

# Infinitívuszi szerkezetek a magyarban: adatvezérelt megközelítés<sup>1</sup>

Kalivoda Ágnes<sup>1</sup>, Prószéky Gábor<sup>1,2</sup>

<sup>1</sup> HUN-REN Nyelvtudományi Kutatóközpont

<sup>2</sup> Pázmány Péter Katolikus Egyetem,  
Információs Technológiai és Bionikai Kar

This paper describes Hungarian infinitive constructions using a data-driven approach. It aims to study these constructions along formal-distributional features, utilizing authentic corpus data and frequency counts. Our findings derive from an open-source dataset extracted from a 776.9-million-word treebank of Hungarian. The resource contains more than 9 million instances of infinitive constructions, annotated for a wide range of linguistic features and metadata. We discuss the following topics: inflected infinitives, verb clusters consisting of multiple infinitives, auxiliary-like lexical items, separable preverbs, the attested word orders of infinitives, finite verbs and their respective preverbs, and finally, detailed distributions of preverbs in infinitive constructions. Our study reveals some trends which would have remained unseen without a quantitative approach.

**Keywords:** infinitive construction, data-driven approach, corpus analysis, formal-distributional features, preverb, word order, Hungarian grammar

**Kulcsszavak:** infinitívuszi szerkezet, adatvezérelt megközelítés, korpuszelemzés, formai és disztribúciós jegyek, igekötő, szórend, magyar nyelvtan

## 1. Bevezetés

Tanulmányunkban a magyar nyelv infinitívuszi szerkezeteivel foglalkozunk, ezeknek a korpuszalapú leírására törekszünk. Vizsgálatunk tárgyát képezik az (1) példában felsorolt, és az ezekhez hasonló esetek. Az infinitívuszt és az igekötőjét vastagon szedéssel, a segédigeszerű elemet és ennek az igekötőjét aláhúzással jelöljük.

---

<sup>1</sup> Szeretnénk köszönetet mondani a két névtelen bírálónak a cikkünk gondos átolvasásáért és a sok hasznos javaslatért. Kalivoda Ágnes kutatása az 142317 számú OTKA PD projekt keretében, az Innovációs és Technológiai Minisztérium Nemzeti Kutatási, Fejlesztési és Innovációs Alapból nyújtott támogatásával, a PD 22 pályázati program finanszírozásában valósult meg.

- (1) a. *sikerült megoldani a feladatot*  
 b. *el akarom végre mondani*  
 c. *képtelen volt rájönni*  
 d. *el sem hajlandó indulni*  
 e. *így felejtette aztán el bevinni*  
 f. *a leveleket alá szokás írni*  
 g. *el kell kezdeni úszni tanulni*

A hagyományos nyelvtanokban az infinitívusz főnévi igenévként ismert, mi azonban – Antal (1977: 22) és Kálmán C. et al. (1989) érvelésével egyetértve – nem tartjuk szerencsésnek ezt az elnevezést. Ha valamit főnévi igenévként hívnánk, az inkább az *-ás/-és* képzős alak lenne. Az 'infinitívusz' elnevezés arra utal, hogy az általunk vizsgált, *-ni* toldalékos alakok időjel nélküliek.

A 'segédige' szintén egy sokat vitatott fogalom a magyar nyelvelméletekben. A szakirodalom nem egységes abban a kérdésben, hogy indokolt-e segédigei kategóriát feltételezni a magyarban, és ha igen, akkor milyen szempontok mentén különíthetők el ennek a csoportnak a tagjai az egyéb igéktől. A szempontok között rendszerint felmerül az, hogy ezek az igék defektívek, azaz hiányosak. Ez a hiányosság lehet hangtani természetű (hangsúlytalanságra törekvés), alaktani (hiányos paradigma), mondattani (viszonylag kötött szórend) és szemantikai is (hiányos vonatkeret). Ezek a jegyek jellemzően kölcsönös összefüggésben állnak egymással. A magyar segédigék kérdésével foglalkozó szakirodalom alapos áttekintését adja Kenesei (2001) és Laczkó (2014). Kálmán C. et al. (1989) formai és disztribúciós szempontok mentén határozza meg a segédigék csoportját, az ő megközelítésükben nemcsak igék, hanem bizonyos névszók is bekerülnek ebbe a kategóriába. A mi tanulmányunknak nem célja, hogy a 'segédige' terminus körüli vitában állást foglaljon. 'Segédigeszerű elem' néven hivatkozunk minden olyan igére, igenévre és névszóra, amelynek van infinitívuszi dependense az általunk használt korpusz függőségi elemzésében.

A kutatásunk adatvezérelt, megfigyeléseink mögött egy több mint 9 millió infinitívuszi szerkezetet tartalmazó, korpuszból kinyert adattábla áll. A célunk az, hogy a vizsgált konstrukciók főbb formai és disztribúciós tulajdonságait a valós nyelvhasználat példáival és részletes gyakorisági adatokkal mutassuk be.

Elsőként a kutatás módszertani hátterét ismertetjük, majd sorra vesszük az infinitívuszi szerkezetek alkotóelemeit: magát az infinitívuszt, a hozzá kapcsolódó segédigeszerű elemeket, valamint az igekötőket, amelyek egyaránt kapcsolódhatnak az infinitívushoz és a finit igéhez is. Végül össze-

foglaljuk a legfőbb eredményeinket, és felvázoljuk a kutatás folytatásának néhány lehetséges irányát.

## 2. Adatok és módszerek

A kutatás során adatvezérelt, kvantitatív fókuszú megközelítést alkalmaztunk, amely a Magyar Nemzeti Szövegtár (Oravecz et al. 2014) vizsgálatából indul ki. Elsőként bemutatjuk a felhasznált korpuszt, majd az adatgyűjtési folyamat főbb lépéseit. Végül ismertetjük a munkánk első, gyakorlati eredményét: egy nyílt hozzáférésű adattáblát, amely több mint 9 millió infinitívuszi szerkezetet tartalmaz.

### 2.1. A felhasznált korpusz

A Magyar Nemzeti Szövegtár 2 (Oravecz et al. 2014) egy 1,04 milliárd szavas korpusz, amely a 20. század végének és a 21. század elejének magyar nyelvhasználatát reprezentálja. A kutatásunkban ennek egy új – megjelenés előtt álló – változatát használtuk fel, az MNSz2-UD korpuszt (Kalivoda et al. 2023). Ez a verzió számos korpusztisztítási lépésen esett át, amelyek közül a legfontosabbak a paragrafus-szintű duplumszűrés, a nem magyar nyelvű bekezdések automatikus azonosítása és törlése, valamint az 500 tokennél hosszabb mondatok kiszűrése. Az utóbbiak jellemzően nem valódi mondatok, hanem kódrészletek, vagy hibásan feldolgozott weboldalak. Ezáltal a szöveg minősége jobb, de a korpusz mérete kisebb az MNSz2-höz képest (776,9 millió szövegszó).

Az MNSz2-UD kapcsán fontos még kiemelni, hogy új elemzési szintekkel gazdagodott: az emMorph (Novák et al. 2016) által létrehozott, részletes morfológiai elemzéssel, valamint függőségi elemzéssel, amely a Universal Dependencies formalizmusát követi (de Marneffe et al. 2021). Ez utóbbi különösen fontos volt a jelen kutatás számára, mivel az infinitívushoz és a finit igéhez is tartozhatnak elvált igeötők, amelyek a függőségi relációk segítségével könnyen a megfelelő ígétőhöz kapcsolhatók. Ez az annotációs szint a legnagyobb segítség akkor is, amikor azt kell megállapítani, hogy az adott infinitívusz melyik finit igéhez, igenévhez vagy névszóhoz tartozik a mondatban. Azonban meg kell jegyeznünk, hogy az automatikus függőségi elemzés minősége eléggé változó.

Az MNSz2-UD korpusz elemzése a Stanzával (Qi et al. 2020) készült, amelynek a magyar tanítóanyaga a mindössze 1,2 millió szavas Szeged Dependency Treebank (Vincze et al. 2010). Elsősorban a kicsi tanítóanyag lehet az oka annak, hogy a Stanza a gyakori függőségi relációkat jól azonosítja (ilyen például az ígétől csak egy-két szóra eltávolodott ige-

kötő), ritkább jelenségek esetében azonban több hibát találunk. Ezek a hibák különösen az igekötős infinitívuszi szerkezetek szórendi mintáinál tűnnek elő (ld. az 5.2. fejezetet). Bár a tanulmányban közölt gyakorisági értékek hibás találatokat is magukban foglalnak – ami ekkora szöveg-mennyiség automatikus feldolgozása során elkerülhetetlen –, a bemutatott jelenségek nagyságrendje, egymáshoz viszonyított arányai megbízhatók.

## 2.2. Az adatgyűjtés folyamata

A korpuszból lekérdeztünk minden olyan mondatot, amely legalább egy infinitívuszt tartalmaz. Ezután kinyertük az egyes infinitívuszok összes olyan morfológiai és szintaktikai jellemzőjét, amelyet nélkülözhetetlenek tartottunk egy formai-disztribúciós leírás szempontjából. Ezek a következők: szám- és személyjelölés; lemma, tehát az inflexiók toldalékok nélküli szóalak; igekötő; vonzatok és bővítmények; az infinitívusszal együtt álló, segédigeszerű elem. Ami az igekötőket illeti, a prefixumként megjelenőket a részletes morfológiai elemzés segítségével tudtuk azonosítani, az elváltakat pedig a függőségi elemzés *compound:preverb* relációja alapján találtuk meg. Meghatároztuk azt is, hogy az elvált igekötő milyen távol áll az infinitívuszi ige-től, tokenszámban mérve. A vonzatokat és bővítményeket úgy állapítottuk meg, hogy az infinitívusz dependenseit kerestük, ezek lemmáját és esetragját nyertük ki. A névutókat is esetragként kezeltük. A segédigeszerű elemet pedig egy ellentétes irányú kereséssel lehetett megtalálni: azt vizsgálva, hogy minek a dependense maga az infinitívusz. Itt további, körültekintő adatfeldolgozásra volt szükség a létige különböző megvalósulásai esetében. Például a *fürödni voltam* szó-sorban a *van* lemma az, amit keresünk, de a *képes voltam*-ban a *képes*, ahogyan a *képes lettem volna* esetében is. A (2) példában összefoglalva láthatók az eddig bemutatott jellemzők, azokkal a változónevekkel, amelyeket később az adattáblában is használunk.

(2) *nekem is **oda kellett volna mennem** nyolcra*

→ inf_form:	mennem
→ inf_lemma:	oda+megy
→ inf_prev:	oda
→ inf_prevpos:	-3
→ inf_stem:	megy
→ inf_persnum:	1Sg
→ fin_lemma:	kell

- argframe\_cases: Dat Sub
- argframe\_long: én[Dat] nyolc[Sub]

A segédigeszerű elemnek hasonló jellemzőit hasonló módszerrel nyertük ki, mint amelyeket az infinitívusz esetében bemutatunk: szám- és személyjelölés; lemma; igekötő; valamint szófaj. Ez utóbbi az infinitívuszok vizsgálatakor nem volt érdekes, mert csak ige lehet, a segédigeszerű elemek azonban lehetnek igenevek és névszók is. A vonzatkeretet nem vizsgáltuk. A (3) példa illusztrálja a segédigeszerű elemek jellemzőit.

(3) *meg sem próbáltuk galériánkba feltenni*

- fin\_form: próbáltuk
- fin\_lemma: meg+próbál
- fin\_prev: meg
- fin\_prevpos: -2
- fin\_stem: próbál
- fin\_wordclass: VERB
- fin\_persnum: 1PI

Az eddig tárgyalt jellemzők ismeretében meg lehetett határozni az infinitívusz, a segédigeszerű elem és ezek igekötőinek sorrendjét is. A felsoroltak közül csak az infinitívusz az, amely minden korpusztalálathoz szerepel, minden más opcionális. A lehetséges sorrendek száma 41, de – ahogy később látni fogjuk – a korpuszban ennél lényegesen kevesebb szórendi variáció fordul elő. A (4) példa azt az esetet szemlélteti, amikor a finit igeének és az infinitívusznak is van igekötője.

(4) *próbálja meg mindenki kihozni magából a maximumot*

- order: fin pv PV INF

Az adatgyűjtés során megőriztük az infinitívuszi szerkezet mondatszintű kontextusát is, valamint a dokumentumszintű metaadatokat: a publikálás évét, a stílusregisztert és a fájlnévet. Ezeket mutatja be az (5) példa.

- (5) *Hátha csak be akarok férközni a bizalmába.*
- doc\_year: 2004
  - doc\_style: literature
  - doc\_id: 1.lit\_hu\_dia\_Hubay\_Miklos-Az\_en\_koromban.vert
  - left\_context: Hátha csak
  - kwic: be akarok férközni
  - right\_context: a bizalmába.

### 2.3. Az elkészült adattábla

Az adatgyűjtési folyamat eredményeként egy olyan adattáblát kaptunk, amely az infinitívuszi szerkezetek összes, MNSz2-UD korpuszban szereplő előfordulását tartalmazza. Ez összesen több mint 9 millió találatot jelent. Az adattábla formátuma TSV (tab separated values): egy egyszerű táblázat, amelyben az első sor a fejléc, a további sorok az egyes korpuszbeli előfordulásoknak felelnek meg, az oszlopok pedig a vizsgált jellemzőknek. Az adatbázis nyílt hozzáférésű, a <https://zenodo.org/record/8417366> linken érhető el.

## 3. Infinitívusz

Az adattábla segítségével vizsgálható nyelvi jelenségek bemutatását az infinitívusszal kezdjük. Elsőként egy morfológiai témával foglalkozunk, az infinitívusz ragozhatóságával, a ragozatlan, valamint a különféle szám- és személyragos alakok gyakoriságával. Ezután egy mondattani témára, az infinitívuszok halmozására térünk át.

### 3.1. Ragozott infinitívusz

A magyar infinitívusz egyik sajátossága az, hogy szám- és személyragot is felvehet. Ekkor a vele együtt álló segédigeszerű elem személytelen, csak az idő- vagy módjel jelenik meg rajta (vö. *lát-ni akar-om* ↔ *lát-n-om kell*). Ez a jelenség egyedi a finnugor nyelvcsaládon belül, de tágabb perspektívában nem számít ritkaságnak. Megtalálható például több újlatin nyelvben – köztük a portugál az, amelyben a legjobban dokumentált –, valamint a walesiben is (Jansegers 2021). A magyar ragozott infinitívusz legrészletesebb jellemzését Tóth (2000) adja.

**1. táblázat:** A ragozatlan és a ragozott infinitívuszok tokengyakorisága.

Személy	Szám	Rag	Gyakoriság
–	–	-ni	7 486 728
3	egyes	-nia/-nie	495 834
3	többes	-niuk/-niük	195 486
1	egyes	-nom/-nem/-nöm	184 675
1	többes	-nunk/-nünk	168 142
2	egyes	-nod/-ned/-nöd	37 740
2	többes	-notok/-netek/-nötök	4 017

Az 1. táblázat a különféle infinitívuszi végződések gyakorisági eloszlását mutatja. Látható, hogy a ragozatlan alak hétszer olyan gyakori, mint a szám- és személyragos alakok együttvéve. A ragozott alakok viszonylatában legmagasabb a harmadik, legritkább a második személy tokengyakorisága, ezeken belül pedig rendre az egyes számú alakok fordulnak elő többször.

### 3.2. Infinitívuszok halmozása

Az eddigi példánk többségében egy segédigeszerű elem és egy infinitívusz alkotta az igei komplexumot. Most egy olyan jelenség felé fordítjuk a figyelmünket, amellyel gyakran találkozni a szakirodalomban – ld. például az É. Kiss – Riemsdijk (2004) kötet számos tanulmányát –, ez az infinitívuszok halmozása egy igei komplexumon belül. A jelenséget a (6) példái szemléltetik.

- (6) a. *Nem tudunk enni adni a családjainknak.*  
 b. *Szeretném elvinni úszni tanítani.*

Elméleti szempontból hatalmas konglomerátumok hozhatók létre az infinitívuszok halmozásával, de a korpuszadatok tanúsága szerint egyáltalán nem szokás háromnál több infinitívuszt használni egy tagmondatban. Az ezzel kapcsolatos adataink összefoglalását a 2. táblázat tartalmazza.

Amint látjuk, már két infinitívusz együttes használata sem túl gyakori (a közel 9 millió esetből mindössze 53,5 ezerben adathozható). Ezek között találunk néhány olyat, amely két infinitívusz szoros, kollokáció-szerű kapcsolata. Ezekben a két elem sorrendje is meglehetősen kötött, pl. *veszni hagyni, élni hagyni, köpni-nyelni*. Ezen kívül jellemző még az, hogy az egyik infinitívusz egyfajta célt fejez ki, a másik pedig a célhoz való eljutást, pl. *dolgozni menni, enni adni, leülni tárgyalni*.

Három infinitívusz együttes megjelenése (összesen 65 esetben) meglehetősen ritka. A lánc első eleme alig mutat változatosságot: a *tudni, kezdeni, próbálni, lenni* (mint névszóhoz tartozó kopula) alakok bukkannak fel itt. Feltételezhető, hogy a beszélők inkább az alárendelt mellékmondatok, kötőmódú alakok és egyéb alternatív kifejezőmódok használatát választják az infinitívuszok halmozása helyett.

**2. táblázat:** A halmozott infinitívuszok tokengyakorisága. A 'Szórend' oszlopban az INF<sub>r</sub> személyragos infinitívuszt jelöl.

Halmozás	Szórend	Gyakoriság	Összesen
<b>2</b>	INF – INF	51 341	<b>53 443</b>
	INF <sub>r</sub> – INF	1 425	
	INF – INF <sub>r</sub>	561	
	INF <sub>r</sub> – INF <sub>r</sub>	116	
<b>3</b>	INF – INF – INF	59	<b>65</b>
	INF <sub>r</sub> – INF – INF	6	

Végül érdemes kitérnünk arra a kérdésre, hogy a halmozott infinitívuszi szerkezetekben milyen eloszlást mutatnak a személyragozott (INF<sub>r</sub>) és a személyrag nélküli (INF) alakok. A 2. táblázatban látható, hogy a személyrag leggyakrabban a lánc első elemén jelenik meg, ld. a (7a) és (7d) példákat. Ugyanakkor az sem kivételes eset, hogy az utolsó infinitívusz a személyragozott, ld. (7b), és ritkábban az is előfordul, hogy több személyragos infinitívusz áll egymás mellett, ld. (7c).



- (7) a. *Sikerült útszöknöm aludni végre kisszobánkba.*  
 b. *Nem ártana megtanulni olvasnod.*  
 c. *Önnek kell tudnia alkalmazkodnia.*  
 d. *Nem kell fáradoznom elmenni szavazni.*

Az utolsóként említett esetben minden személyrag azonos, és szembetűnő, hogy az adatok 93,2%-a beszélt nyelvi szöveg leiratából származik: rádiós beszélgetésből vagy parlamenti felszólalás jegyzőkönyvéből.

#### 4. Infinitívusszal álló segédigeszerű elemek

A nyelvi jelenségek vizsgálatát az infinitívusz mellett megjelenő, segédigeszerű elemekkel folytatjuk. Ahogy említettük, ezek szófajilag sokféle lehetnek: finit igék, igenevek, valamint névszók, amelyek kopulával együtt alkotnak összetett szerkezetet. Először arra teszünk kísérletet, hogy Kálmán C. et al. (1989) mintájára két szempont mentén, igetövűség és személyragozhatóság szerint csoportosítsuk az adattábla leggyakoribb segédigeszerű szavait. Azt találjuk, hogy a választott jellemzők alapján nem mindegyik lexikai elem sorolható be tisztán egy-egy kategóriába. A fejezet második felében ezeket a – Kálmán (2015) szóhasználatát követve – „fél-igéket és alig-igéket” mutatjuk be.

##### 4.1. Igetövűség és személyragozottság

A segédigeszerű elemeket elsőként aszerint csoportosítjuk, hogy igei vagy névszói tövűek-e. Igei tövűek a finit igék és a különféle igenevek, ld. a (8) példasort. Az adattáblában a *tessék* és *tessen* megszilárdult alakokat is ide soroltuk, mivel mindkettő a *tetszik* igére vezethető vissza.

- (8) a. *meg szeretném nézni*  
 b. *mukkanni sem merve*  
 c. *az úszni vágyók*

Névszói tövűek a melléknevek és főnevek, amelyek szintén gyakran állnak infinitívusszal – ld. (9) példasor –, és némelyikük az igei tövű elemekhez nagyon hasonló eloszlást mutat. Kálmán C. et al. (1989: 59) összetett igeként hivatkozik ezekre, abból a megfontolásból, hogy a szerkezet része a kopula is, amely a kijelentő mód jelen idő harmadik személy kivételével megjelenik a névszói főtag mellett. Az általunk vizsgált korpuszban a segédigeszerű elemek 11,2%-a névszói, 88,8%-a igei tövű.

- (9) a. *az ilyet előre meg szokás beszélni*  
 b. *belegondolni is szörnyűség*  
 c. *képtelen volnék végigcsinálni*

A másik csoportosítási mód a segédigeszerű elemeknek azon a jellemzőjén alapul, hogy személyragozhatók-e vagy sem. Ez a tulajdonság független a tőtípustól. Az igék általában személyragozhatók, de van néhány olyan is, amely – legalábbis az infinitívuszi vonzata mellett – csak egyes szám harmadik személyű alakban fordul elő, ld. (10) példa. Idő- és módjelet ezek is felvehetnek, a mellettük álló infinitívusz pedig személyragozható.

- (10) a. *megéri megkóstolni*  
 b. *megéri megkóstolnom*  
 c. *\*megérem megkóstolni*

Névszói tő esetében a személyrag a kopulán jelenhet meg, a számot pedig a kopula és a névszó egyaránt jelöli, ld. (11) példa.

- (11) a. *hajlandók voltak belátni*  
 b. *\*hajlandó voltak belátni*

Az igetövűség és a személyragozottság kapcsolatát Kálmán C. et al. (1989: 59) egy 2x2-es mátrixban jeleníti meg. Ezt a mátrixot megkíséreltük reprodukálni úgy, hogy az adattáblánkban legalább 5000 tokengyakorisággal rendelkező elemeknek megvizsgáltuk ezt a két jegyét. Az eredmény az 1. ábrán látható.

Míg Kálmán C. et al. (1989) minden vizsgált szót egyértelműen be tudott sorolni valamelyik kategóriába intuitív alapon, a korpuszadatok enélkül árnyaltabb képet mutatnak. Az 1. ábrán látható, hogy a *sikerül* infinitívusz mellett jellemzően személytelen az ige, a beszélők egy része mégis ragozza, a *kötelező* pedig szintén inkább személytelen, de előfordul személyragos kopulával is. Legérdekesebbek a *kell*, *muszáj* és *szabad*, amelyeknek nemcsak a személyragozhatósága, hanem az igei és névszói használata is ingadozó. Ezekkel foglalkozunk a következő fejezetben.

		Személyragozott?					
		igen		nem			
Igetövű?	igen	<i>akar</i>	<i>bír</i>	<i>elfelejt</i>	<i>elkezd</i>	<i>árt</i>	
		<i>elmege</i>	<i>enged</i>	<i>fog</i>	<i>hagy</i>	<i>illik</i>	
		<i>jár</i>	<i>jön</i>	<i>igyekszik</i>	<i>készül</i>	<i>megér</i>	
		<i>kezd</i>	<i>kíván</i>	<i>lát</i>	<i>látszik</i>	<i>sikerül</i>	
		<i>megy</i>	<i>mer</i>	<i>megpróbál</i>	<i>mond</i>		
		<i>próbál</i>	<i>segít</i>	<i>megtanul</i>	<i>szokott</i>		
		<i>tud</i>	<i>szeret</i>	<i>szeretne</i>	<i>tervez</i>		
		<i>tetszik</i>	<i>van</i>	<i>szándékozik</i>	<i>vél</i>	<i>kell</i>	
						<i>muszáj</i>	
						<i>szabad</i>	
	nem	<i>hajlandó</i>	<i>hivatott</i>			<i>célszerű</i>	<i>értelmes</i>
		<i>jó</i>	<i>kénytelen</i>			<i>fontos</i>	<i>könnyű</i>
		<i>képes</i>	<i>képtelen</i>			<i>lehetetlen</i>	<i>nehéz</i>
		<i>kész</i>	<i>köteles</i>			<i>szükséges</i>	<i>tilos</i>
						<i>kötelező</i>	

**1. ábra:** A leggyakoribb segédigeszerű elemek csoportosítása aszerint, hogy igitövűek, illetve személyragozottak-e. Öt lexikai elem körül szaggatott vonalat használtunk, ami azt jelenti, hogy az elem ingadozó viselkedést mutat.

#### 4.2. Féligék és alig-igék

A *kell* alapvetően igitövű, de a korpuszban több olyan használata is előfordul, amelyben inkább a névszókra hasonlít. Ekkor mindig *kellesz* alakban áll, amely nem egyes szám második személyű alak, hanem a *kell* + *lesz* összevonásából keletkezett, ld. a (12) példasort. A *kellesz* használata nyelvjárási jelenség, és az északkeleti régióban jellemző. Részletes leírását, valamint a kapcsolódó szakirodalmak alapos áttekintését adja Virovec (2022). A szerző arra a következtetésre jut, hogy a szerkezet használatának gyakoriságára és elfogadhatóságára számos tényező lehet hatással: például az evidencia típusa, a beszélő bizonyossága, a beszélő elkötelezettsége, a megnyilatkozás szándékolt és érzékelt udvariassága, ld. Virovec (2022: 22).

- (12) a. *Megnyugtattak: Nem kellesz **fizetnem!***  
 b. *Persze **bizonyítania** kellesz, amire biztosan meg is lesz a lehetősége.*  
 c. *De akárkinek a bőrébe kellesz **bújnod** legközelebb, [...]*  
 d. ***El** kellesz **engednie** így vagy úgy...*  
 e. *Asszem **meg** kellesz **emelnem** a percdíjat.*

A korpuszadatok azt mutatják, hogy a *kell* a személyragozhatóság tekintetében is ingadozást mutat. Bár az esetek döntő többségében személytelen, adatolható személyragos alakjai is, ld. a (13) példait. É. Kiss (2009: 214) azt írja, hogy az *el köllök menni* csak Magyarország északnyugati részén fordul elő. Bár az MNSz2 használja a régiót mint metaadatot, ez nem nyelvjárási területet jelöl, így ilyen kérdések vizsgálatára nem használható.

- (13) a. *Holnap **be** kellek **menni** a kórházba.*  
 b. *Még **el** kellünk **menni** a postára.*

A *szabad* szintén a névszók és igék közötti szürke zónában helyezkedik el. Kenesei (2001: 26) azt írja róla, hogy az újabb, igeként való használatában megfigyelhető alakok a *szabadott*, *szabadna* és *szabadjon*. Függetlenül attól, hogy az alany milyen számú és személyű, a *szabad* ragozott alakjaiban a harmadik személy az állandó, ld. (14a-b). Névszói használatkor nagyon ritkán előfordul, hogy számban egyezik az alannyal, azaz többes számú alakja is lehet, ld. (14c).

- (14) a. *Szabadna **tudnom** a nevüket?*  
 b. *Nem szabadott volna **kinyitnia**.*  
 c. *A dolgozók szabadok **elhagyni** a céget.*

A *muszáj* igei használatáról röviden említést tesz *A magyar nyelv értelmező kéziszótára* (Bárczi – Országh [főszerk.] 1959), ahol 'ritka' és 'népies' címkével látják el a szócikkírók. A korpuszadatok között a *muszájna* és *muszájjon* alakokra találunk példát a szépirodalmi regiszteren belül, ld. (15a-b). Névszói használata esetén a *muszáj* személytelen a köznyelvben – pl. *nekem muszáj megnézni* –, de a szépirodalmi és beszélt nyelvi regiszterekben akad néhány példa a *muszáj*-jal álló kopula személyragozására, ld. (15c-d). Ingadozás figyelhető meg abban is, hogy a *muszáj* számban egyezik-e a kopulával vagy sem.

- (15) a. *Nem muszájna hagyni.*  
 b. *És hogy azután szülni muszájjon?*  
 c. *Így muszáj voltunk váltogatni a futást a gyaloglással.*  
 d. *Muszájok vagyunk megmutatni, hogy [...]*

Kálmán (2015) nemcsak leírni, de magyarázni is igyekszik a *kell*, *szabad* és *muszáj* esetében megfigyelhető szófaji és morfológiai ingadozásokat. A *kell*-t „féligenek” tekinti amiatt, hogy hiányos a paradigmája, és valamilyen elvont viszonyra utal, nem pedig dinamikus eseményre. Jelentésében közel áll hozzá a *muszáj* és a *szükséges* is, amelyek rendszerint kopulával együtt alkotnak finit igés szerkezetet. Kálmán szerint a *muszáj lesz*, *szükséges lesz* mintájára alakulhatott ki a *kell lesz*, majd az összevont *kellesz*. Említi még a *muszájna*, *szabadjon*, *szabadott* alakokat is, amelyek épp ellentétes irányban változnak, a névszói használatuk mellett az igei is megjelenik.

Arra a kérdésre, hogy hogyan alakultak ki ezek az alakok, Kálmán a következő választ adja: „Jelen időben a személytelen igei állítmány és a névszóval alkotott (pl. *muszáj*, *szabad*) egyformán jelöletlenek, nem jelöli önálló nyelvi jel az egyes szám harmadik személyt – éppen ez ad módot arra, hogy legalább néha ugyanabba a kategóriába soroljuk a kétféle szótót. Múlt idejű alakjuk a személytelen igéknek is van (*kellett*, *lehetett*), viszont jövő idejű alakjuk a magyar igéknek általában nincs. Míg a *muszáj* és a *szabad* esetében a jövő idő kifejezésére rendszeres nyelvi eszköz áll rendelkezésre (*muszáj lesz*, *szabad lesz*), az igetövek esetében nincs a jövő idő kifejezésének rendszeres nyelvi eszköze. Ezért különösen vonzó lehetőség „féligenekként” kezelni, vagyis ideiglenesen névszótónek tekinteni őket, mert így meg lehet alkotni a kifejezetten jövő idejű *kellesz*, *lehet lesz* alakokat.”

Felmerül az a kérdés is, hogy miért éppen ez a három szó mutat ingadozást. Kálmán szerint a válasz a gyakoriságukban keresendő, és az általunk vizsgált korpuszadatok is erre utalnak. A *kell* a leggyakoribb személytelen ige, a *szabad* és a *muszáj* pedig a leggyakoribb, személytelen igés kifejezésben előforduló névszó, tehát azonos mondatkörnyezetekben fordulnak elő. Emellett szemantikailag is hasonló, nagyon általános modalitásokat kifejező szavak, ez pedig szorosan összefügg a kiugró gyakoriságukkal.

## 5. Igekötők és szórendi minták

A magyar nyelvben az igekötőzés a komplex predikátumok alkotásának az egyik, ha nem a leggyakoribb módja. Az igekötő elválhat az igétől,

amelyhez lexikailag tartozik. Ez egy 'finit ige + infinitívusz' szerkezet esetében azt jelenti, hogy akár két elvált igeikötővel, és ebből következően sokféle, összesen 41 lehetséges szórenddel számolhatunk.

Először bemutatjuk az igeikötő és az igező három alapvető szórendjét, ezután rátérünk azoknak a szórendi variációknak az ismertetésére, amelyek az infinitívuszi szerkezetekre jellemzők. Végül azt vizsgáljuk, hogy milyen távol kerülhet az igeikötő az infinitívusztól attól függően, hogy van-e beférkőző lexikai elem a szerkezetben, és ha van, akkor milyen szófajú.

### 5.1. A három alapvető szórend

A magyar nyelv igeikötői háromféle szórendi helyzetben állhatnak az ige-tőhöz viszonyítva. Leggyakrabban közvetlenül az ige előtt, prefixumként jelennek meg. Erre az elrendezésre egyenes szórendként utalunk a továbbiakban, ld. (16a) példa. Megszakított szórendnek nevezzük azt az esetet, amikor az igeikötő az ige előtt áll, de egy vagy több szó elválasztja egymástól a kettőt, ld. (16b). A harmadik lehetőség a fordított szórend, ekkor az igeikötő az ige után helyezkedik el, ld. (16c). Az igeikötős szerkezetek szórendi mintázatait részletesen tárgyalja számos tanulmány, többek között J. Soltész (1959), Ackerman – LeSourd (1997), É. Kiss (2021) és Kalivoda (2021).

- |      |                         |                     |
|------|-------------------------|---------------------|
| (16) | a. <i>kimegyek</i>      | <i>kimenni</i>      |
|      | b. <i>ki sem megyek</i> | <i>ki sem menni</i> |
|      | c. <i>nem megyek ki</i> | <i>nem menni ki</i> |

Ahogy a példasorban látható, a finit ige és az infinitívuszok esetében egyaránt lehetséges mindhárom szórend, de ezek arányai jelentősen eltérnek. A 3. táblázatban Kalivoda (2021: 157) MNSz2 korpuszon végzett vizsgálatának egy részletét emeljük ki, amely ezeket a százalékos arányokat mutatja.

Mindkét szerkezet típusban az a leggyakoribb jelenség, hogy az igeikötő nem válik el az ige-től. Míg megszakított szórend ritkán fordul elő a finit igei szerkezetben, az infinitívuszoknál ez meglehetősen általános. A fordított szórend esetében viszont ennek épp az ellenkezője látható: a finit igeiket a korpuszpéldák közel 40%-ában követi az igeikötő, de az infinitívuszokra nézve ez a szórend elenyészően ritka.

**3. táblázat:** Az igekötők három szórendi helyének aránya a finit igei és az infinitívuszi szerkezetek esetében, Kalivoda (2021: 157) alapján.

Szórend (%)	Finít ige	Infinitívusz
egyenes	58,79	78,10
megszakított	1,63	21,72
fordított	39,58	0,18

## 5.2. Az igekötős infinitívuszi szerkezetek szórendi mintái

A mondat típusa az egyik legfontosabb tényező, amely az igekötő és az igező sorrendjét befolyásolja. A tárgyalás szempontjából megkülönböztethető két fő mondatfajta, a semleges és a nem-semleges mondat megkülönböztetése.

Alberti (2006: 37–38) három nyelvészeti terület irányából közelíti meg a semleges mondat fogalmát. Szemantikai szempontból a semleges mondat „jelentése pusztán a lexikai jelentéstényezőkből áll össze, nincsen szerkezeti többletjelentés, vagy óvatosabban fogalmazva a szerkezetből adódó többletjelentés minimális”. Hangtanilag semleges intonáció jellemzi, vagyis a mondatban a tartalmas szavak az alapértelmezett kezdőszótagi hangsúlyukat viselik, a grammatikai funkciószavak hangsúlytalanok. Pragmatikai szempontból a semleges mondatban bizonyos szavak előfeltevést hordozó elemek, de a szórendi és intonációs sajátosságok nem hoznak további előfeltevéseket. A nem-semleges mondatban ezzel szemben a hangsúly eloszlása nem egyenletes, és speciális jelentéstöbbletet hordoz, amely szerkezetileg is megjelenhet (Alberti 2006: 54).

Az infinitívuszi szerkezetben a segédigeszerű elemeknek is meghatározó szerepe van a szórend alakulásában. Ezeknek inherens lexikai tulajdonsága az, hogy a mondatban hangsúlyos helyzetbe törekednek-e, vagy jellemzően inkább kerülnek azt. Hangsúlykérő igeik például az *imád* és az *utál* (vö. *imádom utazni*, *utálok befizetni a csekkeket*), hangsúlykerülő pl. a *látszik (megoldódni látszott)*, beférkőző, egyúttal hangsúlykerülő pl. az *akar* és a *fog (el akar jönni, meg fogja oldani)*. Az egyes ige típusok részletes leírását és ismérveit adja Komlósy (1989), majd részben erre a munkára építve Kálmán C. et al. (1989) és Demszky (2021).

**4. táblázat:** Az infinitívuszi szerkezetek szórendi variációi. A jelöléseink: **f<sub>in</sub>** = segédigeszerű elem, **p<sub>v</sub>** = a **f<sub>in</sub>** igekötője, **inf** = infinitívusz, **p<sub>v</sub>** = az **inf** igekötője. Nem vettük fel külön sorként az olyan megszakított szórendet, ahol az igekötő és az ige között csak klitikumszerű szavak állnak, valamint azt az esetet, amikor az igekötő prefixumként jelenik meg. Ezeket (p<sub>v</sub>)-vel jelöltük.

Szórend	Gyakoriság	Kontextusok
<b>f<sub>in</sub> (p<sub>v</sub>) inf</b>	4 294 622	(1) semleges mondat hangsúlykérő <b>f<sub>in</sub></b> -nel, pl. <i>imádom össze-<b>rakni</b></i> (2) nem-semleges mondat, ahol az <b>inf</b> nem fókusz vagy topik, és a <b>f<sub>in</sub></b> igekötő nélküli, pl. <i>nem <b>tudott</b> önállóan <b>megélni</b></i>
<b>p<sub>v</sub> f<sub>in</sub> inf</b>	1 944 966	semleges mondat beférkőző <b>f<sub>in</sub></b> -nel, pl. <i>el <b>szeretném</b> <b>mondani</b></i>
<b>(p<sub>v</sub>) inf (p<sub>v</sub>) f<sub>in</sub></b>	689 528	(1) semleges mondat hangsúlykerülő <b>f<sub>in</sub></b> -nel, pl. <i><b>megoldódn</b><b>i</b> <b>látszott</b></i> (2) nem-semleges mondat, ahol az <b>inf</b> fókusz vagy topik, pl. <i><b>megvenni</b> senki nem <b>fogja</b>, már <b>felállni</b> is <b>megtanult</b></i>
<b>(p<sub>v</sub>) inf</b>	434 465	(1) elliptált <b>f<sub>in</sub></b> , pl. <i>ezt egyelőre nem [...] <b>tudni</b></i> (2) eleve nincs <b>f<sub>in</sub></b> , pl. <i><b>lenni</b> vagy nem <b>lenni</b></i> (3) visszakérdés az <b>inf</b> -re, pl. – <i>Muszáj lesz átszámolni. – <b>Átszámolni?</b></i>
<b>f<sub>in</sub> p<sub>v</sub> (p<sub>v</sub>) inf</b>	49 027	nem-semleges mondat igekötős <b>f<sub>in</sub></b> -nel, pl. <i>ezt <b>próbáljuk meg le is fordítani</b></i>
<b>inf p<sub>v</sub></b>	2 169	ld. <b>(p<sub>v</sub>) inf</b> , de nem-semleges mondatban, pl. <i>Gyümölcséről <b>ismerni meg a fát</b></i> .
<b>(p<sub>v</sub>) inf f<sub>in</sub> p<sub>v</sub></b>	1 145	(1) igekötős <b>f<sub>in</sub></b> , <b>inf</b> -fel a fókuszban, pl. <i>Legalább <b>esni tanulj meg!</b></i> (2) az <b>inf</b> topik, az igekötős <b>f<sub>in</sub></b> előtt fókusz áll, pl. <i><b>Lemosni</b> is csak közösen <b>moshatjuk le</b> magunkról.</i>
<b>(p<sub>v</sub>) f<sub>in</sub> inf p<sub>v</sub></b>	433	az <b>inf</b> -nek saját fókusza van, pl. <i><b>igyekszik pontosan adni át</b> az üzenetet</i>
<b>inf f<sub>in</sub> p<sub>v</sub></b>	94	progresszív, azaz éppen zajló esemény, pl. <i><b>söpörni kezdtem le</b> magamról a morzsákat</i>



A 4. táblázat összefoglalja a korpuszban található infinitívuszi szerkezeteknek azokat a szórendi mintáit, amelyek legalább ötször előfordultak,<sup>2</sup> kiegészítve a legjellemzőbb kontextusok felsorolásával. Amint látjuk, egy szórendi minta több kontextusnak is megfelelhet, a ténylegesen előforduló szórendek száma azonban nem közelíti meg azt a nagyságrendet, ami elvileg lehetséges lenne.

### 5.3. Az igekötő és az igető távolsága infinitívuszi szerkezetekben

Végül egy olyan kérdésre keressük a választ, amelyet csak kvantitatív alapon lehet vizsgálni: milyen távol kerülhet az igekötő az infinitívusztól, amelyhez lexikailag tartozik? Az eltávolodást elsősorban a különféle beférkőző, segédigeszerű elemek okozzák, és jelentősen különböző eloszlási mintákat találunk attól függően, hogy ezek az elemek milyen szófajúak. Ez olvasható ki a 5. táblázatból.

**5. táblázat:** Az infinitívushoz tartozó igekötők tokengyakorisága és pozícióinak számszerű jellemzői aszerint, hogy van-e beférkőző segédigeszerű elem a szerkezetben, és ha van, milyen szófajú.

beférkőző elem	gyakoriság	pozíciók adatai			
		átlag	szórás	minimum	maximum
–	2 526 235	0,00	0,09	-3	2
finit ige	958 969	-2,15	0,46	-11	2
névszó	4 564	-2,02	0,14	-4	-2
határozói igenév	178	-2,00	0,10	-3	-2
névszó + kopula	18	-4,00	1,08	-7	-3

<sup>2</sup> Ahogy a 2.1 fejezetben jeleztük, a korpusz függőségi elemzése a ritkább nyelvi jelenségek esetében sokszor hibás volt. Ezért úgy döntöttünk, hogy az 1000 előfordulás alatti mintákat átnézzük és szelektáljuk. A legtöbb, 1000-nél ritkább példával rendelkező mintáról kiderült, hogy hibás, ezeket a táblázatban nem tüntettük fel. Emellett sok szöveghibára is fény derült. Tipikusak voltak az olyan tévesztések, hogy az infinitívusz igekötőjét a finit igével írta egybe a mondat szerzője (pl. *elpróbált menni az el próbált menni* helyett). Ezek javítása túl összetett és időigényes lett volna, ezért nem kezeltük őket, a gyakorisági adatokat némileg torzíthatják.

A leggyakoribb, több mint 2,5 millió példával adathozható szerkezet típus az, amelyikben nem férkőzik be segédigeszerű elem az infinitívusz és a hozzá tartozó igekötő közé. Ekkor az igekötő szinte mindig prefixált. Nagyon ritkán fordul elő megszakított szórend, amelyben csak néhány klitikumszerű elem (*nem, sem, se, is*) áll az igekötő és az ítéto között. A hátravetett igekötős esetek még ennél is ritkábbak: akkor fordulnak elő, ha az infinitívusz tagadott vagy saját fókusza van, segédigeszerű elem pedig jellemzően nem jelenik meg a tagmondatban. A (17) példásor illusztrál minden pozíciót.

- (17) a. *Jobb lett volna **meg** sem is **születnie**.*  
 b. *Ezért **szeretnének** a vételárból valamennyit **vissza is kapni**.*  
 c. *Nem **kell** mindent feltétel nélkül **elhinni!***  
 d. *Nem **látni be** egyik ablakon sem, nem látni semmit.*  
 e. *Nem **venni** rajta észre, hogy műfordítás.*

A beférkőző finit ige adatait – közel egymillió előfordulást – vizsgálva az igekötőknek egészen más eloszlása figyelhető meg. A leggyakoribb a -2 pozíció, tehát az az eset, amikor csak a finit ige férkőzik be az igekötő és az infinitívusz közé. Megszakított szórend esetén az igekötő messzire eltávolodhat az infinitívusztól, de mindig közel marad a finit igehez, az ahhoz tartozó igemódosítói pozíciót foglalja el. Közte és a finit ige között csak klitikumszerű szavak állnak (*is, se, sem, ne, nem*). Ezt a jelenséget már többen leírták, és igyekeztek magyarázattal szolgálni rá: legelőször Kálmán C. et al. (1989), akik rámutatnak, hogy bár az igemódosító az infinitívusszal alkot szóközi egységet, szórendi és prozódiai szempontból szorosabban kapcsolódik a segédigeszerű elemhez. A korpuszvizsgálatunk alátámasztja ezt a megfigyelést, és kiegészíti azzal, hogy ha az 'igekötő + finit ige' egység elhangzott, az infinitívuszt akár az összes bővítmény is megelőzheti a további mondatrészben – mivel a magyarban a finit ige követő tartományban a szórend kötetlen (É. Kiss 2007). Hátravetett szórend esetén az igekötő mozgásteret jóval korlátozottabb, legfeljebb két szónyi távolságra kerül az infinitívusztól. Segédigeszerű elem ekkor is be tud férkőzni a szerkezetbe, de ez szinte kizárólag a *kezd* igeire jellemző. Ez a szórend mindig nem-semleges mondat típusban, azon belül pedig progresszív mondatokban jelenik meg, tehát amelyek éppen zajló eseményt jelölnek. A korpuszban adathozható igekötő-pozíciókat a (18) példa szemlélteti.

- (18) a. *Így annál jobban **be lehet** ezt a vezetés nélkül maradt tömeget később egy világállamba **kényszeríteni**...*  
 b. [...] ***el tudnánk** magunknak a saját szánk íze szerinti tökéletes édességet **készíteni**.*  
 c. [...] ***fel kell** tudniuk az olívaolajtermelésre szánt támogatás egy részét **használni** [...]*  
 d. [...], *természetesen **hozzá lehet** majd az ön ügyrendi indítványához is **szólni**.*  
 e. [...], *ha **össze lehetne** őket többé-kevésbé objektív módon is **mérni**, [...]*  
 f. ***Át kellett** ezt a régi épületet **építeni**.*  
 g. *Hát **fel lehet** ezt józan ésszel **fogni**?*  
 h. [...], *és **elő kéne** már otthon **fizetnem**, [...]*  
 i. *Allandóan **meg kell** önöket **szakítanom**.*  
 j. ***Tönkre akarja tenni** a bútoraimat?*  
 k. ***Tépni kezdi** össze a nagy énekeskönyvet.*

Szembetűnő, hogy a beférkőző finit igék közel egymillió előfordulásával szemben a beférkőző névszókra alig több mint 4500 példát lehetett találni. Ezen a gyakoriságon mindössze néhány névszótípus osztozik, amelyek a következők: *szokás, képes(ek), szükséges, hajlandó(ak), kénytelen(ek), muszáj, érdemes, célszerű*. Hátravetett szórend ebben a szerkezetben nem fordul elő. Az infinitívusztól elvált igekötő minden esetben a névszó elé kerül, legfeljebb klitikumszerű szócskák állnak közöttük, ld. a (19) példásort.

- (19) a. *Akkor **fel sem érdemes** többet **jönni**.*  
 b. *Másutt a fordító nevét **ki sem szokás írni**.*  
 c. *Ha valaki **el képes jutni** ebbe az állapotba, [...]*

Elenyészően kevés (18 darab) példát találtunk arra, hogy a beférkőző névszó mellett a kopula is megjelenik. A *se* vagy *sem* szócskák szinte minden esetben feltűnnek a névszó vagy a kopula előtt, ld. a (20) példáit. A szerkezet minden előfordulása beszélt nyelvi szövegek leiratából vagy szépirodalmi szövegből származik.

- (20) a. [...], *hogy **meg sem szabad lett volna** az egészet **próbálni**.*  
 b. ***El sem szabad lett volna** idáig **jutni**!*  
 c. [...] ***meg sem szabad lett volna próbálnunk**.*  
 d. *Még **meg se volt időm kínálni** az úrfit!*  
 e. *Érzelemmentesen is **el vagy képes olvasni** a jelentést.*

Végül előfordul olyan eset is (összesen 178 tagmondatban), hogy a beférkőző segédigeszerű elem határozói igenév, ld. a (21) példait.

- (21) a. [...], **meg sem akarva hallani** azt.  
 b. [...], **rá próbálva ruházni a felelősséget** [...]

Megfigyelhetjük, hogy ekkor az igekötő a határozói igenév előtti pozíciót foglalja el, de az infinitívusztól összességében nem kerül olyan messze, mint finit igék beférkőzésekor. Ilyen szempontból a határozói igenév a névszókkal mutat párhuzamot, és további hasonlóság az is, hogy ezek egyike sem időjeles.

## 6. Összefoglalás

Az MNSz2-UD korpuszból létrehoztunk egy olyan, nyílt hozzáférésű adattáblát, amely több mint 9 millió infinitívuszi szerkezetet tartalmaz. Ez alapján bemutattuk a vizsgált konstrukciók főbb formai és disztribúciós tulajdonságait.

Az infinitívusz kapcsán először azzal foglalkoztunk, hogy milyen gyakoriak a ragozatlan, valamint a különféle szám- és személyragos alakok. Ezután áttértünk a több infinitívuszt tartalmazó igei komplexumokra. Kimutattuk, hogy nem szokás háromnál több infinitívuszt használni egy tagmondaton belül.

A következő nagyobb témánk a segédigeszerű elemek voltak. Kísérletet tettünk arra, hogy a leggyakoribbakat Kálmán C. et al. (1989) mintájára két szempont mentén csoportosítsuk, igetövűség és személyragozhatóság szerint. Bizonyos elemeket nem lehetett egyértelműen kategorizálni a korpuszadatok alapján. A *kell*, *muszáj* és *szabad* esetében azt találtuk, hogy nemcsak a személyragozhatóságuk, hanem az igei és névszói használatuk is ingadozó.

Végül az igekötős infinitívuszi szerkezetek szórendi mintáit vizsgáltuk. Elvileg 41 szórendi variáció lehetséges, de ennek csak tötrésze adathozható. Foglalkoztunk azzal is, hogy milyen távol kerülhet az igekötő az infinitívusztól. Ehhez a méréshez a beférkőző elemek négy típusát különítettük el: a finit igét, a névszót önmagában vagy kopulával együtt, valamint a határozói igenevet. Lényegesen eltérő eloszlási mintákat találtunk attól függően, hogy ezek az elemek megjelennek-e, és ha igen, milyen szófajúak.

A munkánk során létrehozott adattáblát hasznosítani lehet további, akár gépi tanulási módszereket alkalmazó kutatásokban is. Ez a tanulmány formai és disztribúciós jellemzőket vizsgált, de az adattáblában az infinitívuszos vonzataival és bővítményeivel kapcsolatos információk is elér-

hetők. Érdekes lehetne akár mondattani, akár jelentéstani szempontból alaposabban megvizsgálni ezeket. Egy másik érdekes téma a segédige-szerű elemek vizsgálata. Ebben a cikkben végig így hivatkoztunk rájuk, de tovább lehet árnyalni a képet. Egy publikálás előtt álló tanulmányunk-ban (Kalivoda – Prószéky 2023) olyan kísérletet mutatunk be, amely for-mai és disztribúciós jegyek alapján, klaszterelemzés segítségével rajzolja ki ezeknek a lexikai elemeknek négy csoportját, amelyből kettő fedt le a segédigék Kálmán C. et al. (1989) szerinti osztályát. Az adattábla termé-szetesen úgy is használható, hogy kisebb mintákat veszünk belőle. Ez az egyes nyelvi jelenségek célzott, kvalitatív elemzésére nyújt lehetőséget.

### Irodalom

- Ackerman, Farrell – LeSourd, Philip (1997), Toward a lexical representation of phrasal predicates. In: Alsina, Alex – Bresnan, Joan – Sells, Peter (eds), *Complex predicates*. CSLI Publications, Stanford. 67–106.
- Alberti Gábor (2006), *A magyar mondattan elmélete és gyakorlata generatív megközelítésben III. A háttérelmélet*. Pécsi Tudományegyetem, Pécs.
- Antal László (1977), *Egy új magyar nyelvtan felé*. Magvető Kiadó, Budapest.
- Bárczi Géza – Országh László (főszerk.) (1959), *A magyar nyelv értelmező szótára I–VII*. Akadémiai Kiadó, Budapest.
- de Marneffe, Marie-Catherine – Manning, Christopher D. – Nivre, Joakim – Zeman, Daniel (2021), Universal dependencies. *Computational Linguistics* 47/2: 255–308. [https://doi.org/10.1162/coli\\_a\\_00402](https://doi.org/10.1162/coli_a_00402)
- Demszky, Dorottya (2021), The role of verb semantics in Hungarian Verb–object order. *Proceedings of the Linguistic Society of America* 6/1: 54–68. <https://doi.org/10.3765/plsa.v6i1.4941>
- Jansegers, Marlies (2021), *Inflected infinitives*. Oxford Bibliographies Online, Linguistics. <https://doi.org/10.1093/obo/9780199772810-0216>
- Kalivoda Ágnes (2021), *Igékötös szerkezetek a magyarban*. PPKE BTK, Budapest. Doktori (PhD) értekezés.
- Kalivoda, Ágnes – Malouf, Robert – Ackerman, Farrell – Sass, Bálint (2023), *Building a dependency treebank from the Hungarian Gigaword Corpus*. Kézirat.
- Kalivoda, Ágnes – Prószéky, Gábor (2023), *Hungarian auxiliaries revisited*. *Acta Linguistica Academica* 70/4. Megjelenés alatt.
- Kálmán C. György – Kálmán László – Nádasdy Ádám – Prószéky Gábor (1989), *A magyar segédigék rendszere*. In: Telegdi Zsigmond – Kiefer Ferenc (szerk), *Altalános Nyelvészeti Tanulmányok 17. Tanulmányok a magyar mondattan köréből*. Akadémiai Kiadó, Budapest. 49–103.
- Kálmán László (2015), *Féligék és alig-igék*. *Nyelv és Tudomány, tudományos ismeretterjesztő portál*. <https://m.nyest.hu/hirek/felikek-es-alig-igek> [letöltés: 2023. június 15.]

- Kenesei, István (2001), Criteria for auxiliaries in Hungarian. In: Kenesei, István (ed.), *Argument structure in Hungarian*. Akadémiai Kiadó, Budapest. 73–106.
- É. Kiss Katalin (2007), Az ige utáni szabad szórend magyarázata. *Nyelvtudományi Közlemények* 104: 124–152.
- É. Kiss Katalin (2009), *Nekem el kell menni/el kell mennem/el kell, hogy menjek/el kell menjek/el kellek menni*. In: É. Kiss Katalin – Hegedűs Attila (szerk), *Nyelvelmélet és dialektológia*. PPKE BTK, Piliscsaba. 213–227.
- É. Kiss, Katalin (2021), Predicative PPs. In: É. Kiss, Katalin – Hegedűs, Veronika (eds), *Syntax of Hungarian: postpositions and postpositional phrases*. Amsterdam University Press, Amsterdam. 251–284.
- É. Kiss, Katalin – Riemsdijk, Henk (eds) (2004), *Verb clusters: A study of Hungarian, German and Dutch*. John Benjamins Publishing Company, Amsterdam.
- Komlósy András (1989), Fókuszban az igék. In: Telegdi Zsigmond – Kiefer Ferenc (szerk.), *Általános Nyelvészeti Tanulmányok 17. Tanulmányok a magyar mondattan köréből*. Akadémiai Kiadó, Budapest. 171–182.
- Laczkó, Tibor (2014), On verbs, auxiliaries and Hungarian sentence structure in LFG. *Argumentum* 10: 421–438.
- 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 (eds), *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC 2016)*. European Language Resources Association (ELRA), Portorož, Slovenia. 1315–1322.
- Oravecz, Csaba – Váradí, Tamás – Sass, Bálint (2014), The Hungarian Gigaword Corpus. In: Calzolari, Nicoletta – Choukri, Khalid – Declerck, Thierry – Loftsson, Hrafn – Maegaard, Bente – Mariani, Joseph – Moreno, Asunción – Odijk, Jan – Piperidis, Stelios (eds), *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC 2014)*. European Language Resources Association (ELRA), Reykjavik, Iceland. 1719–1723.
- Qi, Peng – Zhang, Yuhao – Zhang, Yuhui – Bolton, Jason – Manning, Christopher D. (2020), Stanza: A Python natural language processing toolkit for many human languages. In: Celikyilmaz, Asli – Wen, Tsung-Hsien (eds), *Proceedings of the 58th Annual Meeting of the ACL: System Demonstrations (ACL 58)*. ACL, online. 101–108.
- J. Soltész Katalin (1959), *Az ősi magyar igekötők: meg, el, ki, be, fel, le*. Akadémiai Kiadó, Budapest.
- Tóth, Ildikó (2000), *Inflected infinitives in Hungarian*. TILDIL Dissertations Series, Tilburg.

- 
- Vincze, Veronika – Szauter, Dóra – Almási, Attila – Móra, György – Alexin, Zoltán – Csirik, János (2010), Hungarian Dependency Treebank. In: Calzolari, Nicoletta – Choukri, Khalid – Maegaard, Bente – Mariani, Joseph – Odijk, Jan – Piperidis, Stelios – Rosner, Mike – Tapias, Daniel (eds), Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC 2010). European Language Resources Association (ELRA), Valletta, Malta. 1855–1862.
- Virovec Viktória (2022), Mi is az a *kellesz*? Egy empirikus kutatás eredményei. Argumentum 18: 236–259. <https://doi.org/10.34103/ARGUMENTUM/2022/13>

