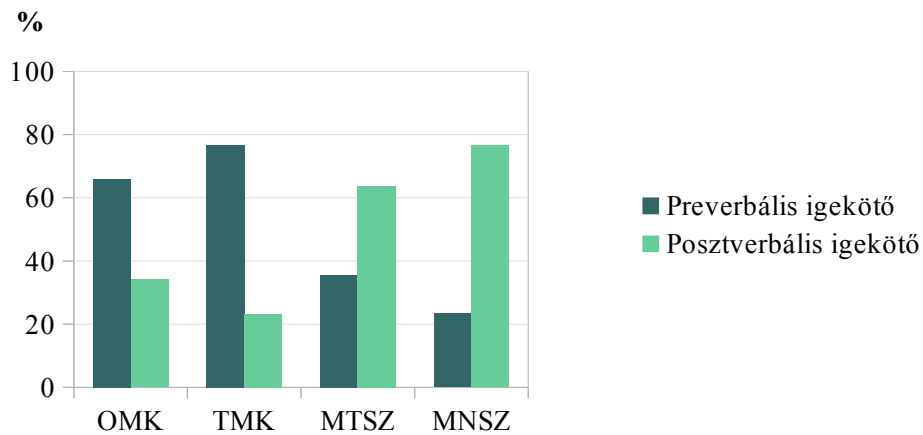


**13. ábra.** A posztverbális nem-semleges mondatok arányának változása a vizsgált korpuszokban.

#### 4.2.3. A kétféle tagadó szórend arányának változása

É. Kiss (2014c) részletesen foglalkozik a kétféle mondattagadó szerkezettel, amelyek közül az archaikusabbra az 'igekötő – tagadószó – ige' mintázat jellemző (megszakított szórend), az újabbra pedig a 'tagadószó – ige – igekötő' (fordított szórend). Az ezzel kapcsolatos megfigyeléseket árnyalja tovább Gugán (2015, 2017). Ő az ó-, közép- és korai újmagyarban foglalkozik a tagadó

mondatok szórendjével. Azt találja, hogy a fordított szórend a kezdetektől jelen van és fokozatosan nyer teret: 1840-50-re – tehát az újmagyar kor középső harmadára – tehető az a szakasz, amikor ez a szerkezet válik gyakoribbá. Ennek kapcsán végeztem egy mérést, hogy látni lehessen, hogyan oszlik meg a tagadó mondatok aránya az egyes korpuszok preverbális és posztverbális igekötős mondatai között. Az eredményt a 14. ábra szemlélteti.



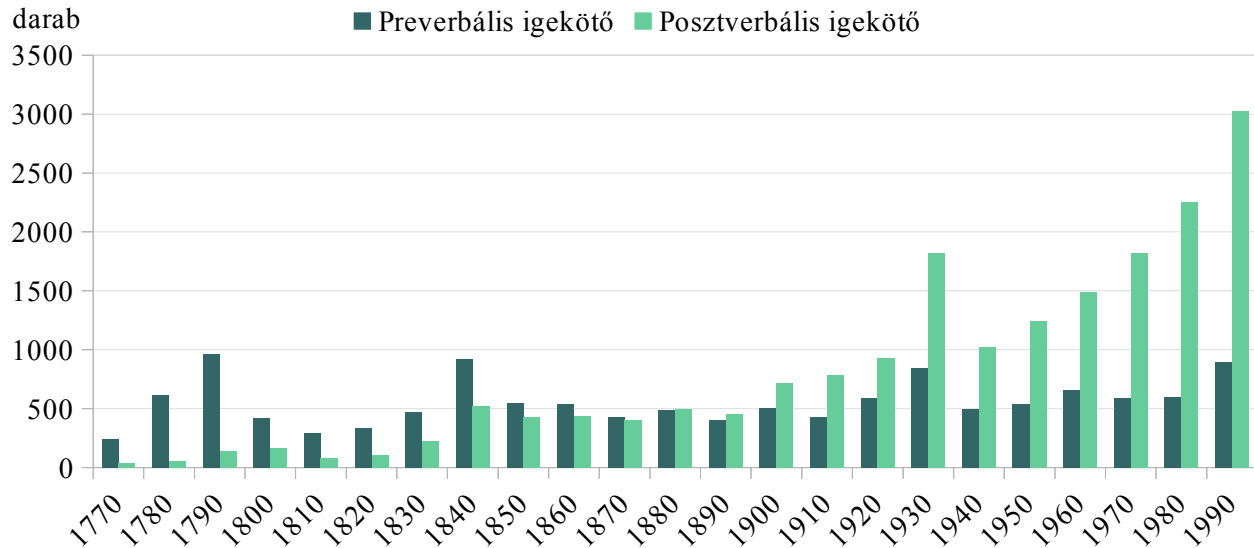
**14. ábra.** Megszakított és fordított szórendű tagadó mondatok százalékos aránya a vizsgált korpuszokban.

A megszakított szórendű tagadás jóval nagyobb arányú a fordítottnál az ó- és középmagyar adatokban, ahogy arra számítani lehetett. Kérdés viszont, hogy a TMK korpuszban miért gyakoribb (körülbelül 10%-kal), mint az OMK-ban. Ha S-görbét nem is, de monoton csökkenést várnánk. Elképzelhető, hogy az OMK adatait torzítják a Müncheneri-kódex mondatai. Ezeknél általánosnak mondható még a közvetlenül preverbális igekötő is a tagadás esetén, ami így sem a megszakított, sem a fordított szórendhez nem számítható be, ld. az (52) példasort. Ez azonban elszigetelt jelenség, és a másik négy forrásban nem található meg.

- (52) a. *nem megoldja*-e ő tehenét  
 b. ki *nem megtöretik*  
 c. ki *nem megtudatik*

Lehetséges magyarázat az is, hogy a beszélt nyelvben erősebb volt az archaikus változat jelenléte, ezért gyakoribb ez a TMK-ban. Az MTSZ korpuszt megvizsgáltam évtizedek szerinti bontásban is, hogy a fordított szórendű tagadás térnyerését pontosabban nyomon lehessen követni (15. ábra). A fordulópont 1880 és 1890 közé esik, ettől kezdve a fordított szórendű tagadás növekedő tendenciát

mutat. Ez kissé eltér Gugán (2017) eredményeitől: nála a fordulópont 40 évvel korábban látszik. A különbség oka valószínűleg az, hogy ő a *meg, el, fel* igeekötőket vizsgálta közvetlenül a finit ige környezetében, míg ez a kutatás további prototipikus igeekötőket is számba vett.



**15. ábra.** Megszakított és fordított szórendű tagadó mondatok arányának változása az MTSZ adatai alapján, évtizedes bontásban.

### 4.3. Összegzés

Ennek a fejezetnek az volt az elsődleges célja, hogy átfogó képet adjon az igeekötők eloszlási mintáiról minden olyan szerkezet esetében, amelynél az igeekötő el tud válni az igtől, igenévtől vagy egyéb igei származéktól, amellyel szóközi egységet alkot. Az eloszlási minták közlésén túl célja volt az is, hogy feltárja azokat a tényezőket, amelyek az igeekötő pozícióját befolyásolják, valamint felhívja a figyelmet a ritkább szintaktikai jelenségekre is.

A fejezet első részében egy szinkrón korpuszvizsgálatot mutattam be, amely az MNSZ2 korpusz 20-21. századi szövegein alapult. A következő szerkezeteket vettem sorra és jellemeztem az igeekötők disztribúciója, valamint a beférkőző szavak szempontjából: finit ige, infinitívusz, határozói igenév, különféle melléknévi igenév-típusok és deverbális főnevek. A kutatásom főbb eredményei a következők:

Kimutattam, hogy a prototipikus igeekötők relatív gyakoriságukat tekintve közel maradnak a finit igehez, míg a periférikusabb igeekötők jobban eltávolodhatnak. Ehhez kapcsolódva megneveztem két olyan tényezőt, amely valószínűleg hatással van arra, hogy a hátravetett igeekötő mennyire

távolodik el a finit igétől. Az egyik a szöveg szerkesztettségi foka: szerkesztetlen – főleg szóban elhangzott – szövegben nagyobb annak a valószínűsége, hogy az igekötő távol kerül a finit igétől, mint gondozott, szerkesztett szöveg esetében. A másik az ige utáni mondatösszetevők, köztük a hátravetett igekötő fonológiai súlya. A méréseim alapján az a tendencia rajzolódik ki, amelyet a növekvő összetevők törvényének vagy Behaghel-törvénynek is neveznek, és É. Kiss (2007) szerint a magyarban a finit ige utáni mondatszakaszra érvényes: a rövidebb összetevő megelőzi a hosszabbat, ha ezt szintaktikai szabály nem gátolja. Ez azzal is egybevág, hogy az egy szótagú, prototipikus igekötők ritkábban távolodnak el a finit igétől, mint a több szótagú, periférikusabb igekötők.

Az infinitívusz kapcsán korpuszméréssel igazoltam, hogy az igekötője preverbálisan nagyon messzire eltávolodhat tőle, de csak akkor, ha segédige-szerű elem – elsősorban finit ige – férkőzik be a szerkezetbe, és az igekötő az ehhez tartozó igemódosítói pozíciót foglalja el. Ez alátámasztja azt a megfigyelést, hogy bár az igekötő az infinitívusszal alkot szótári egységet, szórendi és prosódiai szempontból szorosabban kapcsolódik a segédige-szerű elemhez (Kálmán C. et al. 1989). Megállapítottam, hogy ezzel rokon szintaktikai viselkedése van a létigével passzív szerkezetet alkotó határozói igenévnek is. Ennek az igekötője akkor kerül távoli preverbális helyzetbe, ha a létige előtti igemódosítói pozíciót foglalja el.

A határozói igenevet illetően még azt a megfigyelést tettem, hogy az igekötője hátravethető, de csak akkor, ha az igenév állapot- vagy módhatározói szerepet tölt be. A létigével alkotott passzív konstrukcióban az igekötő mindig preverbális.

Kimutattam, hogy az *el kell, hogy mondjam* típusú szerkezet nemcsak finit igeik igekötőjének a kiemelésével jöhet létre, hanem infinitívuszok és határozói igenevek igekötői is kiemelhetők fölérendelt tagmondatba (pl. *hozzá kell, hogy tudjunk férni* és *össze kell, hogy legyen kötve*). A három szerkezetben az a közös, hogy a mellékmondatban időjeles ige van, amely az infinitívusz esetében egy segédige-szerű elem, a határozói igenév esetében a létige.

A melléknévi igeneveket vizsgálva először is kiderült, hogy ha a -hAtÓ végződésű igenév állítmányi helyzetben van, és a tagmondatban nem hangzik el finit ige, akkor a -hAtÓ teljes mértékben átveszi ennek a helyét: az igekötője ugyanúgy elválhat, mint egy finit igéé, az igekötők disztribúciója is hasonlóan alakul, emellett a szerkezetbe férkőző szavakat tekintve is párhuzam látszik a finit igei szerkezettel. Mindez az ellen szól, hogy a -hAtÓ-t melléknévképzőnek tekintsük – ahogy például Kiefer (2003a) teszi –, hiszen az ezzel képzett szó nagyon erős igei tulajdonságokat őriz.

Az általános szakirodalmi véleménnyel szemben kimutattam, hogy az -AndÓ végződésű melléknévi igenevek állhatnak állítmányi helyzetben, és ekkor az igekötőjük is hátravethető, bár a jelenség kétségtelenül ritka (az esetek 1,72%-ára jellemző, ez a vizsgált korpuszban 1519 példát jelent).

Ahogy várható volt, a leginkább megkötő igeneves szerkezetnek az -Ó és az -(O)(t)t végződésű igenevek bizonyultak. Ezeknek az igekötője csak preverbálisan válik el, és úgy is meglehetősen ritkán: az -Ó végű igeneveknél 98,34%-ban, az -(O)(t)t végűeknek 99,29%-ban közvetlenül az ige-név előtt áll. Ezeknél az igeneveknél a finit igével ellentétes tendenciák látunk olyan szempontból, hogy az elvált igekötők nagy része prototipikus igekötő. Ennek azonban nem szintaktikai okai vannak, hanem arról van szó, hogy a gyakoribb – egyúttal prototipikusabb – igekötők több olyan frazéma-szerű szerkezetben vesznek részt, amely megszakított tagadó szórenddel rögzült (pl. *fel nem tett kérdés, be nem tartott ígéret*). A deverbális főnevek esetében is ugyanez a helyzet, csak itt inkább a periférikus igekötők aránya a szembeűnő (pl. *egyet nem értés, együtt nem működés*).

A fejezet második részében egy diakrón korpuszvizsgálat eredményeiről számoltam be. Ennek során a prototipikus igekötők mondatbeli helyét vizsgáltam a finit ígéhez képest ó-, közép- és újmagyar kori szöveganyagok bevonásával. A diakrón vizsgálat legfontosabb eredménye az, hogy kimutatta a hátravetett igekötőt tartalmazó, nem-semleges mondatok arányának monoton növekedését az ómagyar kortól napjainkig. Ez a tendencia főként azzal magyarázható, hogy a fordított szórendű tagadás fokozatosan teret nyert az archaikus, megszakított szórendű tagadással szemben, ami a korpuszadatok alapján szintén jól nyomon követhető.

Ilyen jellegű szintaktikai vizsgálat korábban nem készült az igekötős szerkezetekről. Bár a kutatás középpontjában az igekötők álltak, bízom abban, hogy ez a munka hozzá tud járulni általában véve az igei szerkezetek adatvezérelt feltérképezéséhez is, és rá tudja irányítani a figyelmet olyan konstrukciókra, amelyek eddig elemzetlenül maradtak.

## 5. Az igekötők produktív kapcsolódási mintái

„[...] a bölcsélet nyelve valósággal elizélődött:  
megizélték, tönkreizélték, agyonszöszölték  
és elmaszlagolták a filozófiát.”

(Határ Győző: Darályvilág Buzdugániában)

Bár a fenti idézet nem bővelkedik tartalmas igékben, mégis tudjuk értelmezni: valakik hatására a filozófia megváltozott, és ez a változás nem volt előnyös. Ezt azért értjük, mert a mondatban szereplő ‘igekötő – képző – vonzatkeret’ hármasoknak (pl. *el-X-elődik valami*, *tönkre-X-el valamit*) önálló jelentése van, és ezt felismerjük akkor is, ha a szerkezetben szereplő ige az *izél*-hez hasonlóan semmitmondó. Ezeknek a hármasoknak a korpuszvezérelt vizsgálata közelebb visz minket ahhoz, hogy megismerjük az igekötők produktív kapcsolódási mintáit. Nem vállalkozom többre ennél a „közelebb vitelnél”, mert a vizsgált jelenség annyira szerteágazó, hogy egy külön disszertáció témája lehetne.

A produktív kapcsolódási minták azt adják meg, hogy az igekötő különféle jelentéseiben milyen igecsoportok tagjaihoz járulhat. Ehhez három tényezőt szükséges figyelembe kell venni (Ladányi 2012: 72-74): (1) Az igekötő szemantikája: Milyen jelentései vannak az igekötőnek? (2) Az igecsoport szemantikája: Mi a közös jelentéseleme azoknak az igéknek, amelyekhez az igekötő egy-egy jelentésében járulhat? (3) A létrejövő igekötős ige vonzatkerete. Amint láttuk, sok esetben már az igekötő és a vonzatkeret együttese elég ahhoz, hogy tudjuk, miféle ige lehet a szerkezetben.

Ebben a fejezetben az első tényezőt vizsgálom meg a legalaposabban azáltal, hogy 56 igekötő termékeny jelentéseit foglalom össze egy ontológiában. A második és harmadik tényezőt, tehát az igecsoportok szemantikáját és az igekötős ige vonzatkeretét automatikus módszerekkel vizsgálom – az előbbi szóbeágyazással, az utóbbit vonzatkeret-kinyerő eljárással –, és az eredményeimet a PREVCONS nevű erőforrásban teszem közzé. A produktív kapcsolódási minták átfogó leírása nem része ennek a kutatásnak, de jövőbeli kutatási irányként egyértelműen következik belőle.

Elsőként a munkafázisokat és a felhasznált eszközöket ismertetem, majd az igealkotás három leggyakoribb módjáról lesz szó: a névszói és igei bemenetű igeképzésről, valamint a hangzásséma-alapú igealkotásról. Ezt a két újonnan létrejött erőforrás, a PREVCONS és az igekötők ontológiájának bemutatása követi. Végül összegzem a legfontosabb eredményeket.

## 5.1. A kutatás módszere

A kutatás alapjául a PREVLEX táblázat szolgált (ld. 2.2.2. fejezet), de ebből csak a hapaxokat, tehát az egyszer előforduló igekötős igéket vettem figyelembe. Azért döntöttem így, mert a hapaxok jellemzően produktív szóalkotás eredményei, így a morfológiai szerkezetük és a jelentésük is kompozicionális – legalább is nagyobb eséllyel az, mint a gyakori szavak esetében. A hapaxok közül is csak azokat őriztem meg, amelyeknél az igekötő terjeszkedő produktivitása 0-nál nagyobb volt (ld. 3.3.1. fejezet). A PREVLEX 122 ilyen igekötőt tartalmaz, ezek összesen 21 038 hapaxban jelennek meg.<sup>36</sup>

A kigyűjtött hapaxokat meglemeztem az emMorph segítségével (Novák et al. 2016, 2017). Ez egy nyílt forráskódú, automatikus morfológiai elemző, amely a felszíni szóalakot morféimákra és morfológiai kódokra bontja fel. Azért volt erre szükség, mert az MNSZ2-ben – amelyből a PREVLEX épült – az igekötős igék sokszor UNKNOWN elemzést kapnak, vagy ha jó az elemzésük, akkor sem látszik a mélyszerkezetük: sem az, hogy hol van az igekötő és ige közötti határ, sem az, hogy az ige adott esetben szóképzés eredménye-e. Az (53) példa azt mutatja, hogy milyen elemzéssel jelenik meg az *elmamlaszkodik* szó az MNSZ2-ben, és milyen elemzés érhető el hozzá az emMorph-ban.<sup>37</sup>

(53) a. *elmamlaszkodik/elmamlaszkodik/IK.IGE.e3*

b. *el[/Prev]=el+mamlasz[/N]=mamlasz+kodik[\_Vbz:k0d/V]=kod+ik[Prs.NDef.3Sg]=ik*

Az emMorph nélkülözhetetlen ehhez a munkához, mivel manuálisan szinte lehetetlen ekkora adathalmazt feldolgozni. A fedése kiemelkedőnek számít 92,6% körüli értékkel (Novák et al. 2017: 73), és pontosság tekintetében is megbízható. Ugyanakkor generál néha olyan elemzést is, amely elvileg lehetséges, de nagyon valószínűtlen. Ilyen például a következő igék elemzése: *ereszkedik* → *eresz+kedik* ‘ereszként viselkedik’, *módosít* → *módos+ít* ‘módossá tesz’. Az ilyen jellegű hibákat nem kezdtem el egyesével javítani, mivel túl sok lett volna az átnézendő adat.

Ez a morfológiai elemző a produktív képzőket ismeri fel aktív – azaz önállóan is használt – tövek mellett. A kutatás előrehaladtával világossá vált, hogy szükségem van olyan szerkezetek

<sup>36</sup>Előfordul, hogy lényegében azonos igelemmák tűnnek fel *-ik-es* és *-ik-telen* alakban is (pl. *ajándékoz* – *ajándékozik*, *akcióz* – *akciózik*). Emiatt „emberi szemmel” nézve kevesebb hapaxot kapnánk, mint automatikus módszerrel.

<sup>37</sup><https://emmorph.herokuapp.com/> (a mélyszerkezet az *analyze* opcióval nyerhető ki)

azonosítására is, amelyek az emMorph-ban nem jelennek meg. Ezek két okból maradhatnak ki: vagy azért, mert passzív – vagy más szóval fiktív – tő van bennük, vagy azért, mert túl szigorú az elemző szabályrendszere, például nem azonosítja az -Ul képzőt, ha az több szótagú szavakhoz kapcsolódik. Az ilyen esetekre saját szabályokat írtam, ezeknek a kimenete az emMorph-éval formailag pontosan egyező elemzés.

Az olyan hapaxok esetében, amelyeknek az igéi szóképzéssel jöttek létre, a képzés bemenetétül szolgáló szavak listáját – például névszói bemenetű igeképzésnél a névszók listáját – megvizsgáltam word2vec keretrendszerrel használó szóbeágyazással is (Siklósi és Novák 2016; Novák et al. 2016). Ennek a lényege az, hogy a szavakat egy valós vektortér pontjaiként képezzük le, és ebben a sokdimenziós térben a hasonló szavak egymáshoz közel, az eltérőek egymástól távol rendeződnek el. A hasonlóság a választott modelltől függően szemantikai vagy szintaktikai jellegű lehet.

A szóbeágyazási modell használatára azért volt szükség, mert az igekötők produktív kapcsolódási mintáit csak az egyes jelentéseikre nézve tudjuk meghatározni (Ladányi 2012: 72). Ezzel az eszközzel lehetőség nyílik olyan objektív szemantikai osztályozás létrehozására, amelyet hagyományos módszerekkel rengeteg idő lenne elérni, ráadásul a legjobb szándékkal sem tudna következetes maradni az osztályozás. Természetesen nem arról van szó, hogy kritika nélkül hagyatkozzunk az automatikusan előállított szemantikai osztályokra, mert ezek néha más jellemzők (pl. a szavak stílusregisztere) alapján állnak össze. A szóbeágyazás alkalmazása mégis olyan alapot ad, amelyet később esetleg csak korrigálni kell.

Végül lekértem az MNSZ2-ből azokat a mondatokat, amelyek a hapaxokat tartalmazzák. Ezeket Sass (2011) módszerével tagmondatokra bontottam, és a releváns tagmondatokból kinyertem a hapaxok feltételezett vonzatkeretét. Ezek az adatok – kiegészítve a fentebb tárgyalt morfológiai és szemantikai információval – alkotják a PREVCONS-ot, amelyet az 5.3. fejezet mutat be részletesen. A PREVCONS segítségével elemezhetővé váltak az igekötők produktív kapcsolódási mintái.

## 5.2. Az igealkotás produktív esetei

*Mézesmázaskodik, markecolgat, succog.* Korábban talán még nem hallottuk ezeket az igéket, mégis van elképzelésünk arról, hogy mit jelentenek. A *mézesmázaskodik* esetében a *-kodik* végződés miatt tudható, hogy az alany valamiként, valahogyan viselkedik. A *markecolgat* igében a *-gat* arra utal, hogy ismétlődve történik valami, még ha nem is tudjuk, mi az (a *markecol* jiddis eredetű



szlengszó, ‘részeg embert kirabol’ jelentéssel). A *succog*-nál pedig szinte halljuk a hangot, amivel a cselekvés jár, bár a pontos cselekvést nehéz volna megnevezni. Ezeknek az igéknek a mintájára képesek vagyunk még rengeteg hasonló szót alkotni. A *mézesmázás* helyére beilleszthetünk bármilyen melléknévi szerkezetet, a *markecol* helyére tetszőleges igét rakhatunk, a *succog* mintájára sok hangutánzó vagy hangulatfestő igét hozhatunk létre. Ez a magyar nyelvben működő, produktív igealkotási módokkal – jelen esetben a névszói és igei bemenetű igeképzéssel, valamint a hangzássémákkal – magyarázható.

Ez a fejezet az említett három igealkotási módot ismerteti. Ezeknek azért van jelentősége a központi témánk szempontjából, mert a produktív szóalkotással létrejött igék csoportja nyitott, tehát amelyik igekötő ilyen igecsoport tagjaihoz kapcsolható, annak minden bizonnyal a produktív kapcsolódási mintáit találjuk meg (Ladányi 2007: 201).

A kutatás kiinduló adathalmaza a 21 038 igekötős igei hapax (ld. az 5.1. fejezetet). Ezeket első körben úgy adtam meg az emMorph-nak, hogy az igekötőt elhagytam belőlük – így 11 602 igitípus morfológiai elemzését kaptam meg. Az igekötő elhagyásának csupán gyakorlati oka volt: az emMorph sok esetben ismeretlennek jelölte a szót, ha igekötős volt, igekötő nélküli formában viszont helyes elemzést rendelt hozzá. A sikeres elemzést követően természetesen összekapcsoltam a kimenetet az eredeti, igekötős alakkal.

A továbbiakban külön alfejezetek foglalkoznak a három igealkotási móddal, a névszói és igei bemenetű igeképzés esetében a képzők, a hangzássémák esetében a különféle alsémák részletes bemutatásán keresztül. Végül arról lesz szó, hogy ezek milyen arányban jelennek meg az igekötős hapaxok halmazában. A téma hatékony tárgyalása miatt szükségessé vált a fonológiai jelölések használata – ezek feloldása a dolgozat elején, az 5. oldalon található. A három igealkotási mód jól áttekinthető, egy oldalas összefoglalását tartalmazza az I. számú melléklet.

### 5.2.1. Névszó + igeképző

A névszói bemenetű igeképzés esetében a névszói bemenet nagyon sokféle lehet: leggyakrabban egyetlen főnév, melléknév vagy számnév (pl. *traktorozik*, *büdösít*, *tizennyolcaz*), de nem ritka az sem, hogy komplex, több szóból álló szerkezet. Lehet például jelzős főnév (pl. *buta liba* → *butalibáskodik*), tulajdonnév (pl. *Sasvári Szilárd* → *sasváriszilárdoz*), sőt előfordul az is, hogy egy teljes üdvözlési forma válik igévé (pl. *Jó reggelt!* → *jóreggeltez*, *Kezit csókolom!* → *kezitcsóko-*

*lomoz*). A névszói bemenet összetettségének elvi korlátja nincs, de performancia-beli van: minél komplexebb, annál nehezebb létrehozni és megérteni is.

A névszókhoz járuló igeképzők két csoportba sorolhatók aszerint, hogy milyen tövekhez kapcsolódnak. Eszerint beszélhetünk valódi (aktív) tövekhez, valamint fiktív (passzív) tövekhez járuló igeképzőkről. Az utóbbi esetben a fő önállóan nem létezik a magyar nyelvben. Sokan fiktív fővel működő igeképzésként elemzik a hangutánzó és hangulatfestő igéket is, én ettől el fogok térni (ennek részleteit ld. az 5.2.3. fejezetben).

A **valódi névszói tövekhez járuló, produktív igeképzőkről** nyújt áttekintést a 30. táblázat. Ezekről részletesen olvashatunk Kiefer és Ladányi (2000c), valamint Ladányi (2007) munkáiban, és a korpuszadatok mellett én is elsősorban ezt a két művet használom fel az egyes igeképzők bemutatásához.

képző	példák	db
-Vz(ik)	<i>zebracsíkoz, doktoroz, kávézik</i>	2558
-(s)Odik	<i>babásodik, jegesedik, vörösödik</i>	638
-VI	<i>falfirkál, offtopikol, chatel</i>	525
-(V)(s)ít	<i>bénít, felhőkarcolósít, vidámít</i>	420
-(s)Ul	<i>halkul, szelídiül, zombul</i>	205
-(s)kOdik	<i>flegmáskodik, izmoskodik, trollkodik</i>	180

**30. táblázat.** Valódi névszói tövekhez járuló, produktív igeképzők néhány példával. A *db* az ige-típusok gyakoriságát mutatja.

A leggyakrabban megjelenő képző, a **-Vz(ik)** alapértelmezettnek tekinthető a denominális igeképzésben. Ladányi (2007: 76) a képző általános jelentését így fogalmazza meg: ‘N-nel kapcsolatban tesz valamit, N-nel foglalkozik’ (ahol az N a névszói bemenetet jelöli). Ez egy annyira kiüresedett jelentés, hogy tulajdonképpen azt is mondhatjuk, a képzőnek csak funkciója van. Az emMorph különválasztja az *-ik*-re végződő variánst, amelyet intranzitívnak tekint (pl. *torrentezik*, tárgy nélkül), és az *-ik* nélkülit, amelyet tranzitívnak elemez (pl. *torrentez valamit*). A korpuszadatok vizsgálatakor úgy láttam, hogy ez a két eset nem válik el ilyen egyértelműen egymástól – tehát intranzitívnak elemzett ige mellett is megjelenik tárgy –, ezért nem különítem el a két változatot.

A **-VI** a -Vz(ik) rivális képzője, mivel a jelentésében ugyanúgy kiüresedett, és elviekben tetszőleges névszóhoz járulhat. A két képző hatóköri konfliktusát jól mutatja az, hogy a vizsgált ige-típusok közül 86 esetben egyezik a névszói bemenet, pl. *botoxol* – *botoxoz*, *szörföl* – *szörföz*,

*filterel* – *filterez*. Ezeknek az igepároknek a tagjai között nem találunk jelentésbeli különbséget, de a nyelvi rendszerben rögzültebb, gyakoribb igéknél már jellemzően igen (ilyen például a *házal* és a *házaz* közti jelentéskülönbség).

A 30. táblázat gyakorisági adatai jól mutatják, hogy a -VI képző sokkal ritkábban jelenik meg, mint a -Vz(ik), ami arra utal, hogy az előbbinek a használati köre szűkebb. Ladányi (2007: 107) ezt a következőkkel magyarázza: a -VI beillesztő képző, amelyet jellemzően akkor választunk, ha a névszó (1) mássalhangzóra végződik (kivéve /l/ vagy /r/) és egy szótagból áll, vagy (2) magas vagy vegyes hangrendű fiktív tő, vagy (3) angol cselekvésigéből létrejött *-ing* végződésű főnév. Ladányi megfigyeléseit a korpuszadataim is alátámasztják.

Ugyancsak rivális képzők a **-(s)Odik** és a **-(s)Ul**, amelyek a következő általános jelentésen osztoznak: ‘N-né változik’. A használati körükben az az eltérés figyelhető meg, hogy egy szótagú alapszóhoz jellemzően a **-(s)Ul** járul, míg több szótagúhoz inkább a **-(s)Odik** (Ladányi 2007: 245). Természetesen ezeknél is előfordulhatnak párhuzamos képzéssel létrejött alakok, pl. *sötétedik* – *sötétül*, *népesedik* – *népesül*. Ennek a két képzőnek a műveltető párja a **-(V)(s)ít**, amelynek az általános jelentése: ‘N-né változtat’ (pl. *sötétít*, *népesít*).

Hangsúlyozni kell, hogy a használati körre vonatkozó megfigyelések tendencia-jellegűek. Ezért igaz ugyan, hogy a **-(s)Ul** általában egy szótagú tövek mellett jelenik meg, de jócskán találunk példát ennek az ellenkezőjére is (pl. *magányosul*, *kristályosul*, *anyagiasul*, *mémesül*, *kereszténydemokratul*). Az ilyen eseteket Ladányi (2007: 253-254) úgy írja le, hogy a vetélkedő képzők közül általános esetben is a speciálisat alkalmazzuk, a képző ezáltal túllépi a hatókörét. Szintén érdekes jelenség, hogy a **-(s)Ul** képző nem kizárólag melléknévhez, hanem néha főnévhez is kapcsolódhat, pl. *szenttehenül*, *pépül*.

Több szakirodalomban – Kiefer és Ladányi (2000c: 154-155), Ladányi (2007: 246) – elhangzik az az állítás, hogy a **-(s)Ul** képző nem produktív, mert a magyarban nem jöhetnek létre új, egy szótagú melléknevek, a nyelvbe bekerült idegen szavakkal pedig nem működik ez a képző. A korpuszadatok ezt cáfolják (pl. *punk* → *(be)punkul*), és a hatókör túllépése is elég gyakori ennek a képzőnek az esetében ahhoz, hogy – legalább is egyelőre – produktívnak lehessen minősíteni.

Ennek a csoportnak a legkevesebb ige típusban megjelenő tagja a **-(s)kOdik** képző, amelynek általános jelentése ‘N-ként viselkedik’. A viszonylagos ritkaságának az lehet az oka, hogy az eddig látott képzőkkel ellentétben itt a bemeneten szemantikai korlát is van: a képző jellemzően élőlényre

utaló névszóhoz kapcsolódik. Élettelen jelentő névszó is szolgálhat bemenetként (pl. *asztalkodik*, *póznáskodik*), de ez nagyon ritka jelenség.

Az **idegen eredetű, fiktív tövekhez járuló igeképzőket** a 31. táblázat foglalja össze. Mivel ezeknél a névszói fő önmagában nem létezik a magyar nyelvben, joggal merülhet fel a kérdés, hogy van-e értelme a fő és a képző szételemezésének (az emMorph például nem jelez szóképzést az ilyen igék esetében). Az adatok között találtam néhány példát arra, hogy magyar szavakon is működnek ezek a képzők (pl. *bolondizál*, *kántorizál*), ezért döntöttem a szételemezés mellett.

képző	példák	db
-ál	<i>annotál, fókuszál, transzponál</i>	303
-izál	<i>motorizál, szolmizál, tematizál</i>	109
-íroz	<i>lavíroz, morfondíroz, políroz</i>	39
-ifikál	<i>kodifikál, purifikál, identifikál</i>	7

**31. táblázat.** Idegen eredetű névszói tövekhez járuló, produktív igeképzők néhány példával. A *db* az ige típusok gyakoriságát mutatja.

A fejezetben tárgyalt tíz képző mind produktív, bár a produktivitásuk mértékében jelentős eltérések mutatkoznak. Zárásképp néhány olyan képzőt említek, amelyek többször felbukkantak ugyan a vizsgált hapaxokban, mégsem tűnnek produktívnak. Ide sorolható az -Ász(ik) képző (pl. *madarászik*, *egerészik*, *bolhászik*), valamint a -(V)gOl (pl. *szamaragol*, *kacsagol*, *elefántogol*).

### 5.2.2. Ige + igeképző

Az igei főhöz kapcsolható képzők illeszkedhetnek eleve képzett igehez (pl. *[[magyar]ít]+gat*), következésképp egymáshoz is (pl. *[[[magyar]ít]gat]+ódik*). Ebbe az alosztályba négy produktív képző tartozik, amelyekről a 32. táblázat ad áttekintést.

képző	példák	db
-(t)At	<i>kocsikáztat, pötyögtet, oszlat</i>	1412
-Ódik	<i>duplázódik, hekkelődik, szövődik</i>	851
-(O)gAt	<i>hintáztat, fuvintgat, búsulgat</i>	510
-(t)Atik	<i>doboltatik, aranyoztatik, rostáltatik</i>	178

**32. táblázat.** Igékhez járuló, produktív igeképzők néhány példával. A *db* az ige típusok gyakoriságát mutatja.

A leggyakoribb a műveltetést kifejező **-(t)At** képző, ennek részletes elemzését adja Komlósy (2000). Ezt az **-Ódik** követi, amely mediális igék képzésére alkalmas. A mediális elnevezés arra utal, hogy az ilyen igék „középúton” vannak a cselekvő és szenvedő igenem között (az igében jelölt esemény megtörténik az alannal, annak akarata vagy aktív közreműködése nélkül).

Az **-(O)gAt** képzőnek két funkciója van: a gyakorítás és a kicsinyítés, vagyis az intenzitás csökkentése. A gyakorító funkció jellemzően az egyszerűséget kifejező igéknél jelenik meg (pl. *ad* → *adogat*), a kicsinyítő pedig azoknál, amelyek hosszabban tartó folyamatot jelölnek (pl. *olvas* → *olvasgat*) (Kiefer és Ladányi 2000b: 199). A két funkció nem független egymástól, sokszor együttesen vannak jelen.

A passzivizáló funkciójú **-(t)Atik** képző jóval elmarad a többitől a gyakoriságát tekintve. Ennek az lehet az oka, hogy a mai nyelvhasználat inkább az összetett passzívt (léltige + -vA képzős igealak), vagy az -Ódik képzős, mediális formát részesíti előnyben. Tehát például a *bearanyoztatik* helyett *be van/lesz aranyozva*, *bearanyozódik* a jellemzőbb.

Az említett négy eseten túl még sok olyan képzőt találunk az adatban, amely igei tőhöz kapcsolódik, de nem – vagy csak nagyon kevésbé – produktív. Különösen népes azoknak a tábora, amelyek gyakorító-kicsinyítő funkcióval bírnak, tehát az **-(O)gAt** alternatívái. Ide sorolható a -kOd (pl. *lopkod*, *tömköd*), a -dOs (pl. *vágdos*, *öldös*), a -dOgÁl (pl. *hulldogál*, *nézdegél*), a -dAl (pl. *rugdal*, *szökdel*), a -dÁcsOl (pl. *bukdácsol*, *nyögdecsel*), az -Ong (pl. *jajong*, *forrong*) és az -igál (pl. *huzigál*, *dobigál*). Ezek a képzők paradigmákba is szerveződnek, erről bővebben ld. Ladányi (2007: 101-105) esettanulmányát.

### 5.2.3. Hangzásséma

Az eddig tárgyalt igealkotási módokat igeképzésnek neveztem, most azonban elhagyom ezt a terminust, és hangzásséma-alapú igealkotásról fogok beszélni. Azért látom indokoltnak a megkülönböztetést, mert az utóbbi esetben nemcsak a végződés kötött, hanem a szótagszám is, és még számos szótagszerkezeti hasonlóság figyelhető meg az így alkotott igéknél. Nem önmagában a szuffixum produktív, hanem a teljes hangzásséma.

Ezzel a megközelítéssel kisebbséget képviselek, ugyanis a szakirodalmak többsége igeképzés-ként tárgyalja ezeket az igealkotási módokat is, azzal a kitételrel, hogy a szótő mindig fiktív. Ilyen elemzést választ pl. Rebrus (2000), Kiefer és Ladányi (2000b) és Ladányi (2007). A képzős elem-

zést eddig nagyon kevesen hagyták el a hangzássémák javára. A kevesek között az egyik legalapossabb munka Benő és Szilágyi N. (2015) cikke, amely az én munkámhoz is nagyban hozzájárult.

Mivel a hangzásséma-alapú elemzés meglehetősen eltér a hagyományosabb megközelítésektől, nem meglepő, hogy az emMorph számára ismeretlen. A hangzássémára illeszkedő igéket csak akkor tudja igeként azonosítani és elemezni, ha szerepelnek a lexikonjában, egyébként ismeretlen szavakként kezeli őket. A hangzássémák azonosítására saját algoritmust írtam, a fejezet további részében ismertetett sémátípusok ennek az algoritmusnak, és nem az emMorph-nak a kimenetei.

A leggyakoribb és legtöbb altípust felmutató hangzásséma az **-Og végű séma** (ld. 33. táblázat). Tagjainak a szerkezeti hasonlóságukon túl közös szemantikai tulajdonságuk az, hogy hosszabb ideig tartó vagy ismétlődő eseményt fejeznek ki.

általános séma és példák	db	megszorított séma és példák	db
(C)(C)VC:Og <i>battyog, hümmög, billeg</i>	62	(C)CO <sub>1</sub> C:O <sub>1</sub> g <i>szötyög, kerreg, csetteg</i>	6
(C)(C)VCOg <i>krákog, büfög, lifeg</i>	59	(C)CO <sub>1</sub> CO <sub>1</sub> g <i>kopog, fröcsög, cseveg</i>	70
(C)(C)VCCOg <i>harsog, csámcsog, prüntyög</i>	13	(C)CO <sub>1</sub> CCO <sub>1</sub> g <i>mormog, dörmög, csemcseg</i>	11
(C)(C)V(C)CVrOg <i>kacsmarog, csavarog, szivárog</i>	6	(C)(C)V(C)CO <sub>1</sub> rO <sub>1</sub> g <i>sutyorog, nyöszörög, sistereg</i>	59
(C)(C)V(C)CVIOg <i>andalog, sandalog, oldalog</i>	5	(C)(C)V(C)CO <sub>1</sub> lO <sub>1</sub> g <i>őgyeleg, édeleg, hízeleg</i>	13
(C)(C)V(C)CV/j/Og <i>kóvályog, nyavalyog, vihajog</i>	4	(C)(C)V(C)CO <sub>1</sub> /j/O <sub>1</sub> g <i>gomolyog, hömpölyög, tévelyeg</i>	10

**33. táblázat.** Az -Og végű hangzássémák, az általános sémák gyakorisága szerinti csökkenő sorrendben. A jobb oldali, megszorított sémák a bal oldali párjuk szigorúbb változatai. A fejléc alatti első sorban szereplő C: (hosszú mássalhangzó) is egyfajta megszorítás, a harmadik sorban látható CC-é (két mássalhangzó), mégis külön sémának tekintem a gyakorisága miatt.

Az ide tartozó sémák állhatnak kettő vagy három szótagból. A két szótagúak között a leggyakoribbak azok, amelyekben a középső mássalhangzó rövid, a két magánhangzó pedig azonos (pl. *potyog, tötyög, mekeg*). A három szótagúak további alsémákra bonthatók az utolsó szótagkezdő mássalhangzójuk alapján. Ez ugyanis csak /r/, /l/ vagy /j/ lehet, ami azért érdekes, mert ezek a mássalhangzók egy csoportot alkotnak a fonetikai hangzósságuk (Kassai 1994: 631) és a fonológiai

jellemzőik alapján is (Siptár 1994: 185).

Bizonyos -Og végű igék -OrÁsz(ik) végződéssel is működnek, ekkor a cselekvés nem pusztán hosszan tartó vagy ismétlődő, hanem alacsony intenzitású is (pl. *vihog* → *vihorász*, *röhög* → *röhörész*, *kacag* → *kacarász*, *tötyög* → *tötyörész*). Ez azonban sokkal korlátozottabb, mint az -Og végű séma (például kérdésesnek tűnik a *fröcsög* → *\*fröcsörész*, *krárog* → *\*krákorász*).

Az -An, -Ant és -int végű sémákat érdemes együtt tárgyalni a szemantikai rokonságuk miatt (ezt a 34. táblázat példái illusztrálják). Mindháromra igaz, hogy mozzanatosak, azaz pillanatnyi eseményt fejeznek ki. Az -An ilyen tekintetben az -Og funkcionális párja (vö. *csattan* – *csattog*, *zöttyen* – *zöttyög*). Az -An végű séma általános jelentése az, hogy ‘X-szel történik valami pillanatnyi’. Az -Ant végű séma ennek a kauzatív megfelelője, tehát alanyként megjelenik a történés okozója: ‘Y tesz X-szel valami pillanatnyit’ (pl. *meglibben a függöny* – *a szél meglibbenti a függönyt*). Az -int az -Ant-hoz hasonlóan kauzatív, és azzal a jelentéstöbblettel bír, hogy az esemény nagyon alacsony intenzitású.

-An		-Ant		-int	
séma és példák	db	séma és példák	db	séma és példák	db
(C)CVC:An <i>csattan, zöttyen</i>	127	(C)CVC:Ant <i>csattant, zöttyent</i>	138	(C)CVC:int <i>csattint, zöttyint</i>	34
(C)CVCCAn <i>sercen</i>	12	(C)CVCCAnt <i>sercent</i>	9	(C)CVCCint <i>sercint</i>	4
(C)CVhAn <i>suhan, zuhan</i>	2			(C)CVhint <i>suhint, zuhint</i>	3

**34. táblázat.** Az -An, -Ant és -int végű hangzássémák a gyakoriságuk szerinti csökkenő sorrendben.

A 34. táblázatban látható, hogy ezeknek a sémáknak minden tagja két szótagú. Az alsémák között csak a középső mássalhangzókat tekintve van különbség. Szembetűnő, hogy csak a /h/ jelenik meg egymagában középső mássalhangzóként. Ezt Benő és Szilágyi N. (2015: 48) úgy magyarázza, hogy a magyarban tulajdonképpen nincsen hosszú /h/, még teljes hasonulás esetén is röviden ejtjük – bár az írása ingadozó –, például *az+hoz* → /ahoz/, *ez+hez* → /ehez/.

Feltűnő a hosszú mássalhangzók gyakorisága is, szemben az -Og végű sémával, ahol éppen a rövid mássalhangzók gyakoribbak. Ennek az lehet az oka, hogy az -An végű sémák jellemzően nagyobb intenzitást jelölnek, és ezt az erősebb hanghatású hosszú mássalhangzók festik le jobban,

míg az -Og végű sémákban a gyengébb hanghatású rövid mássalhangzó fejezi ki az esemény kisebb intenzitását (Benő és Szilágyi N. 2015: 48).

A harmadik nagy csoportot, az **-Ad és -Aszt végű sémákat** szintén a szemantikai rokonságuk miatt tárgyalom együtt (ld. 35. táblázat). Ez a rokonság abban nyilvánul meg, hogy az -Ad végűek általános jelentése ‘X-szel történik valami’, a kauzatív -Aszt végűeké pedig ‘Y tesz X-szel valamit’ (tehát ezek is funkcionális párok).

<b>-Ad</b>		<b>-Aszt</b>	
<b>séma és példák</b>	<b>db</b>	<b>séma és példák</b>	<b>db</b>
(C)(C)VC:Ad <i>fonnyad, pilled</i>	43	(C)(C)VC:Aszt <i>fonnyaszt, pilleszt</i>	31
(C)(C)VCAAd <i>kókad, reked</i>	37	(C)(C)VCAaszt <i>kókaszt, rekeszt</i>	30
(C)(C)VCCAd <i>sorvad, erjed</i>	25	(C)(C)VCCAszt <i>sorvaszt, erjeszt</i>	22

**35. táblázat.** Az -Ad és -Aszt végű hangzássémák a gyakoriságuk szerinti csökkenő sorrendben.

Az -Ad végű séma igéihez az -Ozik képző járulhat, amely a cselekvés tartós vagy ismétlődő jellegét fejezi ki, tehát az 5.2.2. fejezetben tárgyalt -(O)gAt képzőhöz hasonló szerepet tölt be (pl. *fullad+ozik, pukkad+ozik, puffed+ezik, reped+ezik*).

Míg az -Og, -An, -Ant és -int végű sémákba tartozó igéket a szakirodalom általában **hangutánzó vagy hangulatfestő** – összefoglaló néven onomatopoetikus, expresszív – szavaknak tekinti, az -Ad és -Aszt végű sémába illeszkedők ezen a kategórián kívül esnek. Én igyekeztem kerülni mind a ‘hangutánzó’, mind a ‘hangulatfestő’ kifejezést, mert túl szubjektívnek találom (különösen az utóbbit). A hangzásséma ennél konkrétabb, mivel pontosan formalizálható jellemzőkre – szuffixumra, szótagszámra, szótagszerkezetre<sup>38</sup> – illeszkedik.

<sup>38</sup> A szótagszerkezettel kapcsolatban egy megjegyzés: a CV-sémákat nem úgy kell értelmezni, hogy a C helyén tetszőleges mássalhangzó állhat. A szótagnak jólformáltnak kell lennie, tehát a szótagmag felé a szegmentumok hangzósága nő, onnan kifelé pedig csökken. Ennek nem megfelelő szavakat nem is igazán tudnánk kiejteni (pl. *\*rkákog*). Ez a törvényszerűség, amelynek a természetes nyelvek szótagjai legtöbbször megfelelnek, a szonoritási hierarchia – bővebben ld. Törkenczy (1994: 276).



#### 5.2.4. A vizsgált igealkotási módok arányai az igekötős szerkezetekben

A korábbi alfejezetek csak az igékkel foglalkoztak, az igekötőket átmenetileg figyelmen kívül hagytam. Most visszatérek a dolgozat központi témájához, az igekötős szerkezetekhez.

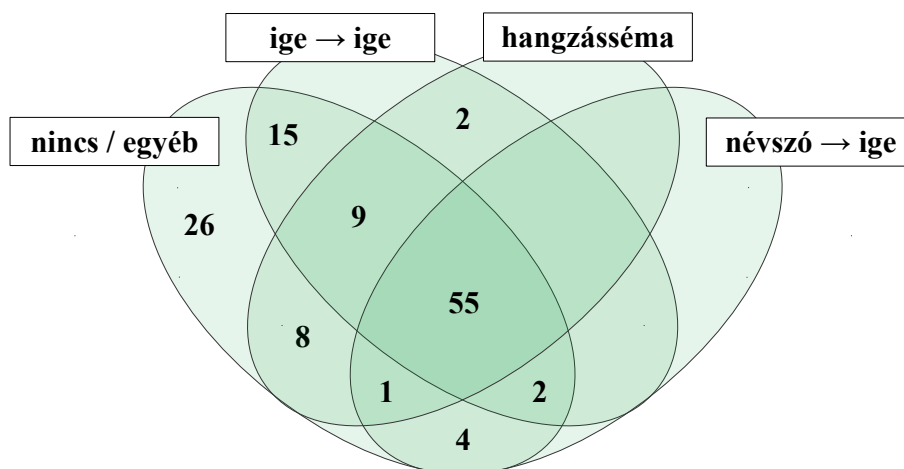
Elsőként érdekesnek tartottam megvizsgálni azt, hogy a három produktív igealkotási mód milyen arányban van jelen az igekötős hapaxoknál, valamint az igekötő elhagyásával kapott igitípusoknál. Ennek az eredményét foglalja össze a 36. táblázat.

	igekötős hapaxok		igitípusok	
	db	%	db	%
<b>névszó → ige</b>	7401	35,2	4985	43,0
<b>ige → ige</b>	4492	21,4	2951	25,4
<b>hangzásséma</b>	1987	9,4	799	6,9
<b>nincs / egyéb</b>	7158	34,0	2867	24,7
	<b>21 038</b>	<b>100</b>	<b>11 602</b>	<b>100</b>

**36. táblázat.** Igealkotási módok gyakorisága az igekötős igei hapaxok, valamint az igekötő elhagyásával kapott igitípusok esetében. A *nincs / egyéb* azokat az eseteket jelzi, ahol nem képzett vagy hangzássémás a szó, vagy pedig nem produktív képzőt tartalmaz.

A legjelentősebbnek a névszói bemenetű igeképzés bizonyult, ez az igekötős hapaxok több mint egyharmadánál (35,2%) figyelhető meg. Ha igekötő nélkül nézzük csak az igitípusokat, akkor ez az adatnak majdnem a felét teszi ki. A hangzássémás igék aránya sem elhanyagolható: az igekötős hapaxok közel egytizedét (9,4%) ezek adják. Ezt fontosnak tartom hangsúlyozni, ugyanis úgy tűnik, hogy a nyelvészeti szakirodalom nagyon kevés jelentőséget tulajdonít az igealkotás ezen módjának. A létjogosultságát mostantól korpuszadatok igazolják.

Ezután azt vizsgáltam meg, hogy az egyes igekötők milyen igealkotási móddal keletkezett igékhez kapcsolhatók. Az eredményt a 16. ábra halmazai szemléltetik. Szinte minden igekötő hapaxai között találunk „egyszerű” igéket, amelyek se nem képzettek, se nem hangzássémára illeszkedők. Általánosnak mondható az igei bemenetű igeképzés megléte is. Ha egy igekötő esetében ez nem jelenik meg, annak csak az adathiány az oka. Erre példa a *hasba* és a *hátba* esete: a hapaxok között szerepel a *hasbabök* és a *hátbabök* is, de képzett ige csak a *hátba* esetében látszik (*hátbabököget*). Ez persze csak esetlegesség, és egy nagyobb korpuszban jó eséllyel a *hasbabököget* is felbukkant volna.



**16. ábra.** A 122 vizsgált igekötő megoszlása aszerint, hogy milyen igealkotási móddal keletkeztek az igék, amelyekhez kapcsolódnak. A halmazokban és azok metszeteiben megadott számok az igekötők darabszámát mutatják. A *nincs / egyéb* azokat az eseteket jelzi, ahol nem képzett vagy hangzássémás az ige, vagy pedig nem produktív képzőt tartalmaz.

Érdekes az is, hogy bár a névszói bemenetű igeképzés van jelen a legnagyobb arányban az igekötős hapaxok között, ez mindössze 62 igekötőnél mutatható ki (az alakváltozatokat összevonva ez a szám csak 56). Ez a kis halmaz nagyjából egybeesik azzal, amit a szakirodalmak többsége az igekötők állományaként jelöl ki, vannak viszont eltérések is. Valószínűleg keveseknek jutna eszébe, hogy igekötőnek minősítse a *szénné* és *szarrá* szavakat, pedig ezek a kapcsolódási mintáikat tekintve sokkal igekötő-szerűbbek, mint például a *közzé* vagy a *véghez*, amelyeket hagyományosan igekötőként sorolnak fel több munkában, ld. Jakab (1976: 9-10) adatait. Az utóbbiak inkább frazémajellegűek, csak egy-két igehez kapcsolódnak.

Az itt bemutatott két mérés alapján belátható: az igealkotás módjára vonatkozó információnak van jelentősége abban, hogy az igekötők produktív kapcsolódási mintáit le tudjuk írni. Ezért döntöttem úgy, hogy ez az adat képezze szerves részét egy igekötős szerkezeteket tartalmazó erőforrásnak, a PREVCONS-nak, amelyet a következő fejezet mutat be.

### 5.3. PREVCONS: Igekötős szerkezetek táblázata

A PREVCONS (*Preverb Constructions*) egy nyílt hozzáférésű táblázat,<sup>39</sup> amely 21 038 igekötős hapaxot tartalmaz azok lehetséges vonzatkeretével együtt. Ezen felül számos meta-adat érhető el benne a hapaxok morfológiai szerkezetéről, szemantikájáról és konkrét szövegbeli előfordulásáról.

<sup>39</sup> <https://github.com/kagnes/prevcons>

Az erőforrásnak sok közös vonása van a Mazsola korpuszlekérdező rendszerrel (Sass 2011): a létrehozásának módjában és a vonzatkeretek leírásához használt formalizmusban is a Mazsolát követi. A lényeges különbség az, hogy a Mazsola igei szerkezetek, igeik és bővítményeik tanulmányozására alkalmas, a PREVCONS pedig az igekötős szerkezetek vizsgálatára használható, akár úgy is, hogy a szerkezetben lévő igt nem vesszük figyelembe. Tehát kigyűjthetünk belőle ‘igekötő – képző – vonzatkeret’ mintákat gyakorisági adataikkal együtt (pl. *össze-N-Vz(ik) -t -vAl*, azaz „összevalamiz valamit valamivel” vagy ugyanez élőlényre utaló vonzatokkal), és egy következő lépésben megnézhetjük, hogy a mintákban milyen igeik állnak a „valamiz” helyén. Ez a módszer lehetővé teszi az igekötős szerkezetek adatvezérelt és szisztematikus vizsgálatát.

A PREVCONS github repozitóriumában két forrásfájl található, egy TSV fájl és egy SQLite adatbázis. Ezek pusztán a formátumukat tekintve térnek el egymástól, tartalmilag megegyeznek. A két formátumot az indokolja, hogy a számítógépes feldolgozáshoz a TSV tűnt praktikusabbnak, a kézi adatelemzéshez pedig az SQLite-alapú megoldás. Az erőforrás 13 oszlopból áll, amelyek a következők:

1. **id**: Az igekötős ige egyedi azonosítója, ennek csak technikai szerepe van.
2. **prev**: Igekötő.
3. **verb**: Igelemma.
4. **argframe**: Az igekötős ige lehetséges vonzatkerete. Ennek a kinyerése automatikus módszerrel történt, egy olyan algoritmussal, amely a Mazsola mintájára készült. A vonzatkeret megállapításához nem a teljes mondatot vettem figyelembe, hanem csak azt a tagmondatot, amelyben az adott igekötős ige szerepelt. A vonzatkeret összeállításával kapcsolatban fontos még tudni a következőket:
  - Csak az olyan esetragokat nyertem ki a tagmondatból, amelyek maximális NP végén szerepelnek, tehát olyan névszói szerkezet végén, amely nem része egyetlen magasabb szintű névszói szerkezetnek sem. Ugyanezen az elven működik a Mazsola is, a különbség csupán annyi, hogy a PREVCONS-nál a névszói szerkezetek azonosítása az emtsv emChunk moduljával történt (Indig et al. 2019; Váradi et al. 2018, 2017), a Mazsolában pedig egy szabályalapú algoritmussal (Sass 2011: 37-40).

- Amikor a kinyert esetrágok bekerülnek a lehetséges vonzatkeretbe, a sorrendjük már nem a tagmondatbeli megjelenésüket követi, hanem abc-rendben szerepelnek.
- Az alanyesetet nem jelölöm.
- Egy speciális vonzat a *magát*, amelyet nem helyettesítek tárgyesetű raggal, hanem teljes szóként őrzök meg. Az 5.4.2. fejezetben látni fogjuk, hogy a *magát* vonzatos igekötős szerkezetek alapjelentése sokszor teljesen eltér azokétól, amelyek egyéb tárgyesetű vonzatot tartalmaznak.

Végül fontos látni, hogy az automatikus adatkinyerés miatt a vonzatkeretek több hibát is tartalmazhatnak. Például nem oldható meg a szabadhatározók elkülönítése a kötelező vonzatoktól, és a *magát* vonzat sem mindig azonosítható helyesen: a *magam*, *magad*, *magunk*, *magatok*, *maguk* alakok egyaránt lehetnek alany- és tárgyesetűek is. Továbbá az sem jószólható meg előre, hogy bizonyos esetrágok mikor lennének kiválthatók egy irányhármast jelölő vonzattal és mikor nem. Például az *el* igekötő mellett megjelenő *-ra*, *-ba* vonzatok helyett *HOVÁ* is állhatna, de a *rá* melletti *-ra* vonzat tovább már nem általánosítható.

5. *prev\_prod*: Az igekötő terjeszkedő produktivitása (erről ld. a 3.3.1. fejezetet).
6. *prev\_hapaxes*: Az igekötőt tartalmazó hapaxok darabszáma.
7. *inputtype*: Az igénél azonosított szóalkotási mód. A *N* névszóból képzett, a *V* igéből képzett, a *PH* pedig hangzássémás igét jelöl. Amennyiben az ige nem produktív szóalkotással jött létre, ebben az oszlopban kötőjel (-) áll.
8. *suffixtype*: Az igénél azonosított képző vagy hangzásséma típusa. Itt is kötőjel (-) áll, ha az ige sem produktív képzővel, sem hangzássémával nem elemezhető. Egyéb esetekben az 5.2. fejezetben használt jelöléseket alkalmazom. A hangzássémás igéknél csak a szuffixum látható ebben az oszlopban, a teljes CV-séma nem. Az utóbbira vonatkozó, részletes adatok egy külön fájlban érhetőek el a PREVCONS github repozitóriumában.<sup>40</sup>
9. *style*: A forrásmondat MNSZ2-ből kinyert stílusregisztere. Mivel hapaxokról van szó, minden mondathoz pontosan egy stílusregiszter tartozik. Ez a következő értékeket veheti

<sup>40</sup> [https://github.com/kagnes/prevcons/raw/master/phono\\_patterns.xlsx](https://github.com/kagnes/prevcons/raw/master/phono_patterns.xlsx)

fel: *hivatalos, sajtó, spok* (beszélt nyelvi, általában rádióműsor leirata), *személyes* (fórum- és blogbejegyzések), *szépirodalom, tudományos*.

10. **region:** A régió, amelyből a forrásmondat származik (szintén az MNSZ2-ben elérhető adat). Lehetséges értékei: *erdélyi, kárpátaljai, magyarországi, szlovákiai, vajdasági*.
11. **autocluster:** Annak a klaszternek az azonosító száma, amelybe a szóképzéssel létrejött ige bemenete – hangzássémás ige esetén a teljes ige – került a szóbeágyazás során (erről bővebben ld. az 5.1. fejezetet). Mivel külön körökben történt a különféle igealkotási módokkal létrejött ige-klaszterezése, a számozás minden `inputtype` esetén 0-tól kezdődik. Az 1000 feletti értékek azt jelentik, hogy az adott ige nem produktív szóalkotással jött létre, vagy klaszteren kívül maradt.
12. **clause:** Az a tagmondat, amely az ige-kötős igét tartalmazza. A tagmondathatárok megállapítása Sass (2011: 34-37) módszerével történt.
13. **sent:** A teljes mondat, amely az ige-kötős igét tartalmazza. Ezen belül kétféle jelölés látható: az *<ncl>* (*new clause*, azaz tagmondathatár) és a *<q>* (*quotation mark*, azaz idézőjel). A két szóból álló, nagybetűvel kezdődő szósorokra rákerestem, és amennyiben tulajdonnevek voltak, anonimizáltam őket XY-ra. Ez alól akkor tettem kivételt, ha az ige-kötős ige-nek csak a tulajdonnév ismeretében volt igazán értelme, ilyen volt például a következő eset: *Ezek túladyendrészik magát Ady Endrét, [...]*

A PREVCONS-szal számos kérdés vizsgálható, például az, hogy melyik ige-kötő milyen stílusregiszterben van jelen a leginkább, vagy egyes ige-kötők mellett mennyire sokszínű vagy épp kötött a vonzatkeret. A legnagyobb ereje mégis az, amiről a fejezet elején is szó volt: segítségével az ‘ige-kötő – képző – vonzatkeret’ hármasok a konkrét ige-k kihagyásával vizsgálhatók. Erre látunk példát a 37. táblázatban.

A hármasok lekérdezése után, egy következő lépésben meg lehet nézni az ezekhez tartozó ige-ket. Célszerű a találatokat az automatikus klaszterszámuk alapján rendezni, hogy a szemantikailag hasonló ige-k egymás mellé kerüljenek. Ezt a 38. táblázat mutatja be.

A legjobban akkor látható, hogy a szóbeágyazás milyen jól ragadja meg a szavak szemantikai hasonlóságát, ha a klaszterszám szerinti rendezést a hangzássémás ige-ken végezzük el. Ekkor két-

igekötő	képző	vonzatkeret	db
<i>el</i>	-(s)Odik		196
<i>el</i>	-Vz(ik)	-t	107
<i>el</i>	-(V)(s)ít	-t	72
<i>el</i>	-Vz(ik)		69
<i>el</i>	-(s)Ul		41

**37. táblázat.** Az *el* igekötő mellett megjelenő öt leggyakoribb névszó → ige képző és vonzatkeret.

igekötő	ige	klaszter	tagmondat
<i>el</i>	<i>emdéefesedik</i>	1	és emdéeefesedett .
<i>el</i>	<i>balosodik</i>	2	ha teljesen elbalosodna ez a fórum .
<i>el</i>	<i>liberálisodik</i>	2	ugyanis reménytelenül elliberálisodtak ,
<i>el</i>	<i>pártosodik</i>	2	s elpártosodott ,
<i>el</i>	<i>ordasodik</i>	3	és elordasodik az ő hazája .
<i>el</i>	<i>simicskásodik</i>	3	Mielőtt még teljesen elsimicskásodnánk itt ,
<i>el</i>	<i>tőkésedik</i>	6	ha a volt elvtársak eltőkésednek ,
<i>el</i>	<i>magyarodik</i>	7	Óhatatlanul elmagyarodott ,
<i>el</i>	<i>amerikaisodik</i>	7	24 óra alatt elamerikaisodtak .
<i>el</i>	<i>európaisodik</i>	7	Amely mostanra szépen eleurópaisodott ,

**38. táblázat.** Az első tíz találat az *el-N-(s)Odik* szerkezetben szereplő igékre, az automatikus klaszterszámukkal és az igéket tartalmazó tagmondatokkal együtt.

féleképp csoportosulnak az igék. Az egyik lehetőség az, hogy tipikusan ki/mi adja ki az igével jellemzett hangot. Ezt látjuk az (54a) és az (54b) példában: az előbbi igéi folyadékok, az utóbbi állatok hangjára utalnak. A másik lehetőség az, hogy milyen hangja/hangulata van egy-egy cselekvésnek. Így például az (54c) igéi élőlény mozgását, az (54d) igéi élőlény nézését írják le.

- (54) a. bugyog csobog zubog csorog csurog csöpög csepeg ...  
b. csivog vartyog nyihog brekeg gágog mekeg rőfög ...  
c. szédeleg kóvályog biceg tipeg csoszog kocog cammog ...  
d. pillant pislog pillog pislant hunyorog mered ...

A PREVCONS lekérdezéseivel betekintést nyerhetünk az igekötők produktív kapcsolódási mintáiba. Ezenfelül szemantikai vizsgálatokra is alkalmas, ugyanis az igekötők főbb jelentései elválnak

benne a különböző vonzatkeretek, automatikus igei klaszterek mentén. Ezeknek a jelentéseknek a leírását, a köztük lévő szemantikai viszonyok feltérképezését mutatja be a következő fejezet.

## 5.4. Az igekötők ontológiája

Bizonyos igekötőknek csak egy produktív jelentése van: ilyen például az *újra*, amely mindig azt jelenti, hogy az igével kifejezett cselekvés megismétlődik (pl. *újradrótoz, újradizájno*).<sup>41</sup> Ennél viszont sokkal gyakoribb jelenség az, hogy egy igekötő poliszém, azaz több, akár öt-hat jelentéssel bír. Ezek a jelentések nem függetlenek egymástól, metaforikus kapcsolatok fedezhetők fel közöttük. Ilyen például az *agyon* történetileg első jelentése, a ‘megöl’ (pl. *agyonver, agyonlő*), és az ebből kialakult ‘túlzásba esve (károsodásig/kimerülésig)’ jelentés, ahol a cselekvés eredménye nem feltétlenül halál, de átvitt értelemben ahhoz közeli állapot (pl. *agyondőgönyöz, agyonsikál*). Nemcsak az egyes igekötők különféle jelentései jellemezhetők hálózatként, hanem az a kapcsolatrendszer is, amely különböző igekötőket köt össze az egyes jelentéseik mentén. Erre kézenfekvő példák az olyan igekötőpárok, amelyek irányjelentésükben ellentétesek, pl. *ki-be, fel-le*, vagy közel azonosak, pl. *be-bele, szét-szerte*. Ez a fejezet egy olyan kísérletet mutat be, amelynek célja az igekötők jelentéseinek és a jelentések közti viszonyoknak a hálózat-szerű ábrázolása a PREVCONS alapján, egy ontológia formájában.

Az ontológia informatikai értelemben olyan tudástár, amely entitásokból és azok relációiból áll. Entitásként jelennek meg az igekötők és a jelentések, relációként pedig három hagyományos jelentésviszony, a szinonímia, az antonímia, valamint a hiperonímia. Az általam létrehozott ontológia szigorúan véve nem teljesíti az ontológiák egyik alapvető kritériumát: nem használ formális leírónyelvet.<sup>42</sup> Ezt azonban nem tartom nagy hiányosságnak, mert a céloom nem a közvetlen nyelvtechnológiai alkalmazhatóság volt, hanem az, hogy nyelvészeti motivációval láttassam és magyarázzam az igekötők szemantikáját. Az ontológia reprezentációs formája egy síkbarajzolható gráf, amely PDF formátumban érhető el.<sup>43</sup>

Bár ilyen jellegű ontológia korábban nem készült a magyar igekötőkről, az igekötők szemanti-

<sup>41</sup> Az ilyen prefixumokat több szakirodalom nem tekinti igekötőnek, erről bővebben ld. Jakab (1976: 9-10) áttekintő táblázatát.

<sup>42</sup> Angol nyelvre több olyan formális ontológia is készült, amelyben kitüntetett figyelmet szenteltek az igekötő-szerű elemek, igei partikulák részletes szemantikai leírásának, erről bővebben ld. Bhatia et al. (2018), McShane et al. (2005).

<sup>43</sup> <https://github.com/kagnes/prevonto/raw/master/PreverbOntology.pdf>

kájának jelentős szakirodalma van. J. Soltész (1959) egy teljes könyvet szentel a hat ősi igekötőnek (*meg, el, ki, be, fel, le*), amelyben részletesen ismerteti a használati körüket, jelentésárnyalataikat és kialakulásuk történetét ún. „hagyományos nyelvészeti” szemlélettel. Ennek a hat igekötőnek a jelentésrendszerét tárgyalja Tolcsvai Nagy (2015: 72-86) is, de teljesen más megközelítéssel, a funkcionális kognitív szemantika keretében. Fontos még megemlíteni Kiefer és Ladányi (2000a: 483-513), valamint Ladányi (2007: 183-237) munkáit, amelyek a *be* ősi igekötő mellett részletesen elemzik két kevésbé régi (*szét, agyon*) és egy kifejezetten új (*tönkre*) igekötő jelentéseit és kapcsolódási mintáit. Az utóbbi mű kitér az igekötők épp formálódó, új jelentéseire is a neologizmusok vizsgálatán keresztül (Ladányi 2007: 258-283). Az említett szakirodalmak közös vonása az, hogy az igekötő-állományból néhány, reprezentatívnak ítélt példányt választanak, és ezeket rendkívül alaposan elemzik, ügyelve a jelentésárnyalatok megkülönböztetésére, ugyanakkor bemutatva a jelentések közti összefüggéseket.

Az általam készített ontológia eltér a fentebb ismertetett munkáktól a vizsgált igekötők halmazát, az elemzés módszerét és az eredmény közlési módját tekintve is. Arra törekedtem, hogy minél jobban lefedjem az igekötő-állományt, így az ontológia 56 igekötőt tartalmaz (amennyiben a *fel-föl* típusú alakváltozatokat külön számoljuk, ez a szám 62). Az igekötők kiválasztásának a szempontjait az 5.4.2. fejezet ismerteti. Az eltérő célkitűzés miatt az ontológia készítésekor nem törekedtem olyan részletgazdagságra, mint az említett szakirodalmak szerzői. Ami az elemzés módszerét illeti, minden következtetésem a PREVCONS adataira épül. Tehát csak produktív jelentéseket állapítok meg, és ezt elsősorban a leggyakrabban megjelenő vonzatkeret-minták alapján, másodsorban az igék automatikus klaszterszáma alapján teszem. Végül az eredményt sematikus formában, gráffal ábrázolom. Az ontológia mostani formájában „emberek számára” értelmezhető, de a jövőben könnyen átalakítható olyan formális leírássá, amellyel nyelvtechnológiai szempontból is használhatóvá válik.

A továbbiakban arról lesz szó, hogy milyen nehézségekbe ütközött az ontológia építése annak ellenére, hogy a munka jól strukturált adathalmazból indult ki. Ezután az ontológia részletes bemutatása következik, előbb az entitások (az igekötők és jelentéseik), majd a jelentések közötti szemantikai viszonyok leírásával.



#### 5.4.1. A poliszémiából adódó nehézségek

A fejezet bevezetőjében már felmerült, hogy az igekötők jelentései között metaforikus kapcsolatok figyelhetők meg. Néha olyan erősek és kiszámíthatóak az ilyen típusú összefüggések, hogy nehéz eldönteni, indokolt-e egyáltalán a külön jelentés felvétele. Előre megjósolható esetek például azok, amikor a konkrét irányjelentések nem csak térbeli irányokat jelölnek (pl. *visszamegy a boltba*), hanem időbelieket is (pl. *visszatekint a gyermekkorra*). Ez az absztrakciós lépés minden irányjelentéssel bíró igekötő esetében megfigyelhető, ezért az ontológiában az időbeli irányt nem választottam el a térbelitől.

Ugyancsak tendencia-szerűnek tűnik az irányjelentésű igekötők kapcsolata a kommunikációs igékkel. Ilyenkor az igekötő azzal járul hozzá a szerkezet jelentéséhez, hogy meghatározza a kommunikáció irányát (pl. *hátraság, utánaszól*). Jellemző viszont az is, hogy egy gyakori, lexikalizált ige mintájára sok új, jelentésében nagyon hasonló ige képződik (pl. a *beleszól* ‘kéretlenül szól hozzá valamihez’ jelentése mintájára a *belejópofiz, beleofof* egy topikba, chatbe, stb.). Itt azt az irányelvet követtem, hogy ha viszonylag sok – öt vagy annál több – hapax illeszkedett ugyanarra a mintára, akkor az igekötőhöz tartozó külön jelentésként adtam meg azt a lexikalizált igét, amelyet mintaadónak véltem. Így tehát a *bele* esetében megjelenik egy konkrét irányjelentés és a ‘beleszól valamibe’ jelentés is.

További kérdéseket vet fel az igekötők tömörítő képessége is. Meglehetősen gyakori eset az, hogy a tagmondatban szereplő egyik bővítményből képzünk igét, és ezt illesztjük egy lexikalizálódott igekötős igei szerkezetben az eredeti ige helyére. Ezek jellemzően inkább egyedi, kreatív szóalkotások eredményei, és nem indokolt produktív jelentésként felvenni őket az ontológiába. A jelenség jobb megértését szolgálják az (55) példasor mondatai. Ezekben a következő tömörítéseket látjuk: az (55a)-ban a *lenéptáncolom* a ‘néptáncolással ledolgozom’, az (55b)-ben a *ledinnyéztem* a ‘dinnyével lefizettem’, az (55c)-ben a *kiuhuskodtak* az ‘uhu szemeihez hasonlóan néztek kifelé’ rövidebb megfogalmazása.

- (55) a. Én inkább elmegyek, és **lenéptáncolom** a felesleget, és tanulok sokmindent.  
b. Azóta természetesen lekenyereztem, akarom mondani, **ledinnyéztem** őket.  
c. De a nagy merev szemek **kiuhuskodtak** az összetoldozott fényképből.

Nem kerültek be az ontológiába azok az esetek sem, amikor egy jelentést nem a konkrét igekötő szintjén, hanem inkább egy meghatározott igekötő-csoport szintjén volna érdemes megadni. Erre példa az ‘irányjelentéses igekötő + ige + *magát*’ konstrukció, amelynek a legáltalánosabb jelentése az, hogy ‘az igével jelzett cselekvéssel elérte, hogy az igekötővel jelzett helyre jusson’. Ezt a mintát figyelhetjük meg az (56) példásor mondataiban.

- (56) a. Sajnos mire kiértünk a boltból, a kisfiú **hazapityeregte magát**.  
b. Természetesen ép eszű ember nem kívánhatja, hogy az izraeli zsidókat tolják a tengerbe [...] ha már **odaerőszakoskodták magukat**, [...]  
c. [...] aki két év kihagyás után, tavaly ősszel állt újra munkába, „**visszaedzette magát** a csapatba”, [...]

Bizonyos igekötők esetében ez a konstrukció sajátos jelentést vehet fel, és ezt már a konkrét igekötő mellett szükséges feltüntetni. Ilyen például az ‘*el* + ige + *magát*’, amelynek a jelentése ‘nem szándékosan csinál valami hirtelent’ (pl. *elhorkantja magát*). Emellett természetesen létezik az általános jelentés is, így pl. a (56a) mondat mintájára működne az is, hogy *a kisfiú elpityeregte magát a cukrászdába*.

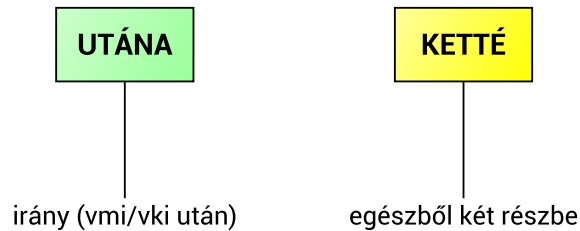
Végül fontos megemlíteni, hogy az aktuális jelentések mindig a szövegkörnyezettől függenek. Így például az (57a) mondatban a *letejez* jelentése az, hogy ‘tejjel véletlenül leönt’, míg az (57b)-ben látható *lespagettiz* azt jelenti, hogy ‘tévesen spagettinek nevez’, bár a *tej* és a *spagetti* szemantikai kategóriája közel azonos, és a vonzatkeret is egyezik.

- (57) a. Reggelijét befejezte // Nyakkendőjét **letejezte** // [...]  
b. [...] amit a hozzáértők simán „**lespagettiznének**”, hiszen ránézésre egyszerű paradicsomos tésztának látszik.

Látjuk tehát, hogy az egyes jelentések megállapítása nem egyszerű, és még az irányelvek gondos meghatározása mellett is befolyásolja az egyéni intuíció. Ha több ember kapta volna a feladatot, hogy a PREVCONS alapján készítsen ontológiát, valószínűleg nem született volna két egyforma végeredmény.

### 5.4.2. Ontológiai entitások: igekötők és jelentések

Az ontológia kétféle entitásból áll – igekötőkből és ezek jelentéseiből –, amelyeknek a gráfban csomópontok felelnek meg (ld. 17. ábra).



**17. ábra.** Két egyszerű példa az entitásokra: az *utána* és a *ketté* igekötők a jelentéseikkel. A zöld háttérszín azt mutatja, hogy az igekötőnek van irányjelentése, a sárga háttér ennek a hiányát jelzi.

Csak azokkal az igekötőkkel foglalkoztam, amelyeknél a PREVCONS-ban megfigyelhető az, hogy névszóból képzett igéhez kapcsolódnak. A döntésemnek gyakorlati oka volt: a kutatás előrehaladtával egyértelművé vált, hogy a névszói bemenetek szemantikai kategóriái mondanak el a legtöbbet az igekötő lehetséges jelentéseiről, ezeknek a vizsgálatával lehet a leghatékonyabban haladni. Miután összevontam az alakváltozatokat,<sup>44</sup> 56 elemzendő igekötő maradt. Természetesen a későbbiekben további igekötőkkel is érdemes bővíteni az ontológiát.

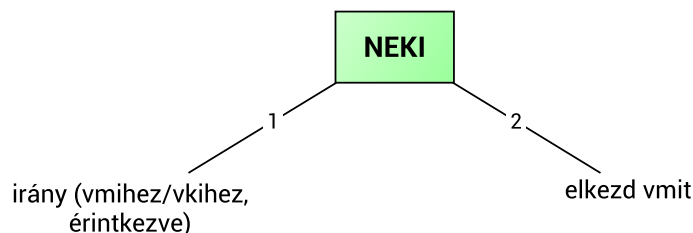
A jelentések kapcsán hangsúlyozni kell, hogy csak a PREVCONS alapján kimutatható – produktív – jelentéseket vettem számba. Ez az oka annak, hogy a *meg* jelentés nélkül szerepel az ontológiában. Az eredeti irányjelentése ('hátra, vissza') ugyanis ma már csak néhány igével mutatható ki, pl. *megad* 'visszaad', *megfordít* 'visszafordít', *megtér* 'visszatér'.<sup>45</sup> A *valami*, *valaki* stb. szavakat a lexikográfiában megszokott módon rövidíttem: *vmi*, *vki*. A jelentések leírásánál törekedtem a minél általánosabb megfogalmazásra, és előfordult, hogy részben önmagával definiáltam egy szót, ha a gondosabb definíció túl bonyolult lett volna (pl. *utána* = *vmi/vki után*).

A 17. ábrán látható, hogy az igekötőt és a jelentést fekete vonalak – a gráfban élek – kötik össze egymással. A legegyszerűbb és legritkább esetben az igekötőhöz egyetlen jelentés tartozik, ekkor az összekötő él nem kap semmilyen címkét. Amikor egy igekötőhöz több jelentés kapcsolódik, az élek számozottak (ld. 18. ábra). A számozási sorrend nem implikál semmit, tehát az 1-es nem jelenti azt, hogy a hozzá tartozó jelentés a legrégebbi vagy legtipikusabb. Ez csak a jelentésekre történő

<sup>44</sup> Az alakváltozatok a következők: *be* – *bé*, *bele* – *belé*, *fel* – *föl*, *felül* – *föül*, *rá* – *reá*, *tele* – *teli*.

<sup>45</sup> További példákhoz ld. J. Soltész (1959: 30-33).

rövid hivatkozást szolgálja:  $neki_1$  = ‘irány (valamihez/valakihez, érintkezve)’ (pl. *nekidöccen az ablaknak*),  $neki_2$  = ‘elkezd valamit’ (pl. *nekidühödik az írásnak*). Az igeekötők jelentéseit minden esetben példamondatok is illusztrálják, ezeknek a listája a II. számú mellékletben látható, valamint online is elérhető.<sup>46</sup>



**18. ábra.** A *neki* igeekötő két produktív jelentése.

Az éleken nem csak számok szerepelhetnek, hanem szöveges tartalom is. A + magát címke azt jelenti, hogy az adott jelentés mindig együtt jár a *magát* vonzat megjelenésével. Ezt látjuk például a ‘*ki* + ige + *magát*’ szerkezetnél, amelynek jelentése ‘eleget csinál valamiből’, ld. az (58) példasort.

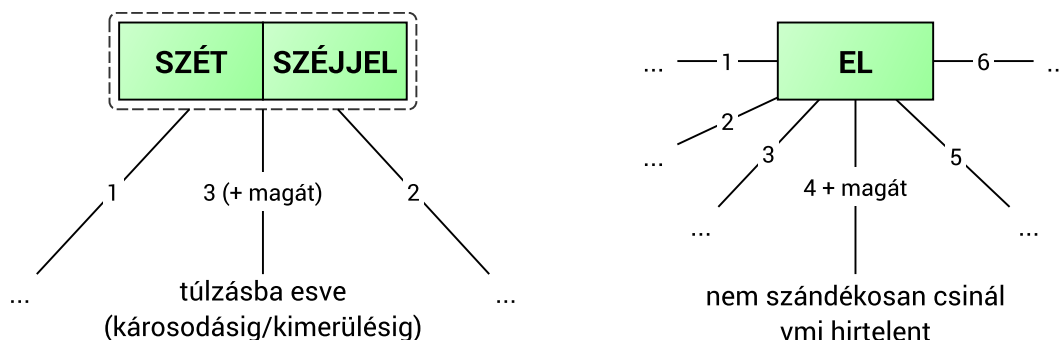
- (58) a. Na polgártársak, miután virtuálisan **kisöröztük magunkat**, döntsük el mi lesz ma este!
- b. S amikor mindenki **kiteniszezte magát**, újra képbe jött a focimeccs, [...]
- c. Külvárosi fiatalok, miután **kimolotovkoktélózták magukat** Párizsban, egy tanyasi motelben húzzák meg magukat.

Ha a + magát címke zárójeles, akkor az adott jelentés gyakran, de nem kötelezően jár együtt a *magát* vonzattal. Ez figyelhető meg például a *szét*, *széjjel* igeekötők ‘túlzásba esve (károsodásig/kimerülésig)’ jelentésénél, amelyet az (59) példasor szemléltet.

- (59) a. Hosszú, hosszú hétvége következik, rengeteg idő lesz kirándulni, felállni a gép elől - aki pedig nem akar az **szétkockulhatja magát**.
- b. Otthon mindenki **széjjelmulatja magát**, [...]
- c. Angolon meg nem hallok meg semmiccse, mer **szétwalkmanezem a fületem**.
- d. Egyébként ha **széjjeloffoljátok itt a topicot**, közétek lövetek.

<sup>46</sup> [https://github.com/kagnes/prevonto/raw/master/Preverb\\_senses\\_examples.docx](https://github.com/kagnes/prevonto/raw/master/Preverb_senses_examples.docx)

Szólnak érvek amellett is, hogy itt nem egy, hanem két jelentésről van szó, mivel az akcióminőségek eltérők. A *magát* vonzat esetében exhaustív ('kifáradásig, kimerülésig'), az egyéb tárgyesetes vonzat esetében intenzív ('nagyon, túlságosan') akcióminőség jelenik meg (Ladányi 2007: 221). Én nem választottam el ezt a két lehetőséget, hanem egy összevont, általánosabb jelentést fogalmaztam meg, mert azt tapasztaltam, hogy ha egy igekötőnél az egyik megvan, előkerül a másik is. A 19. ábra szemléltet minden lehetséges élcímké-típust.



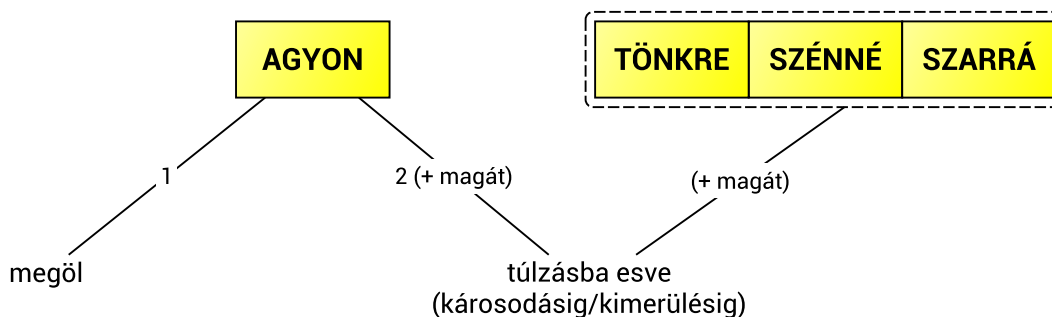
**19. ábra.** A *magát* vonzat megjelenése a *szét*, *széjjel* és *el* igekötőknél. A *szét* és *széjjel* körüli szaggatott vonal azt jelzi, hogy a két igekötő minden jelentésében szinonim, mindössze stílusbeli különbség van közöttük. Csak a tárgyalás szempontjából releváns jelentéseket tüntettem fel az ábrán.

### 5.4.3. Szemantikai relációk

Az ontológia háromféle jelentésviszonyt tartalmaz, ezek a következők: szinonímia (rokonértelműség), antonímia (ellentétes jelentés) és hiperonímia (hierarchikus viszony, amelyben az egyik szó – hiperonima – jelentése magába foglalja a másik szó – hiponima – jelentését). A gráfban ezek élekként jelennek meg, amelyek a jelentéseket kötik össze egymással. A továbbiakban mindhárom jelentésviszonyt részletesen, példákon keresztül veszem sorra.

A szinonímiát az jelzi, ha egy jelentésből több igekötőbe vezetnek élek. Ezt látjuk a 20. ábrán a 'túlzásba esve (károsodásig/kimerülésig)' jelentés esetében. Az ábrán szereplő *agyon*, *tönkre*, *szénné* és *szarrá* igekötőkön túl megtaláljuk ezt a jelentést a *szét*, *széjjel* és *össze* igekötőknél is.

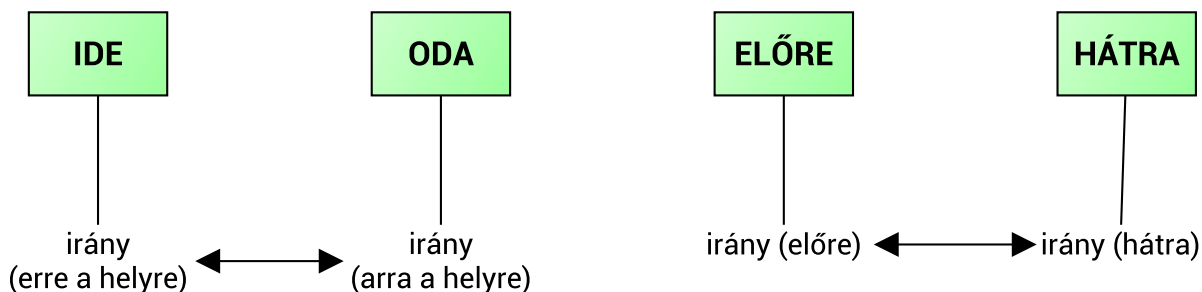
Ezt természetesen nem úgy kell érteni, hogy például a *szénné* bármikor kiválthatja az *agyon* igekötőt annak a második jelentésében anélkül, hogy a mondatban jelentésváltozás történne. Egy kis konnotáció- vagy stílusbeli különbség mindig adódik, ehhez ld. a (60) példasort. A *szénné* és a *szarrá* szlengben fordulnak elő, a *tönkre* kissé informális – bár egyre gyakrabban jelenik meg hivatalos szövegekben is –, az *agyon* pedig semleges, bármilyen stílusregiszterben megtaláljuk.



20. ábra. A szinonímia megjelenítése az ontológia grájfjában: egy jelentésből több igekötőbe fut él.

- (60) a. [...] bár a második itteni hetem végére az idegeim meglehetősen **agyoncincálód**nak.
- b. Én **tönk**reröhögtem magam ezen a hozzászóláson.
- c. [...] ő meg legalább **szénnéizgul**ja magát az úton, vagy az autójában.
- d. Csórikám, ezzel **szarrá**kereshetné magát a nyugati világban [...]

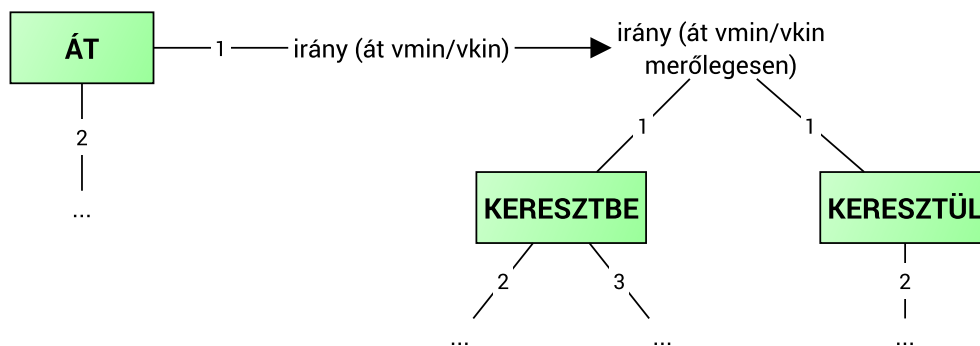
Az antonímiát a jelentések közé húzott kétirányú nyíl jelzi (ld. 21. ábra). Valószínűleg ez a legkönnyebben felismerhető szemantikai viszony az anyanyelvi beszélők többsége számára. Tipikusan a konkrét irányjelentéseknél jelenik meg.



21. ábra. Az *ide* és *oda*, valamint az *előre* és *hátra* jelentései közt fennálló antonim viszony, amelyet kétirányú nyíl jelez.

A hiperonímia jelzésére az egyirányú nyíl szolgál, amely a fölérendelt, általánosabb jelentésből az alárendelt, specifikusabb jelentés felé mutat (ld. 22. ábra). Ilyen viszony figyelhető meg például az *át* és a *keresztbe*, *keresztül* igekötők irányjelentései között, amelyeknek a különbségét a (61) példaszor mutatja be.

A (61a)-ban az *áthajókáztak* ige alapján még nem tudjuk, hogy az alanyok a Bering-szorost szélteben szelték át, vagy hosszában hajókáztak rajta végig. Csak a második tagmondatból derül ki, hogy az utóbbi értelmezés a helyes. A *keresztbe*, *keresztül* irányjelentése specifikusabb az *át*-



**22. ábra.** A hiperonímiát egyirányú nyíl mutatja: az *át* irányjelentése magába foglalja a *keresztbe* és *keresztül* irányjelentését. Csak a tárgyalás szempontjából releváns jelentéseket tüntettem fel az ábrán.

énál, mert azt is magukba foglalják, hogy a mozgás iránya a (61b)-ben az útra, a (61c)-ben az ágyra merőleges. Ez a különbség valószínűleg nem egyformán van jelen az anyanyelvi beszélők intuíciójában, mert néhány korpuszpélda azt sejteti, hogy az *át* és *keresztül* irányjelentése kváziszinonim, és csak a *keresztbe* fejezi ki mindig a merőlegességet.

- (61) a. [...] 1995-ben **áthajókáztak** a Bering-szoroson, majd hosszú arktikus körút után 2000-ben Izland vizeire is eljutottak.
- b. [...] amikor hirtelen **keresztberohant** előtte az úton egy kondányi vaddisznó.
- c. Gyurica **keresztülvetődött** az ágyon, ruhástul.

Az *át* és *keresztül* esetéhez hasonló, apróbb bizonytalanságok ellenére törekedtem arra, hogy csak akkor jelöljek szemantikai relációt, ha a PREVCONS példái alapján mindig fennállni látszott a kapcsolat bizonyos jelentések között. Emiatt például nem került antonima-viszonyba az *egybe* és a *ketté*. Az utóbbi ugyanis azt jelenti, hogy egy egységből pontosan kettő lesz, és ennek az ellentéte az volna, hogy kettőből egy egység lesz. Az *egybe* ennél általánosabb: kettő vagy több rész egyé válását jelenti, ld. a (62) példáit.

- (62) a. Néhány téglasor után jó bőven megöntözték a falakat, a borkősav felforrt, a mész megoltódott, és a téglasorok szinte **egybeégtek**.
- b. Összevissza kutyult levél lett belőle, ahogy újra **egybebookmarkoltam** széthulló gondolataimat, [...]

Zárásképpen fontosnak tartom egy példasorral illusztrálni, hogy a szemantikai viszonyok miatt az egyes jelentések szintjén értelmezhetők, és miért nem az igeekötők szintjén. Létezik ugyanis olyan eset, amikor ugyanannál az igeekötőpárnál megjelenik az antonímia és a szinonímia is, attól függően, hogy melyik jelentésüket vizsgáljuk – ilyen az *össze* és a *szét*. Ezek konkrét irányjelentésükben ellentétesek: az *össze* jelentése ‘több helyről egy helyre’, ld. a (63a) és (63b) példákat, a *szét* igeekötőé pedig ‘egy helyről több helyre’, ld. a (63c) és (63d) példákat.

- (63) a. [...] sokkal erősebb kötelék, mint amivel a világirodalom nagyjait **összelasszózt**a.
- b. A zsúfolt folyosón idegenül, de mégis egyetértő mosollyal pillantgattuk egymást, ahogy **összenyomakodtunk**.
- c. - **Széttologatja** a bútorokat.
- d. Bizony, össze volt az már pöndörödve, de az ember szépen **szétsimítgatta**, [...]

Az *össze* és a *szét* ugyanakkor közel azonos értelmű egy absztraktabb jelentésüket tekintve: ‘túlzás-ba esve (károsodásig/kimerülésig)’. A (64) példasorban a két igeekötő fel is cserélhető egymással úgy, hogy a mondat jelentése nem vagy csak árnyalatnyit változik.

- (64) a. Most jól **összeofftopikoztuk** XY rovatát.
- b. A névből sejtetődő, hogy a vírus **összebabrálja** a menüt, [...]
- c. **Szétgusztustalankodtad** az esküvőjét is, a sok dopeman meg tapsolt hozzá [...]
- d. **Szétkarcoltak** egy hajótestet, elégettek néhány kötelet, lökdösték egymást, [...]

Annak érdekében, hogy az egyes jelentések közötti kapcsolatok megállapítása minél objektívabb legyen, a jövőben érdemes lehet pszicholingvisztikai kísérletekkel, anyanyelvi beszélők nagy csoportján vizsgálni az ontológiában eddig rögzített jelentésviszonyok érvényességét.

## 5.5. Összegzés

Ebben a fejezetben egy sokoldalú témát vizsgáltam, az igeekötők produktív kapcsolódási mintáit. Az adatvezérelt kutatás alapját 21 038 igeekötős hapax képezte. Ahhoz, hogy ezzel a témával érdemben foglalkozni lehessen, több nyelvészeti részterületet is be kellett vonnom a kutatásba. A



produktív igealkotási módokhoz fonológiai és morfológiai, a vonzatkeretek kinyeréséhez szintaktikai, az igekötők és igecsoportok jelentéseihez szemantikai vizsgálatra volt szükség.

A fejezet első részében azzal a kérdéssel foglalkoztam, hogy hogyan hozunk létre új igéket. Az igealkotás számos módja közül azt a hármat vizsgáltam, amely számottevően volt jelen az adatban. Ezek a névszói és igei bemenetű igeképzés – és ezeken belül is a produktív képzők –, valamint a hangzásséma-alapú igealkotás. Az utóbbit Benő és Szilágyi N. (2015) nyomán elkülönítettem az igeképzésektől, és a leírását is máshogyan végeztem el (szótő-képző bontás helyett szótagszerkezeti sémákkal). Ennek a munkának a jelentősége nem új felfedezésekben érhető tetten, hanem a meglévő ismeretek rendszerezésében, gyakorisági adatokkal való bővítésében, és pár ponton a felülbírálásában (például annak kapcsán, hogy a -(s)Ul képző produktív-e vagy sem). Végül megvizsgáltam az igealkotási módok arányait az igekötős hapaxok halmazán, és kimutattam, hogy a legjelentősebb a névszói bemenetű igeképzés, amely a hapaxok 35,2%-át adja, viszont mindössze 62 – az alakváltozatok összevonása után 56 – igekötő kapcsolódik így képzett igéhez.

A fejezet második részében bemutattam a PREVCONS táblázatot, amelyet az igekötős szerkezetek vizsgálatára hoztam létre. Ez az erőforrás a kutatás alapjául szolgáló hapaxokat tartalmazza a morfológiai szerkezetükre, vonzatkeretükre, szemantikájukra és szövegbeli előfordulásukra vonatkozó információval együtt. A PREVCONS jelentőségét három dologban látom. Először is abban, hogy már nagyon jó erőforrások készültek a gyakori igekötős igékre és ezek vonzatkereteire – ld. Sass et al. (2010) és Sass (2011) –, nem készült viszont olyan, amely hasonló információt adna a hapaxokról (ezek tokenszámukat tekintve kevesen, de típusszámukat tekintve nagyon sokan vannak). Másodszor, a PREVCONS-ban összegyűjtött minták lehetővé teszik a lexikai erőforrások hatékony bővítését. Végül, ami a legfontosabb: a PREVCONS a magyar nyelvet kutatók és tanítók számára jó áttekintést és érdekes adatokat szolgáltat az igekötős szerkezetekről.

A harmadik részben arra tettem kísérletet, hogy hálózat-szerűen, gráfként ábrázoljam az igekötők jelentéseit, és a jelentések közt fennálló viszonyokat. Ebből a célból létrehoztam egy ontológiát, amelyben entitásként jelennek meg az igekötők és a jelentések, relációként pedig három hagyományos jelentésviszony, a szinonímia, az antonímia és a hiperonímia. Az ontológia 56 igekötő szemantikáját írja le, és tervezem a további bővítését, javítását is. Jó néhány intuitív vonása ellenére ezt tartom a disszertáció egyik legfontosabb eredményének, elsősorban a benne megjelenő új szemlélet miatt.

## 6. Összefoglalás, új tudományos eredmények

„El kell, hogy engedj!”

(V'Moto-Rock)

A disszertációmban a magyar nyelv igezős szerkezeteit vizsgáltam. A három fő témám a következő volt: (1) az igezők állományának meghatározása, (2) az igezős szerkezetek típusainak és szintaktikai jellemzőinek leírása, valamint (3) az igezők produktív kapcsolódási mintáinak feltárása. A munkám során korpuszvezérelt megközelítést alkalmaztam, amelynek a lényege az, hogy a kutatás nagy mennyiségű szöveganyag automatikus elemzéséből indul ki, és elsődleges célja az introspekcióval nem feltárható jelenségek kimutatása.

A 2. fejezetben a kutatás módszertani hátterét mutattam be. Körüljártam a korpuszvezéreltség fogalmát, majd bemutattam azt a két erőforrást, amelyet a dolgozat egészében felhasználtam. Az első az MNSZ 2.0.4 korpusz (Oravecz et al. 2014) módosított verziója, amelyből kiszűrtem a verseket, a duplumokat és az értelmetlen elemzés nélküli mondatokat. A másik a PREVLEX táblázat, amelyhez az első tézisem kapcsolódik:

1. Az MNSZ2 korpusz felhasználásával létrehoztam a PREVLEX táblázatot, amely az igezős igeik jelenleg legbővebb (53 503 szavas), manuálisan ellenőrzött, nyílt hozzáférésű táblázata. Részét képezik a hapaxok – egyszer előforduló szavak –, valamint azok a szavak is, amelyek a korpuszban UNKNOWN címkével jelennek meg. Az igezős igeik mellett a korpuszban mért tokengyakoriságuk is szerepel.

A 3. fejezetben az igező-állomány meghatározásának problémájával foglalkoztam. A fő kérdésem az volt, hogy mely lexikai elemeket soroljuk az igezők közé, és mi alapján végezzük a besorolást? Abból a megfigyelésből indultam ki, hogy az igezők és egyéb, pusztán névszói igemódosítók között nem húzható éles határ. Ehhez a felfogáshoz leginkább a prototípus-elmélet illeszkedik, amelynek keretében megállapítottam az igezőség morfoszintaktikai jegyeit, és a korpuszméréseim alapján meghatároztam a jegyek értékét minden igező-gyanús szó esetében. Az adatgyűjtés eredményéből Igező × Jegy mátrixokat hoztam létre, amelyek elsősorban a jegyértékek ábrázolásában különböznek. A mátrixok alapján megvizsgáltam az egyes jegypárok közötti

korrelációkat. Végül bemutattam négy módszert, amelyek alkalmasak lehetnek az igekötők jegy-alapú osztályozására. A módszerek összehasonlítása után a halmazos módszer mellett döntöttem, amely alapján négy, leginkább a tárgyalást segítő kategóriát neveztem meg: a prototipikus, a centrális, a félperiférikus és a periférikus igekötőket. Főbb eredményeim a következők:

2. Megállapítottam és korpuszon kimértem 11 olyan jegyet, amellyel az igekötők jellemezhetők – ugyanakkor jeleztem, hogy nem minden jegy egyformán fontos. A korpuszmérések eredményét nyílt hozzáférésű Igekötő  $\times$  Jegy mátrixokban rögzítettem.
3. Az abszolút gyakoriságot tartalmazó Igekötő  $\times$  Jegy mátrixon kiszámoltam az egyes jegypárok közötti korrelációkat. Ezáltal mennyiségi mutatókkal jellemeztem a grammatikalizáció folyamatát az igekötők esetében. A produktivitás erős pozitív korrelációban áll a gyakorisággal, míg a szótagszám és a szerkesztettség ezekkel negatív korrelációt mutat – a gyakori és produktív igekötők jellemzően rövidek és morfológiai szempontból bonthatatlanok. A korreláció-számítást bináris adaton is elvégeztem, és kimutattam, hogy a jegyek közti alapvető összefüggések nem változnak attól, ha a gyakoriságra vonatkozó információt elhagyjuk. Ezt azzal magyaráztam, hogy a gyakoriság történetileg olyan szorosan együtt jár más jegyekkel – éppen a grammatikalizációs folyamat révén –, hogy akkor is érezhető a hatása az igekötőség jegyeiben, ha külön jegyként nem vesszük figyelembe.
4. Az ún. halmazos módszer használatával felállítottam egy tipikalitási skálát, amelynek egyik végpontján a prototipikus igekötők, másik végpontján a periférikus igekötők helyezhetők el. A skála a következőképp alakul ki: minden jegyet halmazként értelmezünk, az adott jeggyel bíró igekötőket pedig az adott halmaz elemeiként. Ezután aszerint sorrendezzük az igekötőket, hogy hány halmaz metszetében jelennek meg. A skála ezután a kutatási igényei szerint kategóriákra bontható.

A 4., mondattani tárgyú fejezetben a célom a következő kérdés megválaszolása volt: Milyen mintázattípusokat mutatnak az igekötős szerkezetek, mikor és mennyire távolodhat el az igekötő az igétől, igenévtől, igei származéktól? Először egy szinkrón korpuszvizsgálatot mutattam be, amely az MNSZ2 korpusz 20-21. századi szövegein alapult. A következő szerkezeteket vettem sorra és jellemeztem az igekötők eloszlási mintái, valamint a beférkőző szavak szempontjából: finit ige,

infinitívus, határozói igenév, különféle melléknévi igenév-típusok és deverbális főnevek. Ezután egy diakrón korpuszvizsgálat eredményeiről számoltam be, amely során a prototipikus igekötők mondatbeli pozícióját mértem ki a finit igéhez viszonyítva az ómagyar kortól napjainkig terjedő szöveganyagon. A kutatásom főbb eredményei a következők:

5. Kimutattam, hogy a prototipikus igekötők relatív gyakoriságukat tekintve közel maradnak a finit igéhez, míg a periférikusabb igekötők jobban eltávolodhatnak. Ehhez kapcsolódva megneveztem két olyan tényezőt, amely valószínűleg hatással van arra, hogy a hátravetett igekötő mennyire távolodik el a finit igétől. Az egyik a szöveg szerkesztettségi foka: szerkesztetlen – főleg szóban elhangzott – szövegben nagyobb annak a valószínűsége, hogy az igekötő távol kerül a finit igétől, mint gondozott, szerkesztett szöveg esetében. A másik az ige utáni mondatösszetevők, köztük a hátravetett igekötő fonológiai súlya. A méréseim alapján az a tendencia rajzolódik ki, amelyet a növekvő összetevők törvényének vagy Behaghel-törvénynek is neveznek, és É. Kiss (2007) szerint a magyarban a finit ige utáni mondatszakaszra érvényes: a rövidebb összetevő megelőzi a hosszabbat, ha ezt szintaktikai szabály nem gátolja. Ez azzal is egybevág, hogy az egy szótagú, prototipikus igekötők ritkábban távolodnak el a finit igétől, mint a több szótagú, periférikusabb igekötők.
6. Korpuszméréssel igazoltam, hogy az infinitívusztól az igekötője preverbálisan nagyon messzire kerülhet, de csak akkor, ha segédige-szerű elem – elsősorban finit ige – férkőzik be a szerkezetbe, és az igekötő az ehhez tartozó igemódosítói pozíciót foglalja el. Ez alátámasztja azt a megfigyelést, hogy bár az igekötő az infinitívussal alkot szótári egységet, szórendi és prozódiai szempontból szorosabban kapcsolódik a segédige-szerű elemhez (Kálmán C. et al. 1989). Megállapítottam, hogy ezzel rokon szintaktikai viselkedése van a létigével passzív szerkezetet alkotó határozói igenévnek is. Ennek az igekötője akkor kerül távoli preverbális helyzetbe, ha a létige előtti igemódosítói pozíciót foglalja el.
7. Megállapítottam, hogy a határozói igenév igekötője csak akkor vethető hátra, ha az igenév állapot- vagy módhatározói szerepet tölt be. A létigével alkotott passzív konstrukcióban az igekötő mindig preverbális.
8. Kimutattam, hogy az *el kell, hogy mondjam* típusú szerkezet nemcsak finit ige igekötőjének a kiemelésével jöhet létre, hanem infinitívusok és határozói igenevek igekötői is kiemelhe-

tők fölérendelt tagmondatba (pl. *hozzá kell, hogy tudjunk férni és össze kell, hogy legyen kötvé*). A három szerkezetben az a közös, hogy a mellékmondatban időjeles ige van, amely az infinitívusz esetében egy segédige-szerű elem, a határozói igenév esetében a létige.

9. A melléknévi igenevek kapcsán a következő megfigyeléseket tettem: (1) Ha -hAtÓ végződésű igenév állítmányi helyzetben van, és a tagmondatban nem hangzik el finit ige, akkor a -hAtÓ teljes mértékben átveszi ennek a helyét. Az igekötője ugyanúgy elválhat, mint egy finit igéé, az igekötők disztribúciója is hasonlóan alakul, emellett a szerkezetbe férkőző szavakat tekintve is párhuzam látszik a finit igei szerkezettel. Mindez az ellen szól, hogy a -hAtÓ-t melléknévképzőnek tekintsük – ahogy például Kiefer (2003a) teszi –, mivel az ezzel képzett szó nagyon erős igei tulajdonságokat őriz. (2) Az általános szakirodalmi véleménynel szemben kimutattam, hogy az -AndÓ végződésű melléknévi igenevek állhatnak állítmányi helyzetben, és ekkor az igekötőjük is hátravethető, bár a jelenség kétségtelenül ritka (az esetek 1,72%-ára jellemző, ez a vizsgált korpuszban 1519 példát jelent).
10. A diakrón vizsgálat keretében kimutattam a hátravetett igekötőt tartalmazó, nem-semleges mondatok arányának monoton növekedését az ómagyar kortól napjainkig. Ez a tendencia főként azzal magyarázható, hogy a fordított szórendű tagadás fokozatosan teret nyert az archaikus, megszakított szórendű tagadással szemben, ami a korpuszadatok alapján szintén jól nyomon követhető.

Az 5. fejezet központi kérdése az volt, hogy hogyan írhatjuk le az igekötők produktív kapcsolódási mintáit, és ez alapján milyen megállapítások tehetők az igekötők szemantikájáról? Ennek megválaszolására olyan módszert dolgoztam ki, amely ‘igekötő – képző – vonzatkeret’ hármasok (pl. *el-X-elődik valami, tönkre-X-el valamit*) korpuszvezérelt vizsgálatára épül. Ismertettem az igealkotás három leggyakoribb módját: a névszói és az igei bemenetű igeképzést, valamint a hangzásséma-alapú igealkotást. Ezután bemutattam a PREVCONS táblázatot, amely 21 038 igekötős igei hapaxot tartalmaz, és a fentebb említett hármasok kereshetőségével lehetővé teszi a produktív kapcsolódási minták feltárását. Végül beszámoltam egy olyan kísérletemről, amelynek célja az igekötők jelentéseinek és a jelentések közti viszonyoknak a hálózat-szerű ábrázolása volt a PREVCONS alapján, egy ontológia formájában. A fejezethez kapcsolódó téziseim a következők:

11. Saját algoritmust írtam a hangzássémára illeszkedő igék azonosítására, amely egyúttal sématisztításokba is rendezte az érintett igéket. Ezáltal kimutattam, hogy bár a nyelvészeti szakirodalom nagyon kevés jelentőséget tulajdonít az igealkotás ezen módjának, a hangzássémás igék aránya nem elhanyagolható: ezek jelentik az igekötős igei hapaxok közel egytizedét (9,4%).
12. Kimutattam, hogy az új igekötős igék létrehozásában a névszói bemenetű igealkotás a legjelentősebb. Ez a hapaxok 35,2%-ánál mutatható ki. Ugyanakkor mindössze 62 – az alakváltozatok összevonása után 56 – igekötő kapcsolódik így képzett igehez.
13. Létrehoztam a PREVCONS táblázatot, amely az igekötős szerkezetek vizsgálatára alkalmas, nyílt hozzáférésű kutatóeszköz. 21 038 igekötős igei hapaxot tartalmaz a morfológiai szerkezetükre, vonzatkeretükre, szemantikájukra és szövegbeli előfordulásukra vonatkozó információval együtt.
14. Létrehoztam egy nyílt hozzáférésű ontológiát, amely 56 igekötő jelentéseit és a jelentéseik közötti viszonyokat jeleníti meg. Entitásként szerepelnek benne az igekötők és a jelentések, relációként pedig három hagyományos jelentésvizony, a szinonímia, az antonímia és a hiperonímia. Az ontológiát síkbarajzolható gráfként ábrázoltam.

A felsorolt tézispontok alapján látható, hogy a munkám elméleti és gyakorlati szempontból is hozott új tudományos eredményeket. A gyakorlati haszna elsősorban a PREVLEX és a PREVCONS létrejötte, amelyek önmagukban is értékes erőforrások. A további szerepük lehet az, hogy elősegítsék a morfológiai elemzők javítását és fejlesztését, emellett a PREVCONS-ban összegyűjtött minták más lexikai erőforrások hatékony bővítését is lehetővé teszik.

A kutatás elméleti téren számos olyan tendenciát, jelenséget mutatott ki, amelyek korpuszvezérelt módszer hiányában észrevétlenül vagy sejtés szintjén maradtak volna. Emellett felhívta a figyelmet olyan jelenségekre, amelyek egyáltalán nem ritkák, mégis csak kevés nyelvész érdeklődésére tarthattak eddig számot (ilyenek például a hangzássémákra illeszkedő igék).

Bízom abban, hogy az adatvezérelt szemlélet egyre inkább teret tud nyerni a nyelvtudományban, és a dolgozatomban leírt módszerek, ötletek hasznosnak bizonyulnak nemcsak az igekötős szerkezetek, hanem más nyelvi jelenségek vizsgálatában is.

## Hivatkozások

- Alberti Gábor 2006. *A magyar mondattan elmélete és gyakorlata generatív megközelítésben III. A háttérelmélet*. Pécsi Tudományegyetem. Pécs.
- Baayen, Harald 1989. *A Corpus-Based Approach to Morphological Productivity (Statistical Analysis and Psycholinguistic Interpretation)*. Doktori értekezés. Centrum voor Wiskunde en Informatica. Amsterdam.
- Baayen, Harald 2009. Corpus linguistics in morphology: morphological productivity. In: Lüdeling, Anke – Kytö, Merja (szerk.): *Corpus Linguistics. An international handbook*. Mouton De Gruyter. Berlin. 900–919.
- Behaghel, Otto 1932. *Deutsche Syntax IV*. Carl Winters. Heidelberg.
- Benő Attila – Szilágyi N. Sándor 2015. Hangzásséma és motiváltság a hangutánzó és hangulatfestő igéink körében. In: Kádár Edit – Szilágyi N. Sándor (szerk.): *Motiváltság és nyelvi ikonicitás*. Erdélyi Múzeum-Egyesület (EME). Kolozsvár. 43–57.
- Bentz, Christian – Ferrer-i Cancho, Ramon 2016. Zipf’s law of abbreviation as a language universal. In: Bentz, Christian – Jäger, Gerhard – Yanovich, Igor (szerk.): *Proceedings of the Leiden Workshop on Capturing Phylogenetic Algorithms for Linguistics*. University of Tübingen, online publication system. Tübingen.
- Bhatia, Archana – Teng, Choh Man – Allen, James F. 2018. Identifying senses of particles in verb-particle constructions. In: Markantonatou, Stella – Ramisch, Carlos – Savary, Agata – Vincze, Veronika (szerk.): *Multiword expressions at length and in depth: Extended papers from the MWE 2017 workshop*. Language Science Press. Berlin. 61–86.
- Cooper, William E. – Ross, John Robert 1975. World order. In: Grossman, Robin E. – San, James L. – Vance, Timothy J. (szerk.): *Papers from the Parasession on Functionalism*. Chicago Linguistic Society. Chicago. 63–111.
- Cser András 2017. *Morfológia*. Pázmány Péter Katolikus Egyetem Bölcsész- és Társadalomtudományi Kar. Budapest – Piliscsaba.

- Dömötör Adrienne 2009. Az ó- és középmagyar kori nyelvhasználat morfológiailag elemzett adatbázisa. In: Fazakas Emese – Juhász Dezső – Terbe Erika – Tamás Csilla (szerk.): *Tér, idő, társadalom és kultúra metszéspontjai a magyar nyelvben*. ELTE Magyar Nyelvtörténeti, Sociolingvisztikai, Dialektológiai Tanszék – Nemzetközi Magyarságtudományi Társaság. Budapest – Kolozsvár. 65–80.
- Dressler, Wolfgang U. 2003. Degrees of grammatical productivity in inflectional morphology. *Rivista di Linguistica (Italian Journal of Linguistics)* 15/1: 31–62.
- Elekfi László 2001. Személyragozott igeekötők? *Magyar Nyelvőr* 125/2.  
<http://www.c3.hu/nyelvor/period/1252/125212.htm>
- É. Kiss Katalin 2003. Az egyszerű mondat szerkezete. In: É. Kiss Katalin – Siptár Péter – Kiefer Ferenc (szerk.): *Új magyar nyelvtan*. Osiris. Budapest. 74–160.
- É. Kiss Katalin 2007. Az ige utáni szabad szórend magyarázata. *Nyelvtudományi Közlemények* 104: 124–152.
- É. Kiss Katalin 2014a. Bevezetés. In: É. Kiss Katalin (szerk.): *Magyar generatív történeti mondattan*. Akadémiai Kiadó. Budapest. 11–13.
- É. Kiss Katalin 2014b. A tagadó és a kérdő mondatok változásai. In: É. Kiss Katalin (szerk.): *Magyar generatív történeti mondattan*. Akadémiai Kiadó. Budapest. 34–49.
- É. Kiss Katalin 2014c. Az ómagyar igeidőrendszer. In: É. Kiss Katalin (szerk.): *Magyar generatív történeti mondattan*. Akadémiai Kiadó. Budapest. 60–72.
- Forgács Tamás 2005. Grammatikalizálódás az igeekötők körében. In: Oszkó Beatrix – Sipos Mária (szerk.): *Uráli grammatizáló*. MTA Nyelvtudományi Intézet. Budapest. 88–116.
- Francis, Gill 1993. A corpus-driven approach to grammar: Principles, methods and examples. In: Baker, Mona – Francis, Gill – Tognini-Bonelli, Elena (szerk.): *Text and Technology*. John Benjamins Publishing Company. Amsterdam. 137–156.
- Goody, Jack 1987. *The Interface Between the Written and the Oral*. Cambridge University Press. Cambridge.



- Gugán Katalin 2015. És mégis: mozog? Tagadás és igemódosítók az ómagyarban és a középmagyarban. *Általános Nyelvészeti Tanulmányok* 27: 153–178.
- Gugán Katalin 2017. A magyar tagadó mondatok szórendje és a konstansráta-hipotézis. In: É. Kiss Katalin – Hegedűs Attila – Pintér Lilla (szerk.): *Nyelvelmélet és diakrónia 3.* Pázmány Péter Katolikus Egyetem BTK, Szt. István Társulat. Budapest – Piliscsaba. 91–110.
- Honti László – H. Varga Márta 2006. Meg van írva! A határozói igenév és a létige alkotta szerkezet funkciójáról és háttéréről. In: Mártonfi Attila – Papp Kornélia – Slíz Mariann (szerk.): *101 írás Pusztai Ferenc tiszteletére.* Argumentum. Budapest. 579–586.
- H. Varga Márta 2015. A ‘passzív’ jelentés és nyelvi kifejezőeszközei a magyarban. *Folia Uralica Debreceniensia* 22: 293–310.
- Indig Balázs – Sass Bálint – Simon Eszter – Mittelholcz Iván – Kundráth Péter – Vadász Noémi 2019. emtsv – Egy formátum mind felett. In: Berend Gábor – Gosztolya Gábor – Vincze Veronika (szerk.): *XV. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2019)* Szegedi Tudományegyetem, TTIK, Informatikai Intézet. Szeged. 235–247.
- Ittész Nóra 2009. A magyar nyelv nagyszótára. In: Fábián Zsuzsanna (szerk.): *Szótárírás és szótárírók.* Lexikográfiai füzetek 4. Akadémiai Kiadó. Budapest. 65–80.
- Jakab István 1976. *A magyar igekötők állományi vizsgálata.* Nyelvtudományi Értekezések 91. Akadémiai Kiadó. Budapest.
- Jakab István 1982. *A magyar igekötő szófajtani útja.* Nyelvtudományi Értekezések 112. Akadémiai Kiadó. Budapest.
- Jakubíček, Miloš – Kilgariff, Adam – Kovář, Vojtěch – Rychlý, Pavel – Suchomel, Vít 2013. The TenTen Corpus Family. *7th International Corpus Linguistics Conference CL 2013.* Online publikáció. Lancaster. 125–127.
- Johnston, Mark – Leslie, Sarah-Jane 2019. Cognitive Psychology and the Metaphysics of Meaning. In: Goldman, Alvin I. – McLaughlin, Brian P. (szerk.): *Metaphysics and Cognitive Science.* Oxford University Press. New York, USA. 183–205.

- J. Soltész Katalin 1959. *Az ősi magyar igeekötők: meg, el, ki, be, fel, le*. Akadémiai Kiadó. Budapest.
- Kádár Edit 2007. Az ige–igeekötő szórend csángó beszélt nyelvi szövegekben. In: Maticsák Sándor – Jankovics József – Kolláth Anna – Nyerges Judit – Péntek János (szerk.): *Nyelv, nemzet, identitás. A VI. Nemzetközi Hungarológiai Kongresszus nyelvészeti előadásai*. Nemzetközi Magyarasszertudományi Társaság. Budapest – Debrecen. 61–80.
- Kalivoda Ágnes 2016. *A magyar igei komplexumok vizsgálata*. Mesterszakos szakdolgozat. Pázmány Péter Katolikus Egyetem. Budapest.
- Kalivoda Ágnes 2017. Az igeekötők gépi annotálásának problémái. In: Ludányi Zsófia (szerk.): *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2017: XI. Alkalmazott Nyelvészeti Doktoranduszkonferencia*. MTA Nyelvtudományi Intézet. Budapest. 100–109.
- Kalivoda Ágnes 2018. Az igeekötős igék szintaxisa korpuszvezérelt megközelítésben. In: Scheibl György (szerk.): *Lingdok 17.: Nyelvésztdoktoranduszok dolgozatai*. Szegedi Tudományegyetem, Nyelvtudományi Doktori Iskola. Szeged. 159–176.
- Kalivoda Ágnes 2019. Véges erőforrás végtelen sok igeekötős igére. In: Berend Gábor – Gosztolya Gábor – Vincze Veronika (szerk.): *XV. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2019)*. Szegedi Tudományegyetem Informatikai Tanszékcsoprt. Szeged. 331–344.
- Kalivoda Ágnes – Vadász Noémi – Indig Balázs 2018. MANÓCSKA: A Unified Verb Frame Database for Hungarian. In: Sojka, Petr – Horák, Aleš – Kopeček, Ivan – Pala, Karel (szerk.): *Proceedings of the 21st International Conference on Text, Speech and Dialogue (TSD)*. Springer-Verlag. Brno. 135–143.
- Kálmán C. György – Kálmán László – Nádasdy Ádám – Prószyky Gábor 1989. A magyar segédigék rendszere. *Általános Nyelvészeti Tanulmányok XVII*: 49–103.
- Kálmán László 2013. *Egyre eljebb terjed*. Online publikáció a Nyelv és Tudomány ismeretterjesztő hírportálon. <https://www.nyest.hu/hirek/egyre-eljebb-terjed>

- Kálmán László – Trón Viktor 2000. A magyar igekötő egyeztetése. In: Büky László – Maleczki Márta (szerk.): *A mai magyar nyelv leírásának újabb módszerei IV.* Szegedi Tudományegyetem. Szeged. 203–211.
- Kassai Ilona 1994. A fonetikai háttér. In: Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 2. Fonológia.* Akadémiai Kiadó. Budapest. 581–666.
- Kemény Gábor 2018. Nem elváló igekötők. *Édes anyanyelvünk* 40/1: 7.
- Kenesei István 2000. Szavak, szófajok, toldalékok. In: Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 3. Morfológia.* Akadémiai Kiadó. Budapest. 75–136.
- Kerekes Judit 2011. Az igekötők meghatározásának problémái. In: Gécseg Zsuzsanna (szerk.): *LingDok10: Nyelvész-doktoranduszok dolgozatai.* Szegedi Tudományegyetem Nyelvtudományi Doktori Iskola. Szeged. 109–131.
- Kiefer Ferenc 1992. Az aspektus és a mondat szerkezete. In: Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 1. Mondattan.* Akadémiai Kiadó. Budapest. 729–809.
- Kiefer Ferenc 1995. Prefix reduplication in Hungarian. *Acta Linguistica Hungarica* 43/1–2: 175–194.
- Kiefer Ferenc 1996. Az igeaspektus areális-tipológiai szempontból. *Magyar Nyelv* 92/3: 257–268.
- Kiefer Ferenc – Ladányi Mária 2000a. Az igekötők. In: Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 3. Morfológia.* Akadémiai Kiadó. Budapest. 453–518.
- Kiefer Ferenc – Ladányi Mária 2000b. A szóképzés. In: Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 3. Morfológia.* Akadémiai Kiadó. Budapest. 137–164.
- Kiefer Ferenc – Ladányi Mária 2000c. Morfoszintaktikailag semleges képzések. In: Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 3. Morfológia.* Akadémiai Kiadó. Budapest. 165–214.
- Kiefer Ferenc 2003a. Alaktan. In: É. Kiss Katalin – Siptár Péter – Kiefer Ferenc (szerk.): *Új magyar nyelvtan.* Osiris. Budapest. 127–199.

- Kiefer Ferenc 2003b. A kétféle igemódosítóról. *Nyelvtudományi közlemények* 100: 177–186.
- Kiefer Ferenc 2007. *Jelentéselmélet*. 2. kiadás. Corvina. Budapest.
- Kiefer Ferenc – Németh Boglárka 2012. Amikor az igekötő nem telicizál. *Nyelvtudományi közlemények* 108: 267–273.
- Komlósy András 1992. Régensek és vonzatok. In: Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 1. Mondattan*. Akadémiai Kiadó. Budapest. 299–527.
- Komlósy András 2000. A műveltetés. In: Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 3. Morfológia*. Akadémiai Kiadó. Budapest. 215–292.
- Kornai András – Nemeskey Dávid Márk – Recski Gábor 2016. Detecting Optional Arguments of Verbs. In: Calzolari, Nicoletta – Choukri, Khalid – Declerck, Thierry – Goggi, Sara – Grobelnik, Marko – Maegaard, Bente – Mariani, Joseph – Mazo, Hélène – Moreno, Asunción – Odijk, Jan – Piperidis, Stelios (szerk.): *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*. European Language Resources Association (ELRA). Portorož. 2815–2818.
- Laczkó Tibor 2000. A melléknévi és határozói igenévképzők. In: Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 3. Morfológia*. Akadémiai Kiadó. Budapest. 409–452.
- Ladányi Mária 2007. *Produktivitás és analógia a szóképzésben: elvek és esetek*. Tinta Könyvkiadó. Budapest.
- Ladányi Mária 2012. Igekötős igék kapcsolódási mintázatai. Vizsgálati lehetőségek. In: Tolcsvai Nagy Gábor – Tátrai Szilárd (szerk.): *Konstrukció és jelentés. Tanulmányok a magyar nyelv funkcionális kognitív leírására*. Eötvös Loránd Tudományegyetem. Budapest. 71–84.
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar*, 2. kötet. Stanford University Press. Palo Alto, Kalifornia, USA.
- McInnes, Leland – Healy, John – Astels, Steve 2017. hdbscan: Hierarchical density based clustering. *The Journal of Open Source Software* 2/11: 205.

- McShane, Marjorie – Beale, Stephen – Nirenburg, Sergei 2005. *Disambiguating Homographous Prepositions and Verbal Particles In An Implemented Ontological Semantic Analyzer*. Technical Report 01-05. University of Maryland Baltimore County, Institute for Language and Information Technologies. Baltimore, USA.
- Nádasdy Ádám 2001. A két primadonna. *Magyar Narancs, Modern Talking* 2001/42. 10. 18. 15. Online publikáció. [https://magyarnarancs.hu/egotripp/nadasdy\\_adam\\_modern\\_talking-55638](https://magyarnarancs.hu/egotripp/nadasdy_adam_modern_talking-55638)
- Nádasdy Ádám 2005. A magyar folyamatos. *Magyar Narancs, Modern Talking* 2005/50. 12. 15. Online publikáció. [https://magyarnarancs.hu/egotripp/modern\\_talking-64968](https://magyarnarancs.hu/egotripp/modern_talking-64968)
- Németh Dezső 2001. A munkamemória szerepe a mondatmegértésben. In: Pléh Csaba – Lukács Ágnes (szerk.): *A magyar morfológia pszicholingvisztikája*. BIP – Osiris. Budapest. 83–118.
- Novák Attila – Rebrus Péter – Ludányi Zsófia 2017. Az emMorph morfológiai elemző annotációs formalizmusa. In: Vincze Veronika (szerk.): *XIII. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2017)*. Szegedi Tudományegyetem Informatikai Intézet. Szeged. 70–78.
- Novák Attila – Siklósi Borbála – Oravecz Csaba 2016. A New Integrated Open-source Morphological Analyzer for Hungarian. In: Calzolari, Nicoletta – Choukri, Khalid – Declerck, Thierry – Goggi, Sara – Grobelnik, Marko – Maegaard, Bente – Mariani, Joseph – Mazo, Hélène – Moreno, Asunción – Odijk, Jan – Piperidis, Stelios (szerk.): *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*. European Language Resources Association (ELRA). Portorož. 1315–1322.
- Nyíri Antal – Bodnár Ferenc – Pálfalvi Etelka – Rácz Endre – Velcsov Mártonné – Végh József Mihály (szerk.) 1971. *A Münchener Kódex 1466-ból. Kritikai szövegkiadás a latin megfelelővel együtt*. Codices Hungarici 7. Akadémiai Kiadó. Budapest.
- O'Connor, Michael Patrick 1978. *Hebrew Verse Structure, 1. kötet*. Eisenbrauns. Indiana, Winona Lake, USA.
- Oravecz Csaba – Várad Tamás – Sass Bálint 2014. The Hungarian Gigaword Corpus. In: Calzolari, Nicoletta – Choukri, Khalid – Declerck, Thierry – Loftsson, Hrafn – Maegaard, Bente –

- Mariani, Joseph – Moreno, Asuncion – Odijk, Jan – Piperidis, Stelios (szerk.): *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC 2014)*. European Language Resources Association (ELRA). Reykjavik. 1719–1723.
- Pais Dezső 1959. Az igekötők mivoltához és keletkezéséhez. *Magyar Nyelv* 55: 181–184.
- Pakerys, Jurgis 2017. *Measuring morphological productivity*. Kiosztvány. Graduate School of Linguistics, Philosophy and Semiotics (GSLPS). Tartu. <https://tinyurl.com/y38xsxmm>
- Pedregosa, Fabian – Varoquaux, Gaël – Gramfort, Alexandre – Michel, Vincent – Thirion, Bertrand – Grisel, Olivier – Blondel, Mathieu – Müller, Andreas – Nothman, Joel – Louppe, Gilles – Prettenhofer, Peter – Weiss, Ron – Dubourg, Vincent – Vanderplas, Jake – Passos, Alexandre – Cournapeau, David – Brucher, Matthieu – Perrot, Matthieu – Duchesnay, Édouard 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12: 2825–2830.
- Peredy Márta 2011. Az ige–igekötő sorrend a Jókai-kódexben. In: Bakró-Nagy Marianne – Forgács Tamás (szerk.): *A nyelvtörténeti kutatások újabb eredményei VI*. Szegedi Tudományegyetem Magyar Nyelvészeti Tanszék. Szeged. 181–197.
- Pléh Csaba 1998. *A mondatmegértés a magyar nyelvben. Pszicholingvisztikai kísérletek és modellek*. Osiris Kiadó. Budapest.
- Prószéky Gábor – Tihanyi László – Ugray Gábor 2004. Moose: a robust high-performance parser and generator. In: Hutchins, John (szerk.): *Proceedings of the 9th EAMT Conference*. Foundation for International Studies. La Valletta. 138–142.
- Rákosi György 2014. A case of disagreement: On plural reduplicating particles in Hungarian. In: Kertész András – Rákosi Csilla (szerk.): *The Evidential Basis of Linguistic Argumentation*. Studies in Language Companion Series 153. John Benjamins Publishing Company. Amsterdam. 179–198.
- Rebrus Péter 2000. Morfofonológiai jelenségek. In: Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 3. Morfológia*. Akadémiai Kiadó. Budapest. 763–948.
- Rosch, Eleanor H. 1973. Natural categories. *Cognitive Psychology* 4/3: 328–350.

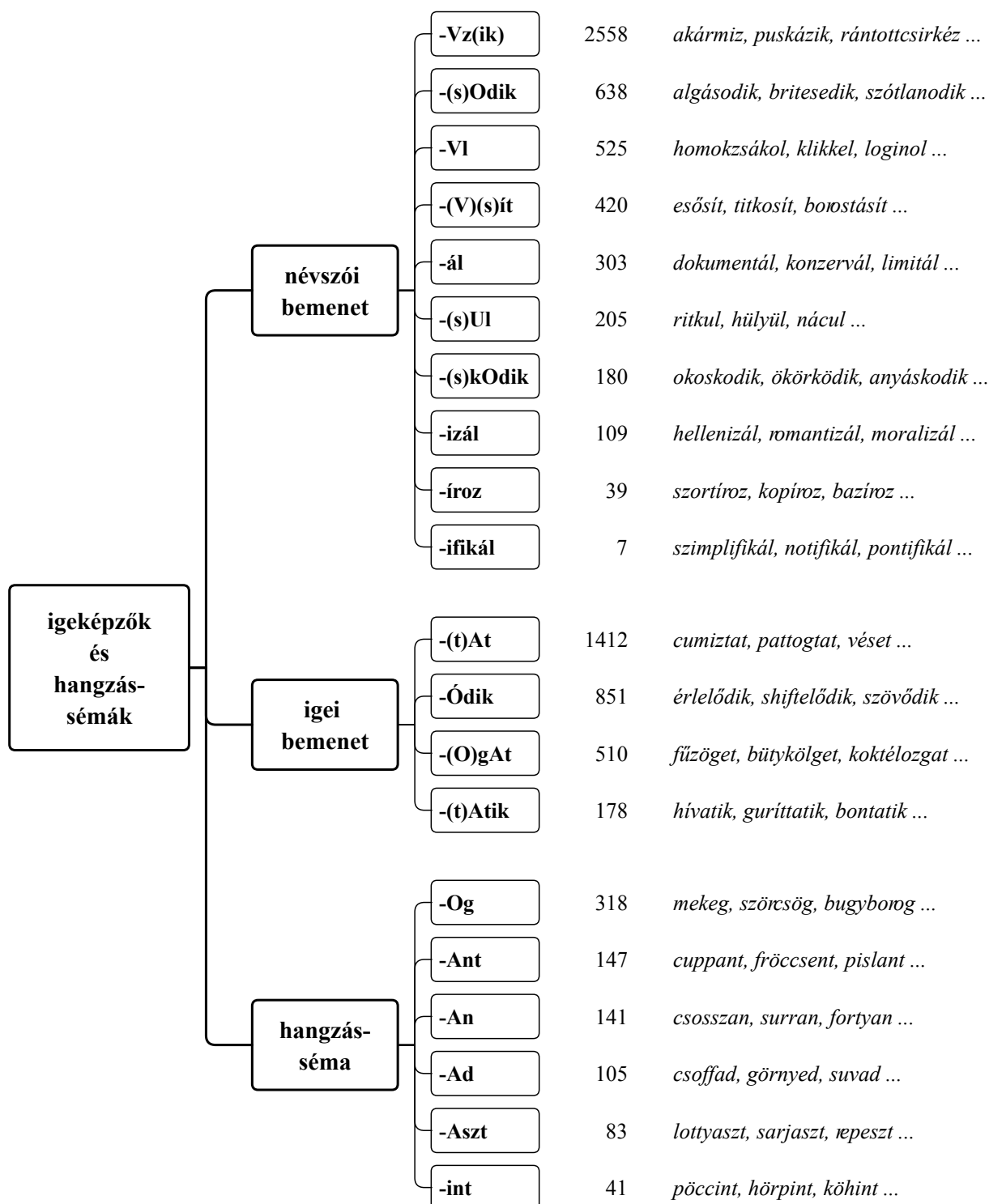
- Sárik Pál 1998. A határozói igenevek néhány problémája. *Magyar Nyelv* 94/4: 423–435.
- Sass Bálint 2011. *Igei szerkezetek gyakorisági szótára – egy automatikus lexikai kinyerő eljárás és alkalmazása*. Doktori értekezés. Pázmány Péter Katolikus Egyetem. Budapest.
- Sass Bálint 2015. 28 millió szintaktikailag elemzett mondat és 500 000 igei szerkezet. In: Tanács Attila – Varga Viktor – Vincze Veronika (szerk.): *XI. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2015)*. Szegedi Tudományegyetem Informatikai Tanszékcsoport. Szeged. 399–403.
- Sass Bálint – Váradi Tamás – Pajzs Júlia – Kiss Margit 2010. *Magyar igei szerkezetek – A leggyakoribb vonzatok és szókapcsolatok szótára*. Tinta Könyvkiadó. Budapest.
- Siklósi Borbála – Novák Attila 2016. Beágyazási modellek alkalmazása lexikai kategorizációs feladatokra. In: Tanács Attila – Varga Viktor – Vincze Veronika (szerk.): *XII. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2016)*. Szegedi Tudományegyetem, TTIK, Informatikai Intézet. Szeged. 3–14.
- Simon Eszter – Sass Bálint 2012. Nyelvtechnológia és kulturális örökség, avagy korpuszépítés ómagyar kódexekből. *Általános Nyelvészeti Tanulmányok* 24: 243–264.
- Siptár Péter 1994. A mássalhangzók. In: Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 2. Fonológia*. Akadémiai Kiadó. Budapest. 183–272.
- Stubbs, Michael 2013. Sequence and order. The neo-Firthian tradition of corpus semantics. In: Hasselgård, Hilde – Ebeling, Jarle – Ebeling, Signe Oksefjell (szerk.): *Corpus Perspectives on Patterns of Lexis*. Studies in Corpus Linguistics 57. John Benjamins Publishing Company. Amsterdam/Philadelphia. 13–34.
- Szabolcsi Anna – Laczkó Tibor 1992. A főnévi csoport szerkezete. In: Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 1. Mondattan*. Akadémiai Kiadó. Budapest. 179–298.
- Tan, Pang-Ning – Steinbach, Michael – Karpatne, Anuj – Kumar, Vipin 2018. Cluster Analysis: Basic Concepts and Algorithms. In: Tan, Pang-Ning – Steinbach, Michael – Karpatne, Anuj –

- Kumar, Vipin (szerk.): *Introduction to Data Mining*. 2. kiadás. Pearson Education. London. 525–612.
- Teubert, Wolfgang 2005. My version of corpus linguistics. *International Journal of Corpus Linguistics* 10/1: 1–13.
- Tolcsvai Nagy Gábor 2015. *Az ige a magyar nyelvben. Funkcionális elemzés*. TINTA Könyvkiadó. Budapest.
- Törkenczy Miklós 1994. A szótag. In: Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 2. Fonológia*. Akadémiai Kiadó. Budapest. 273–392.
- Váradi Tamás – Simon Eszter – Sass Bálint – Gerőcs Mátyás – Mittelholcz 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 (MSZNY 2017)*. Szegedi Tudományegyetem, Informatikai Intézet. Szeged. 49–60.
- Váradi Tamás – Simon Eszter – Sass Bálint – Mittelholcz Iván – Novák Attila – Indig Balázs 2018. E-magyar – A Digital Language Processing System. In: Calzolari, Nicoletta – Choukri, Khalid – Cieri, Christopher – Declerck, Thierry – Goggi, Sara – Hasida, Koiti – Isahara, Hitoshi – Maegaard, Bente – Mariani, Joseph – Mazo, Hélène – Moreno, Asuncion – Odijk, Jan – Piperidis, Stelios – Tokunaga, Takenobu (szerk.): *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. European Language Resources Association (ELRA). Miyazaki. 1307–1312.
- Yeo, In-Kwon – Johnson, Richard A. 2000. A new family of power transformations to improve normality or symmetry. *Biometrika* 87/4: 954–959.
- Zipf, George Kingsley 1932. *Selected studies of the principle of relative frequency in language*. Harvard University Press. Cambridge, Massachusetts, USA.
- Zsibrita János – Vincze Veronika – Farkas Richárd 2013. magyar1anc: A Toolkit for Morphological and Dependency Parsing of Hungarian. In: Angelova, Galia – Bontcheva, Kalina – Mitkov, Ruslan (szerk.): *Recent Advances in Natural Language Processing (RANLP 2013)*. INCOMA Ltd. Sumen. 763–771.



# Mellékletek

## I. számú melléklet: Produktív igeképzés gyakorisági adatokkal és példákkal



## II. számú melléklet: Az igekötők különféle jelentései néhány példamondattal

### **agyon<sub>1</sub> ‘megöl’**

- (1) amikor Simeon a császár parancsára agyonköveztetette.
- (2) és agyongázoltatom magam egy teherautóval
- (3) hogy ott agyonbotozzák őket.

### **agyon<sub>2</sub> (+ magát) ‘túlzásba esve (károsodásig/kimerülésig)’**

- (1) Agyonszótározom ezt a Macbeth-idézetet,
- (2) és így már nyugodtan agyonlassíthatjátok a szerveret.
- (3) őt ugyanis saját pártjának lapja agyontömjénezi
- (4) agyonkészülte magát,
- (5) mert így meg agyonpuffereli magát,
- (6) aki agyonfeketézte magát,

### **alá<sub>1</sub> ‘irány (vmi/vki alá)’**

- (1) és utána alárohantam az autónak...
- (2) Csakhogy a gondos fényező mester alácsiszol a festéknek,
- (3) Aláhajolt a díványnak - senki.

### **alá<sub>2</sub> ‘irány (lefelé)’**

- (1) hogy onnan alálógassak egy sólymot röptében?
- (2) sapkája bojtja aláfityeg,

### **alul ‘rosszabbul, alacsonyabb szinten, nem eléggé’**

- (1) alulprognosztizálja a kiadásokat, a hiányt,
- (2) mert a mosdók számát kicsit aluldimenzionálták a tervezők.
- (3) hogy alulpromotálta az albumot.

### **át<sub>1</sub> ‘irány (át vmin/vkin)’**

- (1) XY árnya áthussant a szobán.
- (2) amikor XY átslisszant a határon.
- (3) áttapogatóznak a londoni ködön,
- (4) és akkor kedélyesen átszólongat a szomszéd asztalokhoz:
- (5) s áthengergeti a déli oldalra.

### **át<sub>2</sub> ‘eltölt egy időszakot vmivel’**

- (1) Az ebédidőt is átvitáztuk.
- (2) hogy a gyűléseket csendben átszóraztunk.
- (3) a nagy XY társaságában éjjeket átiddogáltam,
- (4) akik önmaguk árnyékaiként átsopánkodják életüket,
- (5) így szépen átszundikálta az életveszélyt.

**be<sub>1</sub> 'irány (befelé)'**

- (1) Aztán egy mellékvágányon a kisördög bekoccant a fejembe,
- (2) és kettejük közé becsikorgott a troli.
- (3) én egész eddig csak beoffoztam ide,
- (4) Hogy azonnal beédeskedi magát a nézőtér szívébe.
- (5) A füdőszoba kövén fekszik, vagy begömbikéskedik a mosdókagylóba.
- (6) betraktorozott az erdőbe.
- (7) akkor a tanító néni ezt bepusmogja a védőnőnek,
- (8) Bevakkantottam a telefonközpontba,

**be<sub>2</sub> 'teljes területet/felületet megtelítve'**

- (1) A konyha beillatosodott,
- (2) ha begombostúzhatnám Budapest térképét negyedek, albérletek szerint,
- (3) jól betérerőzné az egész kis medencét,

**be<sub>3</sub> 'nagy/kielégítő mértékben'**

- (1) HA már ennyire "berokkeresedtünk",
- (2) hogy nagyon bekomolyodjék a társadalom,
- (3) ahol most jól bepizzáztunk.
- (4) és majd jól bekarácsonyozunk!
- (5) Lehet, hogy én ovis vagyok, mert minden este begumimacizom... ez már szenvedély...

**be<sub>4</sub> 'intenzíven csinálni/történni kezd'**

- (1) hogy te is bepityeregsz rajta.
- (2) mivel beallergiásodtam a mandulára,
- (3) hogy egy újonnan kipróbált kozmetikumra beallergizál a bőr.

**bele<sub>1</sub> 'irány (befelé)'**

- (1) hiszen azt Pistike belefotosoppolja 5 perc alatt.
- (2) És belebagóznak az ember pofázmányába úton-útfélen,
- (3) hogy beleszurkapiszkáljanak az agyvelőm tekervényeibe.
- (4) Az illuminált vendégek pedig belekurjongattak a mikrofonba.

**bele<sub>2</sub> 'iszik/eszik vmiből'**

- (1) mielőtt meghúzná a flaskót [ az Ikrek után ], csak úgy, mondhatni, tökspontán belenyakal.
- (2) Egészen hazáig kibírtam anélkül, hogy belemajszoljak!
- (3) beleeszel valamibe egy kis kanállal,
- (4) mire a bárshékeken ülők belehörbölnek az italukba

**bele<sub>3</sub> 'beleszól vmibe'**

- (1) BELEKAKUKKOLTAM a nagyok DOLGÁBA.
- (2) A topicba beleszörcsöltem "europa"-t és 0-ást adott!

(3) ha valaki belefanyalog egy topicba,

**célba ‘írány (egy célként kiszemelt helyre)’**

(1) és célbabajonetezik az arabdémonokat...

(2) aki minden öt percben célbaköp az előtte hajlongó hátára,

**egybe<sub>1</sub> ‘írány (több helyről egy helyre)’**

(1) Az adjunktus egybekérte a családot.

(2) Halálisten mester a világ összes halálkaszáját egybehívja,

**egybe<sub>2</sub> ‘részekből egészbe’**

(1) és a téglasorok szinte egybeégtek.

(2) ahogy újra egybemarkoltam széthulló gondolataimat,

(3) és 200 fokon 15-20 perc alatt egybesütjük.

(4) a lassan dolgozó történelem a pluszt a mínusszal végül is egybecsiszolta.

**el<sub>1</sub> ‘írány (elfelé)’**

(1) és a kígyó eltekerődött valahová.

(2) aztán eltrolizunk a Mekiig.

(3) Udvariasan eltisztelgek a hajóskapitánytól.

(4) Na csá, egyelőre elsuccogok.

**el<sub>2</sub> ‘változik/változtat vmivé/vkivé’**

(1) ellegendásodott a történet.

(2) Hogy okosan elskandinávosodjunk.

(3) A mi családuk viszont már teljesen elmekdonálcosodott.

(4) de szörnyűséges tettük miatt "elothellósodnak",

(5) Nem csoda, hogy ennyire elzombulsz, ha ilyeneket hallgatsz...

(6) A film rettentően eltömegesült.

(7) Épp ezért le kellett lépnem, mielőtt végleg elérzelgősítenél.

(8) Fejemben a saját hangomon hallottam a mondatot, ami még jobban eltanácstalanított.

**el<sub>3</sub> ‘elmond vmit’**

(1) Alapeszmém, amit itt elnyervogtam, de azért nem mindig ezen rágódom.

(2) Ők azért néha elpicsogják, miről is van szó.

(3) A főzéssel kapcsolatban ismét elnyüszögném, hogy vannak ám vega nagyvadak is...

**el<sub>4</sub> + magát ‘nem szándékosan csinál vmi hirtelen’**

(1) néha elpisilte magát, de még véletlenül sem aludt volna.

(2) Elköhinti magát.

(3) elbüfögte magát.

(4) aztán elhahotázta magát.

**el<sub>5</sub> ‘tölti vmivel az idejét (általában hosszasan)’**

- (1) valóban egy jó vasárnapi ebéd után eltajtékpipáztak,
- (2) Eldemokráciáztunk volna ki tudja meddig,
- (3) Kicsit elmalmozik itt,
- (4) hogy ehhez neki semmi kedve, olyan jól elminiszterelnökösködik ő enélkül is.
- (5) Egy kicsit elbúvárkodtam az index archívumában.
- (6) és elcsemcsegték azokon a jelenségeken,
- (7) Egy kicsit eltötyörögtek rajta,

**el<sub>6</sub> ‘vmilyen tevékenységgel elveszít/elront vmit’**

- (1) Én sajnálom a pénzt virágföldre, inkább elsörözöm meg elcsipszezem, így tőletek kérek segítséget.
- (2) Nem, de ha elborsozod a rántottát, max. az aznapi reggelid nem lesz olyan finom.
- (3) hogyha a családi ezüsből származó bevételt elvacsorázta a család,
- (4) a magunk életét elbolondozhatjuk, de a gyermekeinkét nem!!
- (5) s elcukrozza meg elkörhintázza keserves garasait.
- (6) mert fizetése nagy részét elinternetespókerezi?
- (7) mert elböszméskedték a dolgot,
- (8) Miután eljégmadárkodta a közös szándékunkat.

**el<sub>1</sub> ‘irány (vmi/vki elé)’**

- (1) hát megfőzi a mindennapi ennivalóját meg elétálalja
- (2) hogy elérúgta volna a kapu előtt a labdát
- (3) aki mint a vörösszőnyeg így elégurítja a legjobb civilszervezetet,

**el<sub>2</sub> ‘irány (vhonnan kifelé, megjelenve)’**

- (1) Erre eléguberált hatvan forintot...
- (2) Eléhalássza végre mind a négyet,
- (3) Elékaparásza zsebkését,

**el<sub>3</sub> ‘időben vmit megelőzve’**

- (1) az az igazság... alig forog a nyelved... két vodkát is elényomtál a söröknek.

**el<sub>0</sub> ‘irány (vhonnan kifelé, megjelenve)’**

- (1) Kiáltozására a kedvesnővérek fapapucs is előkattogott a folyosóról.
- (2) Egy újabb óra múlva a mozdony is előpöfögött valahonnan.
- (3) A szürke hamu alól előparázsluk a tűz.

**el<sub>0</sub> ‘időben vmit megelőzve’**

- (1) Ugyan a darab szerkezetére Hollywood már előtréningezte a nézőt -
- (2) S még félelmetesebb, hogy Kádár országlása alatt azt is megtanultuk, hogy a Vezér szája íze szerint beszéljünk, előre helyeseljük intézkedéseit, magyarán előnyalunk.

(3) Holnap előpihenek a nagy eseményre.

**előre ‘irány (előre)’**

- (1) ha a kerékpárosok a piros lámpánál az autósokat megelőzve előrecsorognak.
- (2) Miután előredöcögtem a kert végéből,
- (3) A csecsemőket előreadogatják a mentőegységekhez,
- (4) A kiskutya előrekaptatott,
- (5) csak előrerikkantunk a sofőrnek, ugyan,

**fel<sub>1</sub> ‘irány (felfelé)’**

- (1) A képernyőre szintén felfrecsegnek a szénsavas üdítők.
- (2) Püspökiből felzötyögök a hatos villamoson a nagy állomásig.
- (3) Felslattyogott a lépcsőn,
- (4) pedig a hozományt is felszekerezte a hegytetőre.

**fel<sub>2</sub> ‘hangot ad ki (általában hirtelen és röviden)’**

- (1) Jolanda elégedetten felröffent,
- (2) ahonnét váratlanul felburrantak a galambok,
- (3) és felberrent a csengő is.
- (4) A tömeg erre felhurrogott.

**fel<sub>3</sub> ‘változik/változtat vmivé/vkivé (pozitívan)’**

- (1) hogy most csak a biogazdálkodást felgépesítsük,
- (2) ám úgy felcavintonozta a választópolgárok memóriáját,
- (3) hogy az eleve izmos gépet még jobban felvadítsák.

**fel<sub>4</sub> ‘telít/telítődik vmivel (általában negatív)’**

- (1) amelyek közül jó néhány annyira felgyomosodott,
- (2) viszont a nem jól tárolt akku hamar felszulfátosodik
- (3) de amikor feldiszkósította a dalait teljesen kiábrándultam belőle...

**félre<sub>1</sub> ‘irány (el az eredeti helyéről)’**

- (1) A pavilon ajtaján a tüllfüggönyt félreakasztották,
- (2) A fej kissé félrebiccen,
- (3) Az őgyelgő népség félreszökdösött.

**félre<sub>2</sub> ‘hibásan, rosszul’**

- (1) Ráadásul félrefeltételez...
- (2) félretisztelték az irodalom peremére.
- (3) és csúnyán félreáraznak dolgokat.
- (4) hogy ezeket a csontokat bizony félrehatározták.
- (5) sőt akit még a környezete is félreistenít,
- (6) hogy a fényképeket félrekategorizálják.

**felül<sub>1</sub> ‘irány (át és fel)’**

- (1) csak felülmaticázzák a billentyűzetet magyarrá.
- (2) A PRV motor hangját "felülszinkronizálták" V8-as motorok hangjaival.
- (3) amelyek felülrétegezték egymást.

**felül<sub>2</sub> ‘jobban, magasabb szinten (esetleg túlozva)’**

- (1) hiszen amennyiben felülkínál azon az áron,
- (2) és mint látható, kissé felülsomagolták a kis adaptert.
- (3) Picasso felülrajzolta önmagát?

**fölé ‘irány (vmi/vki fölé)’**

- (1) vagy maga mellé teszi a bilit, fölé tartja a gyereket,
- (2) s fölé magaslik a palotavárnak is,
- (3) hogy a "rendőrautójával fölegördült a holttestnek".

**hátra ‘irány (hátra)’**

- (1) hátragereblyézte az ujjaival a haját:
- (2) Noha többször hátrakémslelték még később rám,
- (3) a középpályás hátraívelte a labdát,
- (4) Hátrapisszegtem hát:
- (5) Erre mérgesen hátraförmel rám,

**haza ‘irány (hazafelé)’**

- (1) s hazahömpölyögjenek a hét törpék.
- (2) Hajnal 3-kor részegen hazaténfergek,
- (3) és röpke 24 óra alatt hazasíel.
- (4) Elmegyünk vadászni, aztán lassacskán hazavadászatjuk magunkat.
- (5) Hazasettenkednek az adócsalók eltítkolt vagyunkái

**helyre<sub>1</sub> ‘irány (vissza az eredeti helyére)’**

- (1) s helyreusszantják suttyogó csizmájukat.
- (2) - A lány lehiggadt, helyretekerte magán a ruhát.
- (3) A piros Polski továbbrobog, a kéz kinyúl a kilincsért, helyreapja a félig nyitott ajtót menet közben, aztán a motor felbőg, és a kocsí teljes erőből fölvág a Margit utcába, a domb éles meredéken.

**helyre<sub>2</sub> ‘vissza az eredeti (jó) állapotba’**

- (1) Ugyanakkor a Bokros csomag helyrelendítette a költségvetést,
- (2) Még helyreihatjuk a verekedés okozta kijózanodást.
- (3) aki a tankönyvi túlkapásokat helyrefésülné.

**hozzá<sub>1</sub> ‘irány (vmihez/vkihez)’**

- (1) Mihelyt hozzáöttyent a mozdony a szerelvényünkhöz,

- (2) és hozzázsigzalagozom a benzinpumpához,
- (3) Az őszi hernyók gyakran hozzászőnek egy levelet a gyümölcs felületéhez,

**hozzá<sub>2</sub> ‘megjegyzést fűz vmihez’**

- (1) úgyhogy hozzáröffenék.
- (2) hozzádödögött valamit az abbahagyott dörögéshez.
- (3) Hozzáugatok az 1. ponthoz:

**hozzá<sub>3</sub> ‘igazodik/igazít vmihez/vkihez’**

- (1) de a modern ember már hozzászocializálódott ehhez.
- (2) A könyv lassan hozzábutul a televízióhoz.
- (3) hogy a magyar konyhát hozzáfinomították a nemzetközi ízléshez,

**ide ‘irány (erre a helyre)’**

- (1) Mint szellem, idelebegek nektek.
- (2) s ígérem néha idekukkolok.
- (3) (Az ágyból rögtön idepattantam a gép elé.)
- (4) Idepartizánkodod ezt a szóözünt,
- (5) hogy mindenki ideböförogje a véleményét.

**keresztbe<sub>1</sub> ‘irány (át vmin/vkin merőlegesen)’**

- (1) a villanyvezeték az úton keresztbelóg,
- (2) amikor hirtelen keresztberohant előtte az úton egy kondányi vaddisznó.
- (3) ha keresztbeutaztad egész Szófiát.
- (4) Egy idióta vizirendőr pl. keresztbemotorcsónakázott az úszókon,

**keresztbe<sub>2</sub> ‘egymást keresztezve’**

- (1) Na most jól keresztbeírtunk:)))
- (2) hogy az önkormányzatok gyakorlatilag egymás helyett fizessenek, keresztbefinanszírozzanak,

**keresztbe<sub>3</sub> ‘egészből két részbe (merőlegesen)’**

- (1) Léceket keresztbetörte ez a repedés, ez nem valószínű,

**keresztül<sub>1</sub> ‘irány (át vmin/vkin merőlegesen)’**

- (1) amin keresztül bármilyen szenny keresztülszivároghat.
- (2) vagy egyedül keresztülrambózzuk magunkat az egész pályán az ellenség háta mögött.
- (3) akár egy kutyakiképző védőruhán keresztülharapnának a zombik.
- (4) Hangja csak úgy keresztülcsilingel a szobán,
- (5) míg ezen az állapoton keresztülgondolkodod magad,

**keresztül<sub>2</sub> ‘egészből két részbe (merőlegesen)’**

- (1) Olyan hirtelen keresztülszakadt a gát,



- (2) Ez a szubkultúra keresztülmetszette a társadalom osztályszerkezetét, vallási, etnikai és hagyományos, kulturális rétegződését is.
- (3) Keresztülfűrészelik alattunk a földgömböt! -

#### **ketté ‘egészből két részbe’**

- (1) sőt a szobáikat összekötő erkélyt is kettéválasztatta,
- (2) Látszólag taláalomra választ ki férfiakat, akiket aztán igazi hentes módjára kettéaprít egy bárd-dal.
- (3) és tavaszra ügyesen, mérnöki precizitással kettérágják valamennyit,

#### **ki<sub>1</sub> ‘irány (kifelé)’**

- (1) és a gyógyszerészek kisuttyantották őt a laboratórium hátsó kijáratán...
- (2) aztán kikódorgott az ajtón.
- (3) Az öregember kelletlenül kikoszmorgott a gerendaházból,
- (4) úgyhogy kicsasszogok a fürdőbe,

#### **ki<sub>2</sub> ‘kimozdul/kimozdít vmilyen állapotból’**

- (1) ha mindenki kigazdagodna a szegénységből.
- (2) kistratégiázza őket a hatalomból,
- (3) XY kitwitterezte magát az angol válogatottból

#### **ki<sub>3</sub> ‘kimond vmit’**

- (1) Az igazgató ugyan az évmegnyitóban kidörögte,
- (2) nem büntetésből -e talán, amiért kikottyintotta a titkot,
- (3) amint itt kinyekeregtem:

#### **ki<sub>4</sub> + magát ‘eleget csinál vmiből’**

- (1) Kimocorogta magát?
- (2) Miután kellően kiháborogtuk magunkat,
- (3) ahol kedvükre kirettegphetik magukat.
- (4) Na polgártársak, miután virtuálisan kisöröztük magunkat,
- (5) most jól kifórumoztam magam egy ideig...
- (6) kibömbiztem magam egy életre.

#### **ki<sub>5</sub> ‘megtisztít vmit vmivel’**

- (1) ha a budit kidomesztoszolom,
- (2) de tegnap kivízkeőldóztam a Gb-t
- (3) XY szerint az általuk kidolgozott, úgynevezett Széchenyi elektronikus elszámoló rendszer akár még a fekete gazdaságot is khipózhathá.

#### **ki<sub>6</sub> ‘megmutat, láttat vmit’**

- (1) nincs kinek kikencézzem magam
- (2) majd a szempillákat két rétegben kispirálóztam.

- (3) Nincs itthon a nővére, aki eddig vízbe rakta, szekrénybe tette, kivetkőzte, megfésülte,
- (4) kellőképpen kipitykázta-e a templomajtót.

**körbe<sub>1</sub> ‘irány (kezdőponttól a végpontig, a két pont azonos)’**

- (1) körbekoccint a poharával,
- (2) persze azért még körbekonyharuházom,
- (3) Éjszakánként felül az ágyon, körbegöngyöli magán a takarót,

**körbe<sub>2</sub> ‘teljesen, végig’**

- (1) majd körbenyerte vele a fesztiválokat.
- (2) körbenagygyúlésezi a fél országot.
- (3) Apád körbedicsekszi veled a sarki kocsmákat.

**körül<sub>1</sub> ‘irány (kezdőponttól a végpontig, a két pont azonos)’**

- (1) mikor a teraszon körültepegi a fodrász,
- (2) A mozgó testek körülvillogták a borospoharat.
- (3) a csimpánz boldogan körülmakogta őket,
- (4) hogy éjjel körüllocsolják az ágyát petróleummal -
- (5) aki négyszög alakban körülvonszolta.

**körül<sub>2</sub> ‘teljesen, végig’**

- (1) Minél jobban körülpáncélozod magad ítélőfogalmakkal, annál kevésbé vagy kíváncsi.
- (2) Ahogy a roksztárok körülturnézzák a fél világot új lemezükkel,
- (3) aztán ha így körülbecézel,

**közbe<sub>1</sub> ‘irány (két vagy több dolog/ember közé)’**

- (1) Talán egy pillanatra közberakja a kezét,
- (2) azonnal közbefurakodik féltékenyen
- (3) de a főmarsall izgatottan közberontott.

**közbe<sub>2</sub> ‘hanggal/beszéddel megakaszt’**

- (1) és folyvást közbepittyeg a többi hívó.
- (2) és Tarantino máris közbedörrent:
- (3) Ketten-hárman közbevihogtak.
- (4) Most a történelem nekik is közbeköhintett.
- (5) Az osztályból egy-két hang közbevihajgott erre.

**külön ‘elválasztva, elkülönítve’**

- (1) hogy a saját gyászát a menyéétől különkaparja . . .
- (2) Ginát viszont Zsuzsanna azonnal különparancsolta.
- (3) Egy alkalommal XY-nal is különöcsatázott egy nagyot a hazai kapu előtt.

**le<sub>1</sub> ‘irány (lefelé)’**

- (1) hirtelen letottyadt a konyhaszékre,
- (2) és leslisszantam a lépcsőn, hátamon a takarómmal.
- (3) és lebillegjenek az edzőterembe tornázni kicsit.
- (4) Ez utóbbiról színesfémgyűjtők levandálkodták a feliratot, már másodszor.

**le<sub>2</sub> ‘változik/változtat vmivé/vkivé (negatívan)’**

- (1) hogy az egyéniség értékeit leszürkítsük
- (2) Már-már az érdektelenségig lejózanítja a színeket, a formát
- (3) Most nagyon leesősödött az idő,
- (4) csöndben leokkultistásodom.

**le<sub>3</sub> ‘nevez vminek/vkinek (pejoratívan, lekezelően)’**

- (1) A suliban állandóan levidékiparasztotak.
- (2) hogy - teljesen érdemtelenül - lecicaautózták a Pumát.
- (3) aki a Szt. Jobbot letetemcafatozta.
- (4) lemicisapkázzák a Szentkoronát,
- (5) vagy következetesen levadkenderezték ezt a növényt (amit egy magára adó parlagfű kikér magának).
- (6) hogy élből lejulizott,
- (7) Nem érdekelt, hogy lehercegnőzött,
- (8) Miután jól lekabátlopásügyeztük az okirathamisítást,
- (9) hogy leédeskislányozzák a békésen sétálgató nyugdíjas mamikák.

**le<sub>4</sub> ‘teljes területet/felületet lefedve’**

- (1) lecementezték volna ezt a lyukat, (
- (2) Miután lealvázvédőztem mindent,
- (3) hogy ezt a hatalmas nagy területet legyeptéglázzuk,
- (4) tehát végig lemozaikkövezték,

**le<sub>5</sub> ‘véletlenül leönt, leeszik vmivel vmit/vkit’**

- (1) lepálinkázta az aktát,
- (2) Reggelijét befejezte Nyakkendőjét letejezte
- (3) Jól levajaskenyereztem ezt a papírt,
- (4) - Lefagylaltoztad az ingedet!

**meg [csak grammatikai funkció: befejezettség]**

- (1) Ha viszont kétmillió betétes egyszerre megszimmantja,
- (2) s hát nagy későre valaki megcsosszant odabent,
- (3) és megcsemcsegte a magáét.
- (4) hogy a normálisak is meghőzöngösödjének, rossz látni.
- (5) amikor 1964-ben megmusicalosították életét.
- (6) és majd jól megiparikémkedem.

- (7) " Finomak érdekes módon ezt egy picit megpirospaprikázzák.
- (8) Vén marha Rinya - csóválta a fejét az iparőr, aztán a régimódi telefonhoz lépett, és megkurb-lizta.

**mellé<sub>1</sub> 'irány (vmi/vki mellé)'**

- (1) Így mellécsurgattam egy kis narancslevet, citromlevet, és egy csipet fahéjjal fűszereztem.
- (2) ha mellécéloznak ott a kollegák a piszoárnak.
- (3) és mellésodródott egy másik forgó fiatal srác.
- (4) Lajos mellégugyorodott segíteni.

**mellé<sub>2</sub> 'hibásan, rosszul'**

- (1) hogy folyamatosan mellébeszélt Székely ügyében, pontosabban melléhallgatott.
- (2) A zárat valaki figyelmetlenül vagy sietségében mellécsukta.
- (3) Razorról azonban az a hír járja, hogy már többször "mellészagolt", és téves riadót ugatott,
- (4) Ha esetleg mellészavaz, akkor lehet arra hivatkozni, hogy tévedett.
- (5) Az SZDSZ politikai vezetése valamit rettenetesen melléértett.

**neki<sub>1</sub> 'irány (vmihez/vkihez, érintkezve)'**

- (1) és nekihussant a kapu előtti csúszkának.
- (2) A feje nekidöccen az ablaknak,
- (3) zavarában neketetődött egy fatörzsnek.
- (4) ugyanis nekigyalogoltam egy vascöveknek.
- (5) mert ott a tömeg nekipasszíroz a fém korlátnak.

**neki<sub>2</sub> 'elkezd vmit'**

- (1) arca kipirult, majd nekikékült,
- (2) egyszer csak nekihúsosodik.
- (3) Ezért nekidühödtem az írásnak,
- (4) Kati szíve nekimelegedett.
- (5) hogy valamivel nekipeszsdítse a hangulatot.

**nyakon 'irány (vki nyakára)'**

- (1) amikor Adabayor nyakonpöckölte Mascherano-t,
- (2) Harmadszorra nyakonkapta,

**oda 'irány (arra a helyre)'**

- (1) és lassan odaklappogott az ablakhoz a papucsában,
- (2) a fiú odalocsantott egy kis vizet a kannából.
- (3) Odaszerencsétlenkedett a község házra,
- (4) aki odakakáltatja a kuttyáját a betonra.
- (5) mert még a végén odatrüsszent valahova egy másik szörnyűséget.

**odébb 'irány (kicsit elfelé az eredeti helyéről)'**

- (1) aki ilyenkor szemérmesen odébbkanyarodik,
- (2) és odébbhessegettem a lelkünk-üdvéért aggódó Grácikát.
- (3) Odébbcipelte a nyugágyat.

**össze<sub>1</sub> ‘irány (több helyről egy helyre)’**

- (1) néha összefütyölte kutyáit
- (2) hogy összekántálja a félfalut,
- (3) Összetelefonálódott egy társaság azzal,
- (4) mint amivel a világirodalom nagyjait összelassózta.

**össze<sub>2</sub> ‘részekből egészbe’**

- (1) és fékével összemerevült a mozdony
- (2) összeszegeltem a primitív szánkóalkalmatosságot
- (3) Utána kármentés céljából, műgyantával, sziloplaszttal összekókányolta,

**össze<sub>3</sub> ‘túlzásba esve (károsodásig/kimerülésig)’**

- (1) Most jól összeofftopikoztuk XY rovatát.
- (2) Keynes alaposan összederiválta magát!
- (3) akik aztán minden számárságot összehárijánoskodtak, önvédelemből inkább,

**össze<sub>4</sub> ‘érintkezik, összekapcsolódik vmivel/vkivel’**

- (1) - Ti pedig már összeecsókolóztatok a fával is.
- (2) és buggyos zsebe zsákjában ceruzái összezőrrentek köptégelyével.
- (3) Akkor mégiscsak összealkudtatok ezzel az ájtatos, vén csirkefogóval.
- (4) A kiadó neve különösen összeforródott néhány zenésszel,

**össze<sub>5</sub> ‘nagy/teljes területet/felületet érintve’**

- (1) nehogy összekrétázzam az okos homlokom - könnyedén,
- (2) de annyira egyik sem volt jó, hogy összeuborkázzunk vele egy bélszínt.
- (3) ha játék közben összemálnázta öltönyét
- (4) hogy összelekvározzuk meg nutellázzuk a törülközőket,
- (5) Ragacsos (kivéve a fej reklámfelületét) és könnyen összebenzinezi a ruhánkat,
- (6) nem beszélve arról, hogy utálom, ha összerúzsozok valakit,

**össze<sub>6</sub> ‘vmi/vki tömörítésével/kicsinyítésével’**

- (1) összefonnyasztják,
- (2) az öregség összeszikkaszt.
- (3) Mostanra icipicire összezanásodott,

**pofán ‘irány (vki arcára/arcába)’**

- (1) aki eltűri ahelyett hogy pofánnymorna, tökkencs.
- (2) (A fundi külön szám, mikor egy jó nagy, laza hógolyóval pofántalálom:))

**rá<sub>1</sub> ‘irány (vmire/vkire)’**

- (1) A mítoszi madár jó vastagon ráplötytyintett a térképre,
- (2) ha családi házuk falára valaki rágraffitizne tetszőleges szöveget.
- (3) hogy naponta rálötytyentem a teát?
- (4) A szép bőrülésre is rápergett a sár.

**rá<sub>2</sub> ‘rászól, hanggal jelez vminek/vkinek’**

- (1) és néha mérgesen ráhurrognak a fürdőzőkre.
- (2) A tábornok ráröfönt.
- (3) Egy idős asszony rácsörrent a többiekre:
- (4) rámekegtem a kecskére,

**seggbe ‘irány (vki hátsó felébe)’**

- (1) aki eközben folyamatosan seggberugdossa őt.
- (2) XY ősapámat seggbenyilazták valami kis Visegrádi csetepatében 1411-ben.

**szarrá<sub>(+ magát)</sub> ‘túlzásba esve (károsodásig/kimerülésig)’**

- (1) ha ezért szarrábombáznák Mo.-t!
- (2) Mióta visszajöttem két hete csak szarráfagyok,
- (3) Csórikám, ezzel szarrákereshetné magát a nyugati világban mindenféle showműsor keretében.

**széjjel<sub>1</sub> ‘irány (egy helyről több helyre)’**

- (1) A vízfestékek széjjelgurultak, egy kármin el is roppant az Oszkár talpa alatt.
- (2) mintha csak széjjelömlenék a papíron.
- (3) Közben széjjelpillantott az udvaron:

**széjjel<sub>2</sub> ‘egészből részekbe’**

- (1) mint ha a hatáskörök széjjelaprózódnának.
- (2) s ebben a jól végzett munkában Genf vidáman széjjelbomlott.
- (3) hogy széjjelforgácsoljuk az erőnket,

**széjjel<sub>3</sub> (+ magát) ‘túlzásba esve (károsodásig/kimerülésig)’**

- (1) ha széjjeloffoljátok itt a topicot, közétek lövetek.
- (2) hogy 2 perc alatt széjjelkarcolnék akármilyen mobilt,
- (3) - Széjjelég ebben a nagy tűzben!
- (4) XY-t visszafelé hallgatva bármikor széjjelröhögöm magamat.
- (5) Otthon mindenki széjjelmulatja magát
- (6) hogy az életének párja, széjjelstrapálja magát a munkában,

**szembe<sub>1</sub> ‘irány (vmivel/vkivel szemközt)’**

- (1) S akkor Lőrinc, a lombok közül, hirtelen szembebukkant vele,
- (2) ha valaki szembeautózik a forgalommal a kétpályás úton,
- (3) a derengő félhomályban bármikor szembetévedhetne Kosztolányi vagy Márai,

**szembe<sub>2</sub> ‘írány (vki szemébe)’**

- (1) - Hülye gyerek - morogtam rá, és szembefröcsköltem őt, majd kimásztam a vízből.
- (2) szembelöttyintettem a konyakkal.
- (3) Aztán alattomban közelébe lubickolt, és hirtelen szembefricskálta.

**szénné<sub>(+ magát)</sub> ‘túlzásba esve (károsodásig/kimerülésig)’**

- (1) mind titeket szénnécsócsálnak közben.
- (2) a következő talán az lesz, hogy valakit majd szénnélőnek kamu élő adásban kamuból,
- (3) Ezalatt szénnéfényképezik a gyereket.
- (4) már ne haragudj de hogyan él valaki normális életet ha nyáron mondjuk egy értelmes pl. ten-gerparti családi üdülés helyett 1 hétig szénnéissza az agyát,
- (5) de amikor megkaptam a leveleket hát szénnéröhögtem magam,
- (6) Már szénnékísérleteztem magam, de eddig alulmaradtam a szuper menzagrízzel szemben.
- (7) ő meg legalább szénnéizgulja magát az úton,

**szerte<sub>1</sub> ‘írány (egy helyről több helyre)’**

- (1) a fölöttünk keringő récecsapatok szertejvedtek.
- (2) ki-ki szertecammog a maga emeletes ágyaktól zsúfolt vagonjába,
- (3) s alóla hófehér hajgubanca rakoncátlanul szertekandikált

**szerte<sub>2</sub> ‘egészből részekbe’**

- (1) az szinte szerterobban itt 1940-ben.
- (2) szertehullhat az egész politikai építmény.
- (3) hogy az a bűdös, kemény paraszt feje szertehasadjon!

**szét<sub>1</sub> ‘írány (egy helyről több helyre)’**

- (1) s egy pillantás alatt szétsuttogta a suddognivalóját.
- (2) - Széttologatja a bútorokat.
- (3) A vetőmag és a virágpor emellett nagy távolságokon át is széthordódhat.

**szét<sub>2</sub> ‘egészből részekbe’**

- (1) és a szegények között szétoztassék.
- (2) Mivel úribb helyeken a szar és az esővíz szétválasztatik,
- (3) és egyenletesen szétmázolgatja rajta a sarat.

**szét<sub>3 (+ magát)</sub> ‘túlzásba esve (károsodásig/kimerülésig)’**

- (1) mert a receptoraidat szétbutította a körülötted ömlő, primitív, szexuális szenny.
- (2) Szétgusztustalankodtad az esküvőjét is,
- (3) mert különben szétszennyeződik az egész - Linux ide, Linux oda.
- (4) hogy széthallgattatta velünk DKay Barcelona-himnuszát,
- (5) Szétkarcoltak egy hajótestet,
- (6) mer szétwalkmanezem a fületem.

- (7) aki pedig nem akar az szétkockulhatja magát.
- (8) Vagy mert az exnéptáncos végre széttombolhatja magát,
- (9) ha valaki szétkoncentrálja magát.

**tele ‘teljes területet/felületet megtelítve’**

- (1) telepöfögik azzal a rohadt benzingőzzel a levegőt
- (2) amikor telehuhogta a világot rémképeivel,
- (3) Nem rossz emberek, csak teledinnyézik még az anyjukszűzmáriáját is -
- (4) hogy a mondandójukat teleszmájlizzák.
- (5) jól telecsöpögteti a padlót azzal a ragacsos szarral,
- (6) a mellemet is teleszilikonozták,

**tova ‘irány (elfelé, ált. folytatólagosan)’**

- (1) Aztán a kerítés mellett tovasurrant.
- (2) miközben tovazörgött a szekér.
- (3) káromkodva tovatipeg a hídon.

**tovább<sub>1</sub> ‘irány (elfelé, ált. folytatólagosan)’**

- (1) ahonnan továbbszökkenhet a szomszédos épületre.
- (2) Lomhán továbblődörögnek.
- (3) és továbbkopog a kövön.
- (4) Majd továbbszédelegtek.

**tovább<sub>2</sub> ‘folytatva vmit’**

- (1) hogy továbbvéreztetik az önkormányzatokat.
- (2) aztán a posztfeministák továbbvirágoztatták a mítoszt,
- (3) Továbbrepedezik az amerikai pénzügyi rendszer,

**tönkre(+ magát) ‘túlzásba esve (károsodásig/kimerülésig)’**

- (1) hogy tönkremagányosodtam itt a pusztai mezőn, a pangó vizek mentén.
- (2) A rendszerváltáson átnyúló nyálcsápjával tönkrecsócsált egy barátságot.
- (3) évente tönkresírt egy-egy tucát zsebkendő.
- (4) Én tönkreröhögtem magam ezen a hozzászóláson.
- (5) és felelőtlenül tönkrevásárolják magukat.

**túl<sub>1</sub> ‘irány (vmi/vki másik oldalára)’**

- (1) Ha valahol véletlenül túlfestünk a határon,
- (2) meg valami nála hatalmasabb erő túllódította a bóján,
- (3) amikor túlhempereg a fotel peremén.

**túl<sub>2</sub> ‘túlzásba esve’**

- (1) túlhajlongja a szerfölt udvarias professzort,
- (2) és a legtöbben túlpánikolták a helyzetet.



- (3) Nehogy túlalkalmazkodjam magam,
- (4) mintha kicsit túlmenőznék a dolgot,

#### **újjá ‘ismételten (és megújítva)’**

- (1) a templomot kívülről újjáfestették.
- (2) Emellett a könyvben Aragorn már A Gyűrű Szövetségében újjákovácsoltatja a Narsilt,
- (3) a triumvirek újjáosztják a birodalmat

#### **újra ‘ismételten’**

- (1) majd újrakristályosítják az anyagokat,
- (2) Ezek az identitások, ezek folyamatosan újraalkotódnak,
- (3) és újranevesítették a földrészt még a földművelők előtt.
- (4) mert háromszor is újrasminkeltette magát,

#### **utána ‘irány (vmi/vki után)’**

- (1) hogy az ember a szépnek utánacammoghat.
- (2) A villamosjegyek és egy használt papírzsebkendő utánahullanak.
- (3) A limuzint mindenhová utánszállították helikopteren,
- (4) hogy a lépcsőn utánadübörognek-e a rohamrendőrök.
- (5) amikor utánadünnyögtem,
- (6) Innen-onnan utánakurjantottak,

#### **végig<sub>1</sub> ‘irány (kezdőponttól a végpontig)’**

- (1) és végigszökkent a köveken.
- (2) Majd végigballettozott a monitoron,
- (3) ilyenkor a ragadozó nyulak veszedelmes füleiket végiglapítják magukon

#### **végig<sub>2</sub> ‘teljesen, végig’**

- (1) és végigfloodolom a teljes fórumot egy-egy hozzászólással,
- (2) Én ezt végigstudíroztam alaposan,
- (3) és szép sorban végigkritizálja riválisait.

#### **végig<sub>3</sub> ‘eltölt egy időszakot vmivel’**

- (1) s végigfurulyázgatta a hosszú estét.
- (2) Ha végigbájolgom a napot,
- (3) és inkább végigkockulja az éjszakát.
- (4) hogy a srác utána egy egész indiai utat végigagonizált.

#### **vissza<sub>1</sub> ‘irány (egy korábbi helyre)’**

- (1) Aztán visszablattyogtunk a sáron keresztül,
- (2) Visszasuttyantam a szobába.
- (3) és mindenki visszacsődül a táncterre.
- (4) Ahogy polipkarjait egymás után visszakapdosta,

**vissza<sub>2</sub> ‘változik/változtat vmivé/vkivé (egy korábbi állapotra)’**

- (1) és látható módon visszanépfrosztítják ezt a lapot.
- (2) ekkor visszaifjtjuk,
- (3) A csitrik és suhancok nagykorúságban, az öregek visszakiskorúsodtak.
- (4) amellyel az elmagyarosított románokat visszarománosítanak.

**vissza<sub>3</sub> ‘reagál, válaszol vmire’**

- (1) s bár sebesen visszaszervuszolok,
- (2) egypáran visszafasiztáztak,
- (3) meg csak úgy visszapampogjanak!
- (4) elképedve visszadadog:

## Magyar nyelvű összefoglaló

A disszertációmban a magyar igekötők állományát és az igekötős szerkezeteket vizsgálom korpuszvezérelt módszerrel. Fő kutatási kérdéseim a következők: (1) Mely lexikai elemeket soroljuk az igekötők közé, és mi alapján végezzük a besorolást? (2) Milyen mintázattípusokat mutatnak az igekötős szerkezetek, mikor és mennyire távolodhat el az igekötő az igétől, igenévtől, igei származéktól? (3) Hogyan írhatjuk le az igekötők produktív kapcsolódási mintáit, és ez alapján milyen megállapítások tehetők az igekötők szemantikájáról?

Az első kérdést illetően abból indulok ki, hogy az igekötők és egyéb, pusztán névszói igemódosítók között nem húzható éles határ. Ehhez a felfogáshoz leginkább a prototípus-elmélet illeszkedik, amelynek keretében megállapítom az igekötőség morfoszintaktikai jegyeit és a korpuszméréseim alapján a jegyek értékét az egyes igekötő-jelöltekre, majd megvizsgálom a jegypárok közötti korrelációkat. Végül elvégzem az igekötők jegyalapú osztályozását, amelynek alapján négy csoportot különítek el: a prototipikus, a centrális, a félperiférikus és a periférikus igekötőket.

A második kérdés kapcsán először szinkrón korpuszvizsgálatot végzek, amelynek hangsúlyos része a finit igei szerkezetek feltérképezése. Kimutatom, hogy a hátravetett igekötők igétől való távolságát befolyásolhatják fonológiai tényezők és a szöveg szerkesztettsége is. Az igekötők disztribúcióját igenevek, igei származékok esetében is vizsgálom, és ennek során bemutatok néhány ritkán tárgyalt szerkezetet (pl. *szétválni szét lehet, a „be nem indulós Trabant”*). Ezután egy diakrón vizsgálatban mérem ki a prototipikus igekötők mondatbeli helyét az ómagyar kortól napjainkig terjedő szöveganyagon. Így láthatóvá válik a fókuszált összetevőt tartalmazó mondatok arányának növekedése, valamint a fordított szórendű tagadás térnyerése a megszakítottal szemben.

A harmadik kérdés megválaszolására olyan módszert dolgozok ki, amely ‘igekötő – képző – vonzatkeret’ hármasok (pl. *el-X-elődik valami, tönkre-X-el valamit*) korpuszvezérelt vizsgálatán alapszik. Ismertetem az igealkotás három leggyakoribb módját: a névszói és az igei bemenetű igeképzést, valamint a hangzásséma-alapú igealkotást. Ezután bemutatom a PREVCONS táblázatot, amely 21 038 igekötős igei hapaxot tartalmaz, és a fentebb említett hármasok kereshetőségével lehetővé teszi a produktív kapcsolódási minták feltárását. Végül egy olyan kísérletet mutatom be, amelynek célja az igekötők jelentéseinek és a jelentések közti viszonyoknak a hálózat-szerű ábrázolása a PREVCONS alapján, egy ontológia formájában.

## Summary in English

The thesis investigates the class of Hungarian verbal particles and the particle-verb constructions by a corpus-driven method. Its main research questions are as follows: (1) Which lexical items can be classified as verbal particles, and what are the grounds of their classification? (2) What kinds of structural patterns do particle-verb constructions show, when and to what extent can a particle be separated from a finite/non-finite verb or a deverbal element? (3) How can we describe the productive particle-verb patterns, and – based on this – what conclusions can be made about the semantics of verbal particles?

As for the first question, I assume that there is a fuzzy boundary between verbal particles and other bare nominal verb modifiers. The prototype-theory seems to be best suited for the graded categorization needed here. Using this theory, I determine the typical morphosyntactic features of verbal particles and the value of each feature in the case of every possible verbal particle, based on corpus measurements. After that, I measure the correlation of each feature-pair. Finally, I perform a feature-based classification of verbal particles and decide to distinguish four categories: prototypical, central, semi-peripheral and peripheral verbal particles.

Regarding the second question, I first perform a synchronic corpus study, with putting an emphasis on the finite particle-verb constructions. I show that the distance of the finite verb and its postverbal particle can be influenced by phonological factors and by an ‘oral versus written’ distinction. I also examine the distribution of verbal particles in the case of non-finite verbs and deverbal elements. Doing so, I present some structures rarely discussed in linguistic literature. I also conduct a diachronic study which aims to quantify the changes of the prototypical verbal particles’ positions, from the Old Hungarian period to the present day. This study shows a monotonous increase in the proportion of sentences containing a focused constituent. Moreover, it shows how the negative sentences having ‘verb – negative particle – verbal particle’ order make headway against the ones having ‘verbal particle – negative particle – verb’ order.

To answer the third question, I develop a method based on the corpus-driven study of ‘verbal particle – derivational suffix – argument frame’ triplets. I present the three most common ways of word formation in Hungarian: verb formation from nouns and verbs, and thirdly, verb formation using sound patterns. After that, I present the PREVCONS table containing 21 038 particle verb

hapaxes. This resource makes it possible to explore the productive particle-verb patterns by the accessibility of the triplets mentioned above. Finally, I present an experiment which aims to represent the different meanings of verbal particles and the relationships between these meanings in a network-like structure based on PREVCONS, in the form of an ontology.