

## *Tanulmány*

Kalivoda Ágnes

### Az igekötők produktív kapcsolódási mintái

#### Abstract

Productive preverb-verb patterns show which groups of verbs a preverb can be combined with, depending on its various senses. The aim of this paper is to extract these patterns using a method based on the corpus-driven study of ‘preverb – derivational suffix – argument frame’ triplets. First, I take a close look at the most common verb formations in Hungarian: new verbs can easily be created from nouns, verbs and by means of sound patterns. After that, I present the PREVCONS database containing 21,038 preverb-verb hapaxes. This resource makes it possible to explore the productive preverb-verb patterns by the accessibility of the triplets mentioned above. Finally, I describe an experiment which aims to represent the different senses of preverbs and the relationships between these senses in a network-like structure based on PREVCONS, in the form of an ontology.

*Keywords:* preverb, verbal particle, productive patterns, verb formation in Hungarian, ontology

#### 1 Bevezetés

„[...] a bölcelet nyelve valósággal elizélődött: megizélték, tönkreizélték, agyonszöszölték és elmaszlagolták a filozófiát.”<sup>1</sup> Bár ez az 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’ hármasonak (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 hármasonak a korpuszvezérelt vizsgálata közelebb visz minket ahhoz, hogy megismerjük az igekötők<sup>2</sup> produktív kapcsolódási mintáit. A vizsgált jelenség azonban annyira szerteágazó, hogy ennél a közelebb vitelnél többre nem vállalkozhatok.

A produktív kapcsolódási minták azt adják meg, hogy az igekötő különféle jelentéseiben milyen igecsoportok tagjaihoz járulhat. Ehhez három tényezőt szükséges figyelembe 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,

<sup>1</sup> Részlet Határ Győző *Darályvilág Buzdugániában* (2001) című könyvéből.

<sup>2</sup> A tanulmányban szó lesz számos olyan lexikai elemről, amelyet indokoltabb lenne ‘igekötőszerűségnek’ nevezni, mivel a státusza mind a szakirodalom, mind a korpuszadatok alapján bizonytalan. A továbbiakban az egyszerűség kedvéért ezekre is az *igekötő* terminussal utalok. Az igekötő-állomány meghatározásának problémáiról ld. többek között Jakab (1976), Komlósy (1992), Forgács (2005) és Kerekes (2011) munkáit.

sok esetben már az igekötő és a vonzatkeret együttese elég ahhoz, hogy tudjuk, milyen – szemantikailag hasonló – igék lehetnek a szerkezetben.

Ebben a tanulmányban az első tényezőt vizsgálom meg tüzetesen azért, hogy 56 igekötő produktív jelentéseit foglalom össze egy ontológiában. A második és a 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 az igei bemenetű igeképzésről, valamint a hangzásséma-alapú igealkotásról. Ezt a két újonnan létrejött erőforrásnak, a PREVCONS-nak és az igekötők ontológiájának bemutatása követi. Végül összegzem a legfontosabb eredményeket.

## 2 A kutatás módszere

A kutatás alapjául a PREVLEX szolgált, amely az igekötős igék jelenleg legbővebb (53 503 szavas), manuálisan ellenőrzött, nyílt hozzáférésű adatbázisa (Kalivoda 2019). Az igekötős igék mellett fel van tüntetve a Magyar Nemzeti Szövegtár 2.0.4 korpuszban (Oravecz et al. 2014) – a továbbiakban MNSZ2 – mért tokengyakoriságuk is.

A PREVLEX-ben szereplő igekötős igék közül csak a hapaxokat, azaz 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 – legalábbis nagyobb eséllyel az, mint a gyakori szavak esetében. A hapaxok közül is csak azokat őriztem meg, amelyeknél az igekötő terjeszkedő produktivitása<sup>3</sup> 0-nál nagyobb volt. A PREVLEX 122 ilyen igekötőt tartalmaz, ezek összesen 21 038 hapaxban jelennek meg.<sup>4</sup>

A kigyűjtött hapaxokat az emMorph (Novák et al. 2016, 2017) segítségével elemeztem. Ez egy nyílt forráskódú, automatikus morfológiai elemző rendszer, amely a szóalakot morfémákra 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 morfológiai szerkezetük: sem az, hogy hol van az igekötő és ige közötti határ, sem az, hogy az ige adott esetben szóképzés eredménye-e. Az (1) példa azt mutatja, hogy milyen elemzéssel jelenik meg az *elmamlaszkodik* szó az MNSZ2-ben, és milyen elemzés érhető el hozzá az emMorph-ban.<sup>5</sup>

- (1) a. *elmamlaszkodik/elmamlaszkodik/IK.IGE.e3*  
 b. *e1[/Prev]=e1+mamlasz[/N]=mamlasz+kodik[\_Vbz:k0d/V]=kod+ik[Prs.NDef.3Sg]=ik*

<sup>3</sup> A terjeszkedő produktivitás (Baayen 2009: 905) arról ad jóslatot, hogy egy affixumnak várhatóan mekkora szerepe lesz a szóalkotásban a közeljövőben. Ezt úgy kapjuk meg, hogy az affixumot – itt: igekötőt – tartalmazó hapaxok darabszámát elosztjuk a korpuszban – itt: a PREVLEX-ben – található összes hapax darabszámával.

<sup>4</sup> Előfordul, hogy lényegében azonos igelemmák tűnnek fel *-ik-es* és *-ik-telen* alakban is (pl. *ajándékozik* – *ajándékozik*, *akcióz* – *akciózik*). Emiatt „emberi szemmel” nézve kevesebb hapaxot kapnánk, mint automatikus módszerrel.

<sup>5</sup> <https://emmorph.herokuapp.com/> (A morfológiai szerkezet az *analyze* opcióval nyerhető ki.)

Az emMorph nélkülözhetetlen ehhez a munkához, mivel manuálisan szinte lehetetlen ekkora adathalmazt feldolgozni. A fedése kiemelkedőnek számít: 92,6% körüli értékű (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 hibák javítása csak manuális munkával lehetséges, amelyre az átnézendő adat nagy mennyisége miatt nem vállalkoztam. Érdemes tehát szem előtt tartani, hogy a kutatás eredményei nem hibátlan adathalmaz alapján álltak elő, ezeket kissé módosíthatja a szóban forgó javítási munka elvégzése.

Az emMorph a produktív képzőket ismeri fel aktív – azaz önállóan is használt – tövek mellett. A kutatás előrehaladtával világossá vált, hogy szükségem van olyan szerkezetek 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 a -(s)Ul képzőt, ha az több szótagú szóhoz 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 bemenetétől szolgáló szavak listáját – például névszói bemenetű igeképzésnél a névszók listáját – megvizsgáltam word2vec szóbeágyazással is (Siklósi & Novák 2016; Novák et al. 2016). Ennek az eljárásnak 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 osztályozásra, amelyet hagyományos módszerekkel hatalmas időbefektetés árán lehetne 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 a MNSZ2-ből azokat a mondatokat, amelyek a hapaxokat tartalmazzák. Ezeket Sass (2011) módszerével tagmondatokra bontottam, és a releváns tagmondatokból ki nyertem a hapaxok feltételezett vonzatkeretét. Ezek az adatok – kiegészítve a fentebb tárgyalt morfológiai és szemantikai információval – alkotják a PREVCONS-ot, amelyet a 4. fejezet mutat be részletesen. A PREVCONS segítségével elemezhetővé váltak az igekötők produktív kapcsolódási mintái.

### 3 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* esetében pedig szinte halljuk

a hangot, amivel a cselekvés jár, bár a pontos cselekvést nehéz volna megnevezni. Ezeknek az igéknek a mintájára képesek vagyunk még számos 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 jelentőségük 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 adataival a 21 038 igekötős igei hapax (ld. a 2. fejezetet). Ezeket úgy adtam meg az emMorph-nak, hogy az igekötőt elhagytam belőlük – így 11 602 igealak morfologiai 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 alakokkal.

A továbbiakban külön alfejezetek foglalkoznak a három igealkotási móddal, a névszói és az 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 az 1. számú mellékletben található. A három igealkotási mód jól áttekinthető, egyoldalas összefoglalását tartalmazza a 2. számú melléklet.

### 3.1 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!* → *kezitcsókolomoz*).

A névszóhoz járuló igeképzők két csoportba sorolhatók annak alapján, hogy milyen tövekhez kapcsolódnak. Eszerint beszélhetünk valódi (aktív) tövekhez, valamint fiktív (passzív) tövekhez járuló igeképzőkről. Az előbbi esetben a tö önállóan is létezik a magyar nyelvben, míg az utóbbiban mindig képzővel együtt fordul elő. Sokan fiktív tövel működő igeképzésként elemzik a hangutánzó és hangulatfestő igéket is, én ettől el fogok térni (ennek részleteit ld. a 3.3. fejezetben).

A **valódi névszói tövekhez járuló, produktív igeképzőkről** nyújt áttekintést az 1. táblázat. Ezekről részletesen olvashatunk Kiefer és Ladányi (2000a), valamint Ladányi (2007) munkáiban, és a korpuszadatok mellett én is elsősorban ezt a két művet használom fel az egyes igeképzők bemutatásához.

A leggyakrabban megjelenő képző, a **-Vz(ik)** alapértelmezettnek tekinthető a denominális igeképzésben. Ladányi a képző általános jelentését így fogalmazza meg: ‘N-nel kapcsolatban tesz valamit, N-nel foglalkozik’ (ahol az N a névszói bemenetet jelöli) (2007: 76). Ez egy annyira kiüresedett jelentés, hogy tulajdonképpen azt is mondhatjuk, a képzőnek csak funkciója van. Az emMorph különválasztja az *-ik*-re végződő variánst, amelyet intranszitivnak 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 – intranzitívnak elemzett ige mellett is megjelenik tárgy –, ezért nem különítem el a két változatot.

képző	példák	db
-Vz(ik)	<i>zebracsíkoz, doktoroz, kávézik</i>	2 558
-(s)Odik	<i>babásodik, jegesedik, vörösödik</i>	638
-VI	<i>falfirkál, offitopikol, chatel</i>	525
-(V)(s)ít	<i>bénít, felhőkarcolósít, vidámít</i>	420
-(s)Ul	<i>halkul, szelidül, zombul</i>	205
-(s)kOdik	<i>flegmáskodik, izmoskodik, trollkodik</i>	180

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

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 igepároknak 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).

Az 1. 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 bár igaz, hogy a **-(s)Ul** általában egy szótagú tövek mellett jelenik meg, jócskán találunk példát ennek az ellenkezőjére is (pl. *magányosul, kristályosul, anyagiasul, mémesül, kereszténydemokratul*). Az ilyen eseteket Ladányi úgy írja le, hogy a vetélkedő képzők közül általános esetben is a speciálisat alkalmazzuk, a képző ezáltal túllépi a hatókörét (2007: 253–254). Szintén érdekes jelenség, hogy a **-(s)Ul** képző nem kizárólag melléknévhez, hanem néha főnévhez is kapcsolódhat, pl. *szenttehenül, pépül*.

Több szakirodalomban – Kiefer és Ladányi (2000a: 154–155), Ladányi (2007: 246) – olvasható 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 – legalábbis 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. Életlent jelentő névszó is szolgálhat bemenetként (pl. *asztalkodik*, *póznáskodik*), de ez nagyon ritka jelenség.

Az **idegen eredetű, fiktív tövekhez járuló igeképzőket** a 2. táblázat foglalja össze. Mivel ezekben a névszói fő önmagában nem létezik a magyar nyelvben, joggal merülhet fel a kérdés, hogy van-e értelme a fő és a képző különválasztásának (az emMorph például nem jelez szóképzést az ilyen igeik esetében). Az adatok között találtam néhány példát arra, hogy magyar szavakban is működnek ezek a képzők (pl. *bolondizál*, *kántorizál*), ezért döntöttem a különválasztásuk 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

2. 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 produktívoknak. Ide sorolható az **-Ász(ik)** képző (pl. *madarászik*, *egerészik*, *bolhászik*), valamint a **-(V)gOl** (pl. *szamaragol*, *kacsagol*, *elefántogol*).

### 3.2 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 3. táblázat ad áttekintést.

A leggyakoribb a műveltetést kifejező **-(t)At** képző, ennek – és a műveltetés egyéb kifejezési módjainak – részletes elemzését adja Komlós (2000). Ezt az **-Ódik** követi, amely mediális igeik képzésére alkalmas. A mediális elnevezés arra utal, hogy az ilyen igeik „középuton” vannak a cselekvő és szenvedő igenem között (az igeiben 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 igeiknél jelenik meg (pl. *ad* → *adogat*), a kicsinyítő pedig azoknál, amelyek hosszabban tartó folyamatot jelölnek (pl.

*olvas* → *olvasgat*) (Kiefer és Ladányi 2000a: 199). A két funkció nem független egymástól, sokszor együttesen vannak jelen.

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

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

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 (létige + -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. Nagy csoportot alkotnak azok, amelyek gyakorító-kicsinyítő funkcióval bírnak, tehát az -(O)gAt alternatívái. Ide sorolható a -kOd (pl. *lopkod, tömköd*), a -dOs (pl. *vágdos, öldös*), a -dOgAl (pl. *hulldogál, nézdegél*), a -dAl (pl. *rugdal, szökdel*), a -dÁcsOl (pl. *bukdácsol, nyögdecsel*), az -Ong (pl. *jajong, forrong*) és az -igál (pl. *huzigál, dobígál*). Ezek a képzők paradigmákba is szerveződnek, erről bővebben ld. Ladányi (2007: 101–105) esettanulmányát.

### 3.3 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.

A szakirodalom jelentős része igeképzésként tárgyalja ezeket az igealkotási módokat is, azzal a kitételrel, hogy a szótó mindig fiktív. Ilyen elemzést választ pl. Rebrus (2000: 778–779), Kiefer és Ladányi (2000a: 201) és Ladányi (2007: 244–245). A képzős elemzés helyett eddig kevesen választották a hangzásséma-modellt. A kevesek között az egyik legalaposabb munka Benő és Szilágyi N. cikke (2015), amely az én munkámhoz is nagyban hozzájárult.

A hangzásséma-alapú elemzés meglehetősen eltér a hagyományosabb megközelítésektől, ezért nem meglepő, hogy nem nyerhető ki az emMorph-ból. A hangzássémába 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. 4. táblázat). Tagjainak a szerkezeti hasonlóságukon túl közös szemantikai tulajdonságuk az, hogy hosszabb ideig tartó vagy ismétlődő eseményt fejeznek ki.

Kalivoda Ágnes:  
 Az igekötők produktív kapcsolódási mintái  
 Argumentum 17 (2021), 56-82  
 Debreceni Egyetemi Kiadó  
 DOI: 10.34103/ARGUMENTUM/2021/4

általános séma és példák	db	megszorított séma és példák	db
(C)(C)VC:Og <i>battyog, hűmmög, billeg</i>	62	(C)CO <sub>1</sub> C:O <sub>1</sub> g <i>szötyög, kerreg, csetteg</i>	6
(C)(C)VCOg <i>krárog, 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>ögyeleg, 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

4. 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 hangzóságuk (Kassai 1994: 631) és a fonológiai jellemzőik alapján is (Siptár 1994: 185).

Bizonyos -Og végű igék -OrÁsz(ik) végződésel 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 az 5. táblázat példái illusztrálják). Mindháromra igaz, hogy mozzanatosak, azaz pillanatnyi eseményt fejeznek ki. Az -An ilyen tekintetben az -Og funkcionális párja (vö. *csattan* – *csattog*, *zöttyen* – *zötyög*). 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. *meглиbben a függöny* – *a szél meглиbbenti a függönyt*). Az -int az -Ant-hoz hasonlóan kauzatív, és azzal a jelentéstartással bír, hogy az esemény nagyon alacsony intenzitású.

Az 5. 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. 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/ (2015: 48).



Kalivoda Ágnes:  
*Az igekötők produktív kapcsolódási mintái*  
*Argumentum 17 (2021), 56-82*  
 Debreceni Egyetemi Kiadó  
 DOI: 10.34103/ARGUMENTUM/2021/4

<b>-An</b>		<b>-Ant</b>		<b>-int</b>	
<b>séma és példák</b>	<b>db</b>	<b>séma és példák</b>	<b>db</b>	<b>séma és példák</b>	<b>db</b>
(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

5. táblázat. Az -An, -Ant és -int végű hangzássémák gyakoriságuk szerinti csökkenő sorrendben.

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. 6. 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

6. 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 a 3.2. 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>6</sup> – illeszkedik.

### 3.4 A vizsgált igealkotási módok arányai az igekötős szerkezetekben

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

Elsőként érdemesnek tartottam megvizsgálni azt, hogy a három produktív igealkotási mód milyen arányban van jelen az igekötős hapaxoknál, valamint az igekötő elhagyásával kapott ige típusoknál. Ennek az eredményét foglalja össze 7. táblázat.

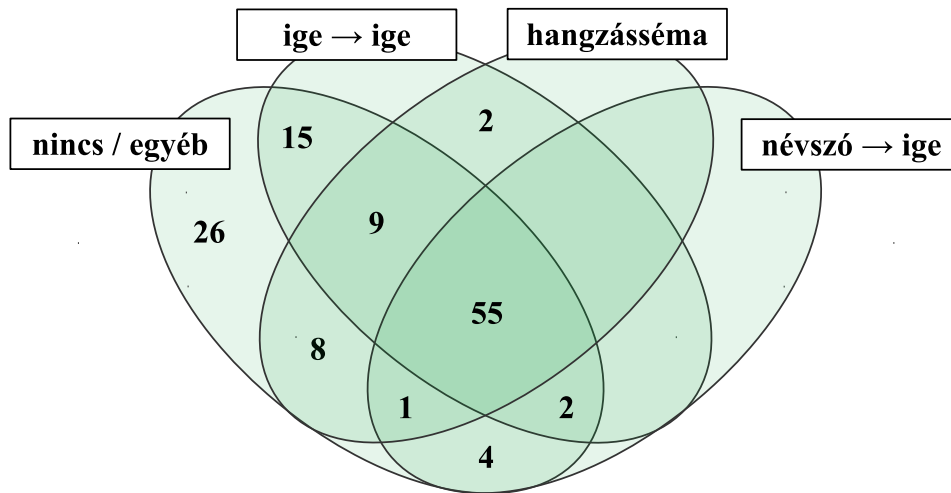
	igekötős hapaxok		igetípusok	
	db	%	db	%
<b>névszó → ige</b>	7 401	35,2	4 985	43,0
<b>ige → ige</b>	4 492	21,4	2 951	25,4
<b>hangzásséma</b>	1 987	9,4	799	6,9
<b>nincs / egyéb</b>	7 158	34,0	2 867	24,7
	<b>21 038</b>	<b>100</b>	<b>11 602</b>	<b>100</b>

7. táblázat. Igealkotási módok gyakorisága az igekötős igei hapaxok, valamint az igekötő elhagyásával kapott ige típusok 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 ige típusokat, akkor ez az adatnak majdnem a felét teszi ki. A hangzássémás ige aránya sem elhanyagolható: az igekötős hapaxok közel egytizedét (9,4%) ezek adják. Ezt fontosnak tartom hangsúlyozni, ugyanis úgy tűnik, hogy a nyelvészeti szakirodalom 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 az 1. 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.

<sup>6</sup> A szótagszerkezettel kapcsolatban egy megjegyzés: a CV-sémákat nem úgy kell értelmezni, hogy a C helyén tetszőleges mássalhangzó állhat. A szótagnak jólformálnak kell lennie, tehát a szótagmag felé a szegmentumok hangzóssága 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).



1. ábra. A 122 vizsgált igekötő megoszlása aszerint, hogy milyen igealkotási móddal keletkeztek az igék, amelyekhez kapcsolódnak. A halmazokban és azok metszeiteiben 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 igekötőszerűbbek, 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.

A 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.

#### 4 PREVCONS: Igekötős szerkezetek adatbázisa

A PREVCONS (*Preverb Constructions*) egy nyílt hozzáférésű adatbázis,<sup>7</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.

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, igék és bővítményeik tanulmányozására alkalmas, a PREVCONS pedig az igekötős szerkezetek vizsgálatára használható, akár úgy is, hogy a szerkezetben lévő igét nem vesszük figyelembe. Tehát kigyűjthetünk belőle ‘igekötő – képző – vonzatkeret’ mintákat gyakorisági adataikkal együtt (pl. *össze-N-*

<sup>7</sup> <https://github.com/kagnes/prevcons>

*Vz(ik) -t -vAl*, azaz „összevalamiz valamit valamivel” vagy ugyanez élőlényre utaló vonzatokkal), és egy következő lépésben megnézhetjük, hogy a mintákban milyen igék 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 praktikusabb lehet a TSV, a kézi adatelemzéshez pedig az SQLite-alapú megoldás. Az erőforrás 16 oszlopból áll, amelyek a következők:

1. *id*: Az igekötős ige egyedi azonosítója, ennek csak technikai szerepe van.
2. *prev*: Igekötő.
3. *verb*: Igelemma.
4. *actform*: Az igekötős ige szóalakja.
5. *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 csak 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 esetragok 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 helyettesíték tárgyesetű raggal, hanem teljes szóként őrzök meg. Az 5. 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 jósolható meg előre, hogy bizonyos esetragok 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.

6. *prev\_prod*: Az igekötő terjeszkedő produktivitása (bővebben ld. a 2. fejezetet).

7. `prev_hapaxes`: Az igekötőt tartalmazó hapaxok darabszáma.
8. `inputtype`: Az ige esetében azonosított szóalkotási mód. A *N* névszóból képzett, a *V* igéből képzett, a *PH* pedig hangzássémás ígét jelöl. Amennyiben az ige nem produktív szóalkotással jött létre, ebben az oszlopban kötőjel (-) áll.
9. `suffixtype`: A képző vagy hangzásséma típusa. Itt is kötőjel (-) áll, ha az ige sem produktív képzővel, sem hangzássémával nem elemezhető. Egyéb esetekben a 3. fejezetben használt jelöléseket alkalmazom.
10. `phonotype`: Amennyiben az ige hangzássémába illeszkedik, itt látható a teljes CV-séma.
11. `style`: A forrásmondat MNSZ2-ből kinyert stílusregisztere. Mivel hapaxokról van szó, minden mondathoz pontosan egy stílusregiszter tartozik. Ez a következő értékeket veheti fel: *hivatalos*, *sajtó*, *spok* (beszélt nyelvi, általában rádióműsor leírata), *személyes* (fórum-és blogbejegyzések), *szépirodalom*, *tudományos*.
12. `region`: A régió, amelyből a forrásmondat származik (szintén az MNSZ2-ben elérhető adat). Lehetséges értékei: *erdélyi*, *kárpátaljai*, *magyarországi*, *szlovákiai*, *vajdasági*.
13. `w2v_input`: Az automatikus klaszterezés bemeneti szava. Ide szóképzéssel létrejött ige esetében az alapszó kerül (pl. *trollkodik* → *troll*), hangzássémás ige esetében pedig az igelemmával egyezik meg a mező tartalma.
14. `w2v_cluster`: Annak a klaszternek az azonosító száma, amelybe a `w2v_input` mező szava került a szóbeágyazás során (erről bővebben ld. a 2. fejezetet). Mivel külön lépésekben történt a különféle igealkotási módokkal létrejött igék 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.
15. `clause`: Az a tagmondat, amely az igekötős ígét tartalmazza. A tagmondathatárok megállapítása Sass (2011: 34–37) módszerével történt.
16. `sent`: A teljes mondat, amely az igekötős ígét 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 igekötős ige csak a tulajdonnév ismeretében volt igazán értelme, ilyen volt például a következő eset: *Ezek túladyendrézik magát Ady Endrét, [...]*

A PREVCONS-szal számos kérdés vizsgálható, például az, hogy melyik igekötő milyen stílusregiszterben van jelen leginkább, vagy egyes igekötők mellett mennyire sokszínű vagy épp kötött a vonzatkeret. A legnagyobb jelentősége mégis az, amiről a fejezet elején is szó volt: segítségével az ‘igekötő – képző – vonzatkeret’ hármasok a konkrét igék kihagyásával vizsgálhatók. Erre látunk példát a 8. táblázatban.

Kalivoda Ágnes:  
*Az igekötők produktív kapcsolódási mintái*  
*Argumentum 17 (2021), 56-82*  
 Debreceni Egyetemi Kiadó  
 DOI: 10.34103/ARGUMENTUM/2021/4

<b>igekötő</b>	<b>képző</b>	<b>vonzatkeret</b>	<b>db</b>
<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)UI		41

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

A hármasok lekérdezése után, egy következő lépésben meg lehet nézni az ezekhez tartozó igéket. Célszerű a találatokat az automatikus klaszterszámuk alapján rendezni, hogy a szemantikailag hasonló igék egymás mellé kerüljenek. Ezt a 9. táblázat mutatja be.

<b>igekötő</b>	<b>ige</b>	<b>klaszter</b>	<b>tagmondat</b>
<i>el</i>	<i>emdéefesedik</i>	1	és elemdéefesedett.
<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,

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

A legjobban akkor látható, hogy a szóbeágyazás milyen jól ragadja meg a szavak szemantikai hasonlóságát, ha a klaszterszám szerinti rendezést a hangzássémás igéken végezzük el. Ekkor kétféleképp csoportosulnak az igék. Az egyik lehetőség az, hogy tipikusan ki/mi adja ki az igével jellemzett hangot. Ezt látjuk a (2a) és a (2b) példában: az előbbi igéi folyadékok, az utóbbi állatok hangjára utalnak. A másik lehetőség az, hogy milyen hangja/hangulata van egy-egy cselekvésnek. Így például a (2c) igéi élőlény mozgását, a (2d) igéi élőlény nézését írják le.

- (2) 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 Az igekötők ontológiája

Az igekötők jellemzően poliszém, azaz több jelentéssel bíró lexikai elemek. Az egyes jelentéseik rendszerint 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 kapcsolattrendszer 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 hálózatszerű 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 leíró nyelvet.<sup>8</sup> Ezt azonban nem tartom nagy hiányosságnak, mert a céloom nem a közvetlen nyelvtechnológiai alkalmazhatóság volt, hanem az, hogy nyelvészeti motivációval látassam és magyarázzam az igekötők szemantikáját. Az ontológia reprezentációs formája egy síkba rajzolható gráf, amely PDF formátumban érhető el.<sup>9</sup>

Bár ilyen jellegű ontológia korábban nem készült a magyar igekötőkről, az igekötők szemantikájának 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 (2000b: 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). Hegedűs (2020) az *újra*, *újjá* és *vissza* igekötők szintaktikai és szemantikai jellegzetességeit ismerteti.

<sup>8</sup> Az angol nyelvre több olyan formális ontológia is készült, amelyben kitüntetett figyelmet szenteltek az igekötőszerű elemek, igei partikulák részletes szemantikai leírásának, erről bővebben ld. Bhatia et al. (2018), McShane et al. (2005).

<sup>9</sup> <https://github.com/kagnes/prevonto>

Az említett szakirodalmak közös vonása tehát az, hogy az igekötő-állományból néhány lexikai elemet 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 halmozását, 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 5.2. 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 vonzatkezminták 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.

### 5.1 *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 az 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 tendenciaszerűnek 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, beleoffol* 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 a 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 a (3) példásor mondatai. Ezekben a következő tömörítéseket látjuk: a (3a)-ban a *lenéptáncolom* a ‘néptáncolással ledolgozom’, a (3b)-ben a *ledinnyéztem* a ‘dinnyével lefizettem’, a (3c)-ben a *kiuhuskodtak* az ‘uhu szemeihez hasonlóan néztek kifelé’ rövidebb megfogalmazása.



- (3) 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 igekötőcsoport 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 a (4) példásor mondataiban.

- (4) 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 (4a) mondat mintájára működne az is, hogy *a kisfiú elpityeregte magát a cukrászdába*.

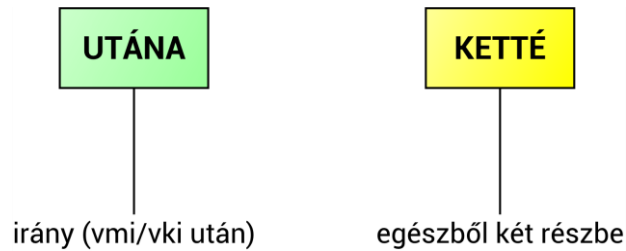
Végül fontos megemlíteni, hogy az aktuális jelentések mindig a szövegkörnyezettől függenek. Így például az (5a) mondatban a *letejez* jelentése az, hogy ‘tejjel véletlenül leönt’, míg az (5b)-ben látható *lespagettiz* azt jelenti, hogy ‘tévesen spagettinek nevez’, bár a *tej* és a *spagetti* szemantikai kategóriája közel azonos, és a vonzatkeret is egyezik.

- (5) 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 befolyásoló tényező az egyéni intuíció. Ha több ember kapta volna a feladatot, hogy a PREVCONS alapján készítsen ontológiát, valószínűleg nem született volna két egyforma végeredmény.

## 5.2 Ontológiai entitások: igekötők és jelentések

Az ontológia kétféle entitásból áll – igekötőkből és ezek jelentéseiből –, amelyeknek a gráfban csomópontok felelnek meg (ld. 2. ábra).

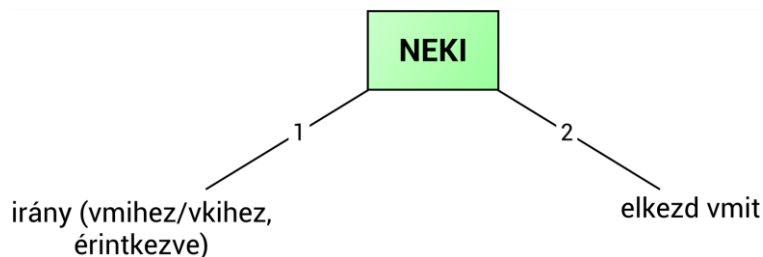


2. ábra. Két egyszerű példa az entitásokra: az utána és a ketté igeekötők a jelentéseikkel. A zöld háttér azt mutatja, hogy az igeekötőnek van irányjelentése, a sárga háttér ennek a hiányát jelzi.

Csak azokkal az igeekötőkkel foglalkoztam, amelyeknél a PREVCONS-ban megfigyelhető az, hogy névszóból képzett igehez kapcsolódnak. A döntésemnek gyakorlati oka volt: a kutatás előrehaladtával egyértelművé vált, hogy a névszói bemenetek szemantikai kategóriái mondanak el a legtöbbet az igeekötő lehetséges jelentéseiről, ezeknek a vizsgálatával lehet a leghatékonyabban haladni. Miután összevontam az alakváltozatokat,<sup>10</sup> 56 elemzendő igeekötő maradt. Természetesen a későbbiekben további igeekötőkkel 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>11</sup> A *valami*, *valaki* stb. szavakat a lexikográfiában megszokott módon rövidítem: *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 2. ábrán látható, hogy az igeekötőt és a jelentést vonalak – a gráfban élek – kötik össze egymással. A legegyszerűbb és legritkább esetben az igeekötőhöz egyetlen jelentés tartozik, ekkor az összekötő él nem kap semmilyen címkét. Amikor egy igeekötőhöz több jelentés kapcsolódik, az élek számozottak (ld. 3. ábra). A számozási sorrend nem implicál semmit, tehát az 1-es nem jelenti azt, hogy a hozzá tartozó jelentés a legrégebb vagy legtipikusabb. Ez csak a jelentésekre történő 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 az ontológia github repozitóriumában érhető el.



3. ábra. A neki igeekötő két produktív jelentése.

<sup>10</sup> Az alakváltozatok a következők: *be* – *bé*, *bele* – *belé*, *fel* – *föl*, *felül* – *fölül*, *rá* – *reá*, *tele* – *teli*.

<sup>11</sup> További példákhoz ld. J. Soltész (1959: 30–33).

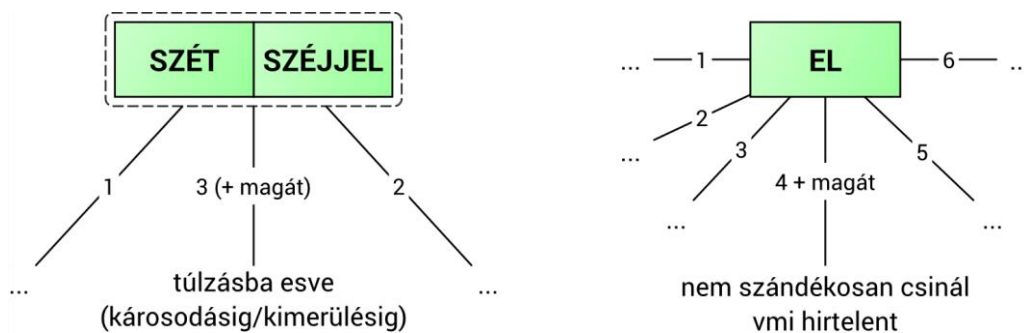
Az éleken nem csak számok szerepelhetnek, hanem szöveges tartalom is. A + magát címke azt jelenti, hogy az adott jelentés mindig együtt jár a *magát* vonzat megjelenésével. Ezt látjuk például a ‘*ki + ige + magát*’ szerkezetnél, amelynek jelentése ‘eleget csinál valamiből’, ld. a (6) példasort.

- (6) 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 **kimolotovkocktélózták 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* igekötők ‘túlzásba esve (károsodásig/kimerülésig)’ jelentésénél, amelyet a (7) példasor szemléltet.

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

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 tárgyeesetes 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 4. ábra szemléltet minden lehetséges élcímketípust.



4. ábra. A magát vonzat megjelenése a szét, széjjel és el igekötők esetében. 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.

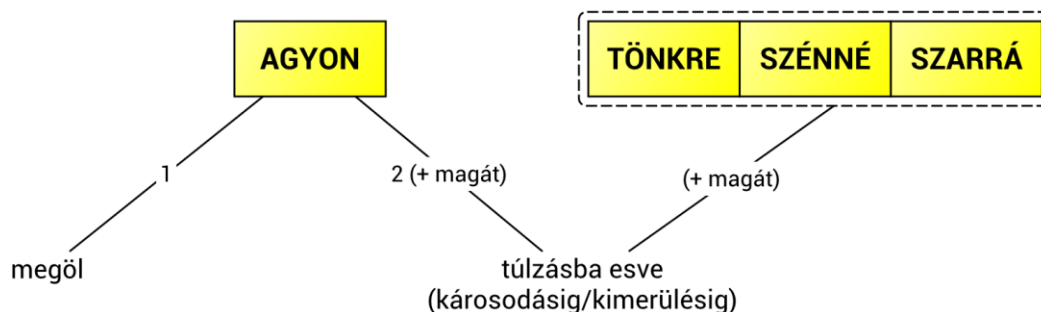
### 5.3 Szemantikai relációk

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 igeikötőbe vezetnek élek. Ezt látjuk az 5. ábrán a ‘túlzásba esve (károsodásig/kimerülésig)’ jelentés esetében. Az ábrán szereplő *agyon*, *tönk्रे*, *szénné* és *szarrá* igeikötőkön túl megtaláljuk ezt a jelentést a *szét*, *széjjel* és *össze* igeikötőknél is.

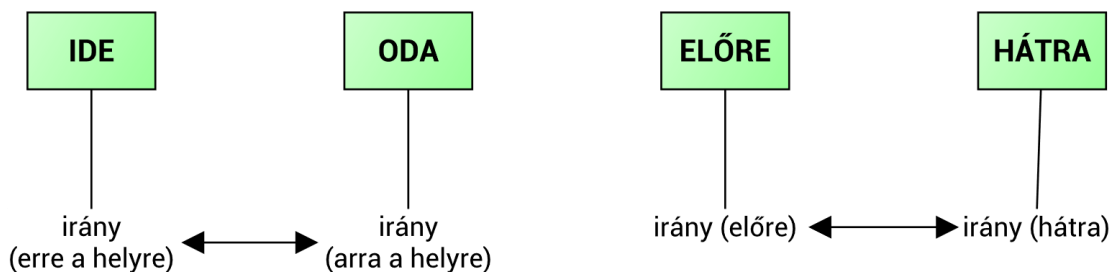
Ezt természetesen nem úgy kell érteni, hogy például a *szénné* bármikor kiválthatja az *agyon* igeikötőt annak a második jelentésében anélkül, hogy a mondatban jelentésváltozás történne. Egy kis konnotációbeli vagy stílusbeli különbség mindig adódik, ehhez ld. a (8) példasort. A *szénné* és a *szarrá* szlengben fordulnak elő, a *tönkre* kissé informális – bár egyre gyakrabban jelenik meg hivatalos szövegekben is –, az *agyon* pedig semleges, bármilyen stílusregiszterben megtaláljuk.



5. ábra. A szinonímia megjelenítése az ontológia grájfjában: egy jelentésből több igeikötőbe fut el.

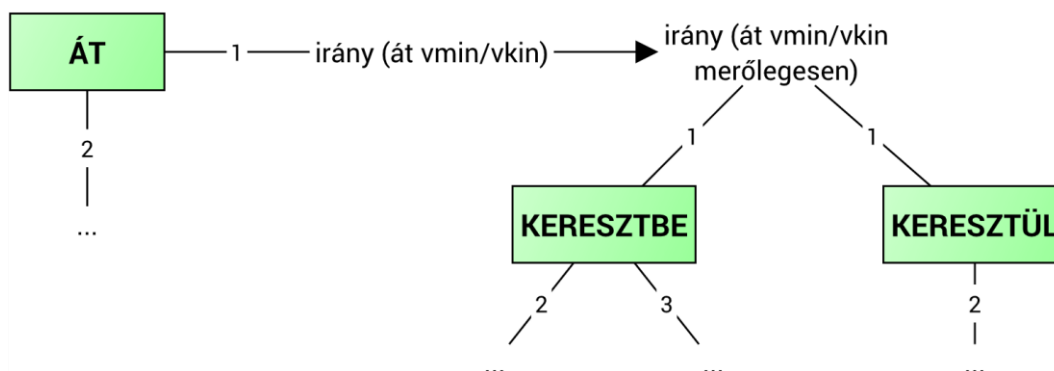
- (8) a. [...] a második itteni hetem végére az idegeim meglehetősen **agyoncincálódnak**.  
 b. *Én **tönkreröhögtem** magam ezen a hozzászóláson.*  
 c. [...] ő meg legalább **szénnéizgulja** magát az úton, vagy az autójában.  
 d. Csórikám, ezzel **szarrákereshetné** magát a nyugati világban [...]

Az antonímiát a jelentések közé húzott kétirányú nyíl jelzi (ld. 6. ábra). Valószínűleg ez a legkönnyebben felismerhető szemantikai viszony az anyanyelvi beszélők többsége számára. Tipikusan a konkrét irányjelentéseknél jelenik meg.



6. ábra. Az ide és oda, valamint az előre és hátra jelentései közti 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. 7. ábra). Ilyen viszony figyelhető meg például az *át* és a *keresztbe*, *keresztül* igeikötők irányjelentései között, amelyeknek a különbségét a (9) példasor mutatja be.



7. 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.

A (9a)-ban az *áthajókáztak* ige alapján még nem tudjuk, hogy az alanyok a Bering-szorost szél-tében 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*-énál, mert azt is magukba foglalják, hogy a mozgás iránya a (9b)-ben az útra, a (9c)-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 kvázi-szinonim, és csak a *keresztbe* fejezi ki mindig a merőlegességet.

- (9) 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 (10) példáit.

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

Zárásképpen fontosnak tartom egy példasorral illusztrálni, hogy a szemantikai viszonyok miatt az egyes jelentések szintjén értelmezhetők, és miért nem az igeikötők szintjén. Létezik ugyanis olyan eset, amikor ugyanannál az igeikötőpárnál 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 (11a) és (11b) példákat, a *szét* igeikötő pedig ‘egy helyről több helyre’, ld. a (11c) és (11d) példákat.

- (11) a. [...] *sokkal erősebb kötelék, mint amivel a világirodalom nagyjait **összelassozta**.*  
 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: ‘túlzásba esve (károsodásig/kimerülésig)’. A (12) példában a két igeekötő fel is cserélhető egymással úgy, hogy a mondat jelentése nem vagy csak árnyalatnyit változik.

- (12) 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.

## 6 Összegzés

A tanulmány egy sokoldalú témát vizsgált, 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 vonni a kutatásba. A produktív igealkotási módokhoz fonológiai és morfológiai, a vonzatkeretek kinyeréséhez szintaktikai, az igeekötők és ige csoportok jelentéseihez szemantikai vizsgálatra volt szükség.

A tanulmány első részében azzal a kérdéssel foglalkoztam, hogy hogyan hozunk létre új igeeket. Az igealkotás számos módja közül azt a hármat vizsgáltam, amely számottevően volt jelen az adatban. Ezek a névszói és az igei bemenetű ige képzés – és ezeken belül is a produktív képzők –, valamint a hangzásséma-alapú igealkotás. Az utóbbit Benő és Szilágyi N. (2015) nyomán elkülönítettem az ige képzésektől, és a leírását is máshogyan végeztem el (szótó-képző bontás helyett szótagszerkezeti sémákkal). Ennek a munkaszakasznak a jelentősége a meglévő ismeretek rendszerezésében, gyakorisági adatokkal való bővítésében, és pár ponton az átértékelésében mutatkozik meg (például annak kapcsán, hogy a -(s)Ul képző produktív-e vagy sem). Végül megvizsgáltam az igealkotási módok arányait az igeekötős hapaxok halmazán, és kimutattam, hogy a legjelentősebb a névszói bemenetű ige képzés, amely a hapaxok 35,2%-át adja, viszont mindössze 62 – az alakváltozatok összevonása után 56 – igeekötő kapcsolódik így képzett igehez.

A tanulmány második részében bemutattam a PREVCONS adatbázist, amelyet az igeekötős 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 kiváló erőforrások készültek a gyakori igeekötős igeekre 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 típuszámú-

kat tekintve nagyon sokan vannak). Másodszer, 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 hálózatszerűen, 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. Számos intuitív vonása ellenére az ontológia a kutatás fontos eredményének tekinthető, főként a benne megjelenő új szemlélet miatt.

## Irodalom

- Baayen, R. H. (2009): Corpus linguistics in morphology: morphological productivity. In: Lüdeling, A. & Kytö, M. (eds.): *Corpus Linguistics. An international handbook, Volume 2*. Berlin: Mouton De Gruyter, 900–919.
- Benő, A. & Szilágyi N., S. (2015): Hangzásséma és motiváltság a hangutánzó és hangulatfestő igéink körében. In: Kádár, E. & Szilágyi N., S. (szerk.): *Motiváltság és nyelvi ikonicitás*. Kolozsvár: Erdélyi Múzeum-Egyesület (EME), 43–57.
- Bhatia, A., Teng, C. M. & Allen, J. F. (2018): Identifying senses of particles in verb-particle constructions. In: Markantonatou, S., Ramisch, C., Savary, A. & Vincze, V. (eds.): *Multiword expressions at length and in depth: Extended papers from the MWE 2017 workshop*. Berlin: Language Science Press, 61–86.
- Forgács, T. (2005): Grammatikalizálódás az igekötők körében. In: Oszkó, B. & Sipos, M. (szerk.): *Uráli grammatizáló*. Budapest: MTA Nyelvtudományi Intézet, 88–116.
- Gerstner, K. (2006): A magyar nyelv szókészlete. In: Kiefer, F. (szerk.): *Magyar nyelv*. Budapest: Akadémiai Kiadó, 437–480.
- Hegedűs, V. (2020): Back to restitutes (again): A syntactic account of restitutive and counterdirectional verbal particles in Hungarian. *Acta Linguistica Academica* 67.3, 319–345. <https://doi.org/10.1556/2062.2020.00017>
- Indig, B., Sass, B., Simon, E., Mittelholcz, I., Kundrath, P. & Vadász, N. (2019): emtsv – Egy formátum mind felett. In: Berend, G., Gosztolya, G. & Vincze, V. (szerk.): *XV. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2019)*. Szeged: Szegedi Tudományegyetem, TTIK, Informatikai Intézet, 235–247.
- Jakab, I. (1976): *A magyar igekötők állományi vizsgálata*. Nyelvtudományi Értekezések 91. Budapest: Akadémiai Kiadó.
- J. Soltész, K. (1959): *Az ősi magyar igekötők: meg, el, ki, be, fel, le*. Budapest: Akadémiai Kiadó.
- Kalivoda, Á. (2019): Véges erőforrás végtelen sok igekötős ígére. In: Berend, G., Gosztolya, G. & Vincze, V. (szerk.): *XV. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2019)*. Szeged: Szegedi Tudományegyetem, TTIK, Informatikai Intézet, 331–344.
- Kassai, I. (1994): A fonetikai háttér. In: Kiefer, F. (szerk.): *Strukturális magyar nyelvtan 2. Fonológia*. Budapest: Akadémiai Kiadó, 581–666.

- Kerekes, J. (2011): Az igekötők meghatározásának problémái. In: Gécseg, Zs. (szerk.): *LingDok10: Nyelvészdoktoranduszok dolgozatai*. Szeged: Szegedi Tudományegyetem, Nyelvtudományi Doktori Iskola, 109–131.
- Kiefer, F. & Ladányi, M. (2000a): Morfoszintaktikailag semleges képzések. In: Kiefer, F. (szerk.): *Strukturális magyar nyelvtan 3., Morfológia*. Budapest: Akadémiai Kiadó, 165–214.
- Kiefer, F. & Ladányi, M. (2000b): Az igekötők. In: Kiefer, F. (szerk.): *Strukturális magyar nyelvtan 3., Morfológia*. Budapest: Akadémiai Kiadó, 453–518.
- Komlósy, A. (1992): Régensek és vonzatok. In: Kiefer, F. (szerk.): *Strukturális magyar nyelvtan 1. Mondattan*. Budapest: Akadémiai Kiadó, 299–527.
- Komlósy, A. (2000): A műveltetés. In: Kiefer, F. (szerk.): *Strukturális magyar nyelvtan 3., Morfológia*. Budapest: Akadémiai Kiadó, 215–292.
- Ladányi, M. (2007): *Produktivitás és analógia a szóképzésben: elvek és esetek*. Budapest: Tinta Könyvkiadó.
- Ladányi, M. (2008): Rendszeres igei poliszémia. In: Kiefer, F. (szerk.): *Strukturális magyar nyelvtan 4., A szótár szerkezete*. Budapest: Akadémiai Kiadó, 269–322.
- Ladányi, M. (2012): Igekötős igei kapcsolódási mintázatai. Vizsgálati lehetőségek. In: Tolcsvai Nagy, G. & Tátrai, Sz. (szerk.): *Konstrukció és jelentés. Tanulmányok a magyar nyelv funkcionális kognitív leírására*. Budapest: Eötvös Loránd Tudományegyetem, 71–84.
- McShane, M., Beale, S. & Nirenburg, S. (2005): *Disambiguating Homographous Prepositions and Verbal Particles In An Implemented Ontological Semantic Analyzer*. Szakmai jelentés #01-05. Baltimore: University of Maryland Baltimore County, Institute for Language and Information Technologies.
- Novák, A., Siklósi, B. & Oravecz, Cs. (2016): A New Integrated Open-source Morphological Analyzer for Hungarian. In: Calzolari, N., Choukri, K., Declerck, T., Goggi, S., Grobelnik, M., Maegaard, B., Mariani, J., Mazo, H., Moreno, A., Odijk, J. & Piperidis, S. (eds.): *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*. Párizs: European Language Resources Association (ELRA), 1315–1322.
- Novák, A., Rebrus, P. & Ludányi, Zs. (2017): Az emMorph morfológiai elemző annotációs formalizmusa. In: Vincze, V. (szerk.): *XIII. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2017)*. Szeged: Szegedi Tudományegyetem Informatikai Intézet, 70–78.
- Oravecz, Cs., Váradi, T. & Sass, B. (2014): The Hungarian Gigaword Corpus. In: Calzolari, N., Choukri, Kh., Declerck, Th., Loftsson, H., Maegaard, B., Mariani, J., Moreno, A., Odijk, J. & Piperidis, S. (eds.): *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC 2014)*. Reykjavík: European Language Resources Association (ELRA), 1719–1723.
- Rebrus, P. (2000): Morfofonológiai jelenségek. In: Kiefer, F. (szerk.): *Strukturális magyar nyelvtan 3., Morfológia*. Budapest: Akadémiai Kiadó, 763–948.
- Sass, B., Váradi, T., Pajzs, J. & Kiss, M. (2010): *Magyar igei szerkezetek – A leggyakoribb vonzatok és szókapcsolatok szótára*. Budapest: Tinta Könyvkiadó.
- Sass, B. (2011): *Igei szerkezetek gyakorisági szótára – egy automatikus lexikai kinyerő eljárás és alkalmazása*. Disszertáció. Budapest: Pázmány Péter Katolikus Egyetem.
- Siklósi, B. & Novák, A. (2016): Beágyazási modellek alkalmazása lexikai kategorizációs feladatokra. In: Tanács, A., Varga, V. & Vincze, V. (szerk.): *XII. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2016)*. Szeged: Szegedi Tudományegyetem, TTIK, Informatikai Intézet, 3–14.



Kalivoda Ágnes:  
*Az igekötők produktív kapcsolódási mintái*  
*Argumentum 17 (2021), 56-82*  
Debreceni Egyetemi Kiadó  
DOI: 10.34103/ARGUMENTUM/2021/4

---

- Siptár, P. (1994): A mássalhangzók. In: Kiefer, F. (szerk.): *Strukturális magyar nyelvtan 2. Fonológia*. Budapest: Akadémiai Kiadó, 183–272.
- Tolcsvai Nagy, G. (2015): *Az ige a magyar nyelvben. Funkcionális elemzés*. Budapest: TINTA Könyvkiadó.
- Törkenczy, M. (1994): A szótag. In: Kiefer, F. (szerk.): *Strukturális magyar nyelvtan 2. Fonológia*. Budapest: Akadémiai Kiadó, 273–392.
- Váradi, T., Simon, E., Sass, B., Gerőcs, M., Mittelholcz, I., Novák, A., Indig, B., Prószéky, G., Farkas, R. & Vincze, V. (2017): Az e-magyar digitális nyelvfeldolgozó rendszer. In: Vincze, V. (szerk.): *XIII. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2017)*. Szeged: Szegedi Tudományegyetem, Informatikai Intézet, 49–60.
- Váradi, T., Simon, E., Sass, B., Mittelholcz, I., Novák, A. & Indig, B. (2018): E-magyar – A Digital Language Processing System. In: Calzolari, N., Choukri, Kh., Cieri, Ch., Declerck, Th., Goggi, S., Hasida, K., Isahara, H., Maegaard, B., Mariani, J., Mazo, H., Moreno, A., Odijk, J., Piperidis, S. & Tokunaga, T. (eds.): *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. Miyazaki: European Language Resources Association (ELRA), 1307–1312.

Kalivoda Ágnes  
Nyelvtudományi Kutatóközpont, Lexikológiai Intézet, Lexikai tudásreprezentáció kutatócsoport  
Pázmány Péter Katolikus Egyetem, Nyelvtudományi Doktori Iskola  
H-1394 Budapest  
Pf. 360  
kalivoda.agnes@nytud.hu

## 1. számú melléklet: A fonológiai jelölések jegyzéke

C	mássalhangzó
C:	hosszú mássalhangzó
V	magánhangzó
V <sub>1</sub> ... V <sub>1</sub>	azonos magánhangzók
A	<i>a</i> vagy <i>e</i>
O	<i>o</i> , <i>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

## 2. számú melléklet: A produktív igealkotási módok áttekintő táblázata

<b>típus</b>	<b>altípus</b>	<b>db</b>	<b>példák</b>
<b>névszó + igeképző</b>	-Vz(ik)	2558	<i>akármiz, puskázik, rántottcsirkéz ...</i>
	-(s)Odik	638	<i>algásodik, britesedik, szótlanodik ...</i>
	-Vl	525	<i>homokzsákol, klikkel, loginol ...</i>
	-(V)(s)ít	420	<i>esősít, titkosít, borostásít ...</i>
	-ál	303	<i>dokumentál, konzervál, limitál ...</i>
	-(s)Ul	205	<i>ritkul, hülyül, nácul ...</i>
	-(s)kOdik	180	<i>okoskodik, ökörködik, anyáskodik ...</i>
	-izál	109	<i>hellenizál, romantizál, moralizál ...</i>
	-íroz	39	<i>szortíroz, kopíroz, bázíroz ...</i>
	-ifikál	7	<i>szimplifikál, notifikál, pontifikál ...</i>
<b>ige + igeképző</b>	-(t)At	1412	<i>cumiztat, pattogtat, véset ...</i>
	-Ódik	851	<i>érlelődik, shiftelődik, szövődik ...</i>
	-(O)gAt	510	<i>fűzőget, bütykölget, koktélozgat ...</i>
	-(t)Atik	178	<i>hívatik, guríttatik, bontatik ...</i>
<b>hangzásséma</b>	-Og	318	<i>mekeg, szörcsög, bugyborog ...</i>
	-Ant	147	<i>cuppant, fröccsent, pislant ...</i>
	-An	141	<i>csosszan, surran, fortyan ...</i>
	-Ad	105	<i>csoffad, görnyed, suvad ...</i>
	-Aszt	83	<i>lottyaszt, sarjaszt, repeszt ...</i>
	-int	41	<i>pöccint, hörpint, köhint ...</i>