

Hány VAN nincs? A létige – zéró váltakozás korpuszvezérelt vizsgálata

Dömötör Andrea

PPKE BTK Nyelvtudományi Doktori Iskola
domotor.andrea@itk.ppke.hu

Kivonat: A kutatás korpuszvezérelt módszerrel vizsgálja, hogy a hagyományos értelemben vett zéró kopulán kívül milyen szintaktikai és szemantikai feltételekhez kötött a létige elhagyhatósága. A korpuszadatok szerint a létige lokatív mondatokban is lehet zéró alakú, különösen ha mutató névmás mellett áll. Ez arra utal, hogy a mutató névmásoknak – az angolhoz hasonlóan – közük van a létezés kifejezéséhez. Tagadó jelentésű egzisztenciális és birtokos mondatokból a *nincs* is elhagyható a tagadó névmások (*semmi, sehol*) mellől. A kutatás ezen tapasztalatai alapján a függőségi nyelvtanon alapuló automatikus mondatelemzők modelljének építéséhez a funkciófej-megközelítés (Simkó–Vincze 2016) tűnik jól alkalmazhatónak, a létige tartalmassága ugyanis nem feltétlenül meghatározó tényező az elhagyhatóság kérdésében. Az elemző eszközöket a formális nyelvléírásból ismert tipikus kopulákon kívül más lehetséges zéró alakú létigék jelenlétére is fel kell készíteni.

1 Bevezetés

A zéró kopula jelensége számos nyelvben ismert, de nincs univerzális elv arra, hogy a kopula mikor lehet testetlen. A magyarban az úgynevezett Third Person Parameter (Stassen 1994) érvényesül, azaz a jelen idő harmadik személy esetén nem jelenik meg testesen a létige. A kopula fogalmának értelmezésében azonban eltér a magyar és az angolszász szakirodalom. A magyar leíró nyelvészet hagyományosan csak a nominatívusi névszói predikátum esetén beszél kopuláról (Kádár 2011), ezzel szemben a nemzetközi, például az angol (pl. Heggie 1988) vagy a spanyol (pl. Maienborn 2005) irodalom ide sorolja a létige + PP szerkezeteket is.

Megfigyelhető, hogy a létige-zéró váltakozás valójában a magyarban sem korlátozódik csupán a nominatívusi névszói predikátumokra. Az ige hiánya lokatív állítmányú mondatokra is jellemző, például: *Hol a könyved?, Kulcs a lábtörő alatt.*

A kutatás azt vizsgálja, hogy a hagyományos értelemben vett zéró kopulán kívül milyen szintaktikai és szemantikai feltételekhez kötött a létige elhagyhatósága. A vizsgált létigetípusokat az 1. táblázat tartalmazza.

Típus	Példa
kopula	<i>Az előadás érdekes volt.</i>
lokatív	<i>A macska az udvaron van.</i>
egzisztenciális	<i>Van egy macska az udvaron.</i>
birtokos	<i>A szomszédnak zöld kertje van.</i>

1. táblázat. A dolgozatban vizsgált létigetípusok

Az előfeltevés szerint a lokatív argumentumot tartalmazó létigés mondatok viselkedése a névszói predikátumosokhoz hasonló lesz, azaz a Third Person Parameter érvényesül, és csak akkor hagyható el a létige, ha az alany specifikus (azaz a mondat nem egzisztenciális jelentésű).

2 Motiváció

A kutatás eredményeit egy függőségi nyelvtanon alapuló automatikus mondatelemző modelljének tervezéséhez szeretnénk felhasználni. A tervezett elemző (Prószéky–Indig 2015) pszicholingvisztikai indíttatású, azaz a gépi szövegfeldolgozást az emberi értéshez próbálja hasonlóná tenni. Működés során az elemző szigorúan balról jobbra halad, a szöveg beolvasása szavanként (szóközönként vagy írásjelenként) történik. Az éppen elemzendő egységet az elemző különböző szálai (morfológiai elemző, különböző grammatikai jelenségeket azonosító szájak, korpuszgyakorisági szájak, anaforafeloldó szájak, fókuszazonosító szál stb.) egy időben, párhuzamosan vizsgálják, így egymás eredményeit is javíthatják, pontosíthatják. Minden szó beolvasása után első feldolgozási szálként elindul a morfológiai elemzés, amely létrehozza a szó lemmáját, és megadja a további elemzéshez szükséges morfoszintaktikai jegyeket. Ezek lehetnek kereslet típusúak, azaz meghatározhatják, hogy milyen jegyekre tartanak igényt, vagy kínálat formájában kielégíthetik más szájak keresletigényét. Egy finit ige például a vonzatkeret-szótárból indít el esetvégződésre vonatkozó keresleti szájakat, amelyekkel aztán a megfelelő esetragos főnevek (NP-k) kapcsolódni tudnak. A morfológiai és szintaktikai relációkat leíró szájak mellett párhuzamosan egyéb információk (statisztikai információk, korpuszgyakoriságok, ontológiák, világismeret) is részt vesznek az elemzésben. Az elemzés kimenete a függőségi nyelvtanokra emlékeztető irányítottgráf-reprezentáció, amely a megnyilatkozás által jelölt nyelvi szituáció minden fontos elemét igyekszik azonosítani.

A számítógépes nyelvészek többféle módon próbálták a *van* létigét függőségi nyelvtani keretben kezelni. Simkó és Vincze (2016) három lehetséges elmélet eredményeit hasonlítja össze a Szeged Korpusz adatain. A funkciófej-elemzés a mondatban a funkciószavakat tekinti fejnek. Ennek megfelelően ha a mondatban nincs ragozott ige, akkor egy üres igealakot kell beszúrni a szerkezetbe. Ezzel szemben a tartalmasfej-megközelítés a jelentéses egységeket tekinti a mondat fő elemeinek. Ebben az értelmezésben csak a „tartalmas” létige lesz fej, kopula esetén a névszói predikátum a mondat feje, és a létige ehhez kapcsolódik (ha testes). A harmadik lehetséges módszer a komplex címkés elemzés, amely lényegében megegyezik a funkciófej-elemzéssel, csak a zéró kopula esetén tér el attól. Ha a

kopula nem jelenik meg testesen a felszíni szerkezetben, akkor a mondat feje a névszói predikátum lesz, és a hiányzó ige ennek címkéjén és a belőle kiinduló éleken lesz jelölve.

Simkó és Vincze (2016) kísérleteiben a tartalmasfe-elemzés érte el a legjobb eredményt, azaz a gépi tanulásra alapozott automatikus elemző ezzel a címkézési módszerrel tudta a legnagyobb pontossággal felismerni a zéró kopulát. Egyúttal a szerzők hangsúlyozzák az elméleti keret átgondolásának jelentőségét. Jelen esetben fontos kérdés, hogy mely létigéket szeretnénk tartalmas egységnek tekinteni, és hogy pontosan milyen esetekben kell felkészülni a zéró létigék meg(nem)jelenésére.

3 Elmélet és terminológia

Amint a bevezetőben is említettem, a magyar formális nyelveírás szerint „(az elliptikus szerkezeteket leszámítva) csak az azonosító és predikatív mondatok VAN-ja lehet zéró (hangalakú), azaz csak ezek alkotnak nominális mondat szerkezetet” (Kádár 2011: 7) Ugyanakkor szintén a bevezetőben láthattunk olyan lokatív példákat is, amelyek szerint ez nem igaz. Úgy tűnik, a létige más esetekben is lehet opcionálisan zéró hangalakú. Mivel a kutatás korpuszvezérelt, nem állítok előzetes hipotéziseket arra vonatkozóan, hogy melyek ezek az esetek.

A kérdés vizsgálata azonban felvet egy taxonómiai problémát, nevezetesen azt a sokat vitatott kérdést, hogy hány létige van a magyarban, vagy pontosítva: hányféle funkcióval bírhat a *van* szótő. Jelen dolgozat Clark (1978) alapján (részletesen I. 5. fejezet) az 1. táblázatban is ismertetett négy fő típust különbözteti meg. Kopula alatt a nominális mondat segédigéjét értem. Az ezen kívüli funkcióban használt létigéket a lokatív, egzisztenciális és birtokos csoportokba osztom fel. Fontos megjegyezni, hogy egzisztenciális mondatnak is lehet helyhatározói bővítője, lokatív mondatról azonban csak specifikus alany esetén beszélünk.

További terminológiai megjegyzés még, hogy létige, illetve VAN alatt a *van*, *nincs(en)*, *sincs(en)* szótövek bármely alakját érthetjük. Kádár (2011: 7) alapján nem tipikus zéró létigének kell tehát tekinteni minden olyan esetet, ahol a VAN nem kopula, de a felszíni szerkezetben nem jelenik meg hangzó alakban. Erre fókuszálunk ebben a tanulmányban.

4 Anyag és módszer

A kutatás módszere korpuszvezérelt, azaz nagy mennyiségű szöveges adat vizsgálatán alapszik, nincsenek kiinduló hipotézisei, minden következtetést az adatokból próbál levonni. A felhasznált adatbázis az MNSz. (Oravecz–Váradi–Sass 2014) korpusz volt, amely különböző műfajú, morfológiailag annotált szövegeket tartalmaz. Jelenlegi, 2.0.4-es verziója körülbelül 1,04 milliárd szövegszóból áll. Az adatbázishoz tartozó keresőfelületen lehetőség van szóalak, szótő és szófajcímke szerint is keresni, valamint a kapott találatokra további szűrőket is be lehet állítani. A korpuszon alapuló kutatások nagy nehézsége, hogy a morfológiai címkék gépi eszközzel készültek, így sok esetben hibásak, nem megbízhatóak (például a *vagy* létigét sokszor kötőszónak, sőt: az *által* igét főnévnek címkézi), így az adatgyűjtést sajnos nem lehetett teljesen automatizálni.

Első lépésben olyan mondatokat kerestem, amelyekben nincs finit ige. Ez a korpusz keresőfelületén reguláris kifejezéssel megoldható ugyan, ám a kapott eredmények alig voltak használhatók, a találatok nagy része ugyanis nem valódi szöveget tartalmazott, hanem HTML-fájlokból származott. Ezért inkább egy 100 000 mondatos random adatból Python-script segítségével szűrtem ki azokat a ragozott ige nélküli mondatokat, amelyekben van nominativus. Így 17 040 találatot kaptam, azonban korántsem minden esetben zéró létige okozta az ige hiányát. A keresést sokszor címek, az egyes műfajokra (pl. országgyűlési naplók, színművek rendezési instrukciói) jellemző nominális stílus, ellipszis, hiányos mondatok vagy éppen hibás címkék zavarták. Emiatt elkerülhetetlen volt a találatok kézi ellenőrzése, és így jelenleg 1000 mondatnyi megbízható adat áll rendelkezésre.

4.1 Kiinduló eredmények

A kiinduló adatok mindegyikében zéró hangalakú létige található, melyek közül 154 (15,4%) nem hagyományos értelemben vett kopula. Bár a minta nem nagy, annyi azért jól látszik, hogy a jelenség semmiképpen sem mondható marginálisnak. Ezenkívül a véletlenszerűen kapott 154 adat arra is jó volt, hogy irányt mutasson a további keresésekhez.

A kiinduló eredmények mondatfajta szerinti megoszlását a 2. táblázat tartalmazza.

Típus	Altípus	Gyakoriság
Lokatív	itt/ott/hol	27
	egyéb	28
	irodalmi	63
Egzisztenciális	semmi/sehol/soha	15
	egyéb	3
Birtokos	semmi/mi	8
	egyéb (vége)	5
Egyéb		5

2. táblázat. A kiinduló korpuszadatok

Ezekből az látható, hogy a nem tipikus zéró létigék az irodalmi szövegek leírásaiban a legjellemzőbbek (pl. *Az árokparton cseresznyefasor*). Jelen kutatás ezzel a típussal nem kíván foglalkozni. Érdekes azonban, hogy a „köznyelvi” adatok között is találunk 31 egzisztenciális vagy birtokos, illetve 55 lokatív mondatot. Ezeket közelebbről vizsgálva az is észrevehető, hogy előbbiekből több mint kétharmada (23 darab) valamilyen tagadó vagy kérdő névmást (*sehol, semmi, soha, mi*) tartalmaz, például: *Címekre semmi szükségem*. A birtokos esetén érdekes még, hogy ezen az altípuson kívül csak a *vége* fordul elő testes létige nélküli mondatokban (pl. *A teremtsnek vége*). A lokatív mondatokra szintén nagy arányban (az adatok nagyjából felében) jellemző a névmás (*itt, ott, hol*) jelenléte, például: *Ott az a nagy gumikerék*. Az „Egyéb” típusban a *minek* és a *miért* kezdetű kérdések fordultak elő (pl. *De akkor miért ez a nagy termelés?*).

Ezek alapján további korpuszbeli keresésre és vizsgálatra érdemesnek tünnek *a semmi* + főnév és *az itt/ott/hol* + NP szerkezetek. A lokatívok esetén ez ugyan számszerűen csak az összes eset felét fedi le, a kiinduló találatok másik fele azonban nehezen körülírható, és sokszor műfajspecifikus. A sajtónyelvben például gyakoriak a hasonló példák: *Intenzív osztályon a Bielefeld labdarúgója*. A kiinduló kis adathalmazból tehát úgy tűnik, hogy a köznyelvi zéró lokatív létigék jellemzően az említett névmásokkal együtt fordulnak elő.

4.2 További keresések

A mutató névmást tartalmazó lokatív mondatokat a következő alakú reguláris kifejezésekkel kerestem:

```
[word="[Hh]ol"] [word="az"]? [msd="DET|SZN.NOM"] [msd="MN.NOM"] *
[msd="FN.(PS[et]\d.)?NOM"]
```

Erre [1] bármelyik példája illeszkedik. Hasonló kereséseket végeztem az *itt*-re és az *ott*-ra is. A kapott találatokból egy Pythonban írt szűrővel zártam ki azokat, amelyek a keresőkifejezésre ugyan szintén illeszkednek, de tartalmaznak valamilyen (igei vagy névszói) állítmányt. Ilyenek [2] példái.

- [1] a. *Hol a macska?*
 b. *Hol a fehér macskád?*
 c. *Hol az a büdös macska?*
- [2] a. *Miért lóg ott az a bádogsillag?*
 b. *Ott a víz feneketlen.*

A szűrés után az egyes névmásokra a 3. táblázat szerinti találatszámokat kapjuk.

Típus	Határozott névelős találatok száma	Határozatlan névelős találatok száma	Összesen
hol	6758	250	7008
ott	6783	563	7346
itt	26574	2181	28755

3. táblázat. Találatok az *itt/ott/hol* típusokra

Ezek a számok önmagukban még nem mondanak sokat, annyi látszik biztosan, hogy a határozatlan névelő előfordulása viszonylag ritka. A többes számú NP-kre is végeztem keresést, erre azonban egyetlen találat volt: *Ha angolul van, hol az adatvédelmi beállítások?*

Az egzisztenciálisok és birtokosok keresőkifejezései a következők voltak:

```
[lemma="[Ss]emmi"] [msd="FN.NOM"]
[lemma="[Ss]emmi"] [msd="FN.PS[et]\d.NOM"]
```

Emellett a korpusz szűrő funkcióját alkalmaztam a finit igei tartalmazó találatok szűrésére. Így 14 444 zéró létigét tartalmazó egzisztenciális, és 17 331 birtokos mondatot kaptam.

5 Az eredmények értelmezése

A korpuszból kapott eredményekkel kapcsolatban a legfontosabb kérdés, hogy: mihez viszonyítsuk a számokat, mennyire számít soknak a hét- vagy a tizenkétezer találat? Ennek eldöntésére Clark (1978) zérólétige-tipológiáját alkalmaztam egyetlen nyelven belül, esetünkben a magyar nyelvre.

Clark (1978) szerint a létigés mondatoknak van egy univerzális hierarchiája, amely megmutatja, hogy mennyire valószínű az ige nélküli szerkesztésmód az adott mondattípusban. Az általa felállított sorrend a következő:

- predikatív
- azonosító
- birtoklásmondat2 (be-possessive)
- lokatív
- egzisztenciális
- birtoklásmondat1 (have-possessive)

Az első három mondattípus az 1. táblázat kopula típusának felel meg. A be-possessive alatt az *Ez a macska a szomszédomé* típusú kopulás mondatokat értjük. A hierarchia elsősorban tipológiai megfigyeléseken alapszik, de egy nyelven belül is jelen lehet. Úgy tűnik, a magyarban bizonyos feltételek mellett a hierarchia egy pontjái kötelező a zéró kopula, utána pedig opcionális. Ez nem egyedülálló eset, például az oroszban is minden jelen idejű predikatív, azonosító és lokatív mondatban testetlen a létige, de az ige nélküli szerkesztésmód opcionálisan a többi mondattípusban is előfordulhat. (I. Kádár 2011:34)

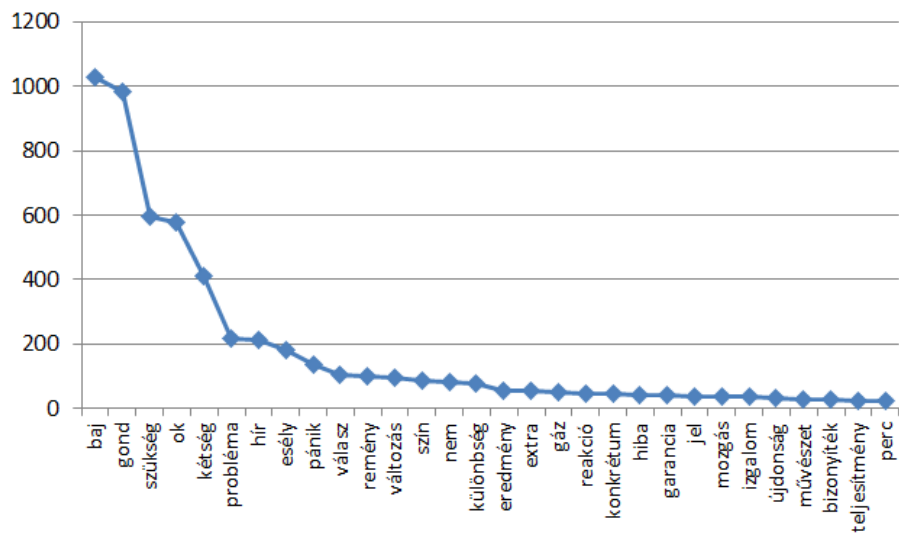
A kutatás alapvető, számítógépes nyelvészeti motivációjához is éppen annak megállapítására van szükség, hogy az egyes mondattípusokban mennyire valószínű a zéró létigés szerkesztésmód. Ez a korpusz segítségével úgy vizsgálható, ha a 4.2-ben ismertetett keresőkifejezésekbe beillesztjük a hiányzó létigét, és az így kapott találatok számát vetjük össze az azonos szerkezetű, ige nélküli mondatok számával. Ezen keresések eredményeit a 4. táblázat tartalmazza.

Típus	Testes létigék száma	Zéró létigék száma	Zéró létigék aránya
kopula	n. a.	n. a.	~100%
hol	5062	7008	58%
ott	9336	7346	44%
itt	10788	28755	73%
lokatív össz.	25186	43109	63%
egzisztenciális	8710	12213	58%
birtokos	13079	17331	57%

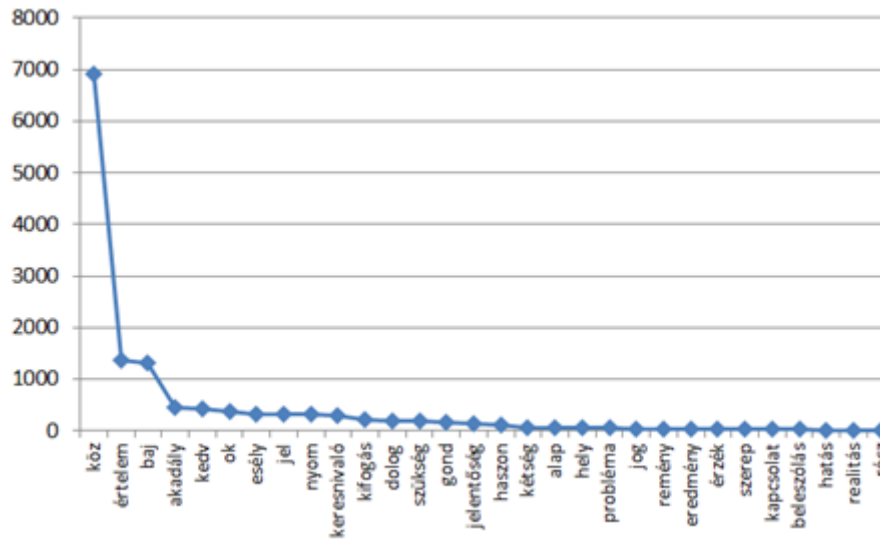
4. táblázat. A testes és zéró létigék eloszlása a vizsgált mondattípusokban

A zéró létige arányát természetesen úgy kell érteni, hogy az adott típusban a megfelelő feltételek mellett (jelen idő, harmadik személy) milyen arányban testetlen a létige. A lokatív, az egzisztenciális és a birtokos mondatok esetén ezek az arányok a Clark-hierarchiának megfelelő sorrendben követik egymást, de a különbségek nem olyan markánsak. Azt is meg kell jegyezni, hogy a vizsgálat nem vonatkozott az

összes lokatív, egzisztenciális és birtokos mondatra, csak azok egy-egy sajátos változatára. Ezeknél a mondattípusoknál tehát a hagyományos kopulához képest jóval több a megszorítás. Az eredmények mégis meglepőek, úgy tűnik ugyanis, hogy a névmások mellől gyakrabban hiányzik a testes létige ezekben az esetekben, mint ahányszor jelen van. Különösen az egzisztenciális és a birtokos mondatok nagyarányú igeelhagyása feltűnő. A jelenség okának felderítésére elvégeztem ezen mondatok alanyainak gyakorisági vizsgálatát, vagyis azt mértem, mely alanyok esetén milyen gyakorisággal jellemző a nincs létige elhagyása. Az eredmények grafikonos szemléltetését az 1. és 2. ábrák tartalmazzák.



1. ábra. Az egzisztenciális mondatok alanyainak gyakorisági megoszlása (30 leggyakoribb szótó)



2. ábra. A birtokos mondatok alanyainak gyakorisági megoszlása (30 leggyakoribb szótó)

A 4. táblázatban bemutatott értékekkel ellentétben az alanyok gyakorisági vizsgálata már mutat látványos különbséget az egzisztenciális és a birtokos mondatok között. A vizsgált zéró létigét tartalmazó egzisztenciális mondatokban összesen 2496-féle szótó fordul elő az alany helyén, a birtokos mondatoknál ezzel szemben mindössze 418, holott a birtokosok eleve majdnem harmincezerrel többen voltak az egzisztenciálisoknál. Amint a 2. ábra görbéjén is látszik, a birtokos mondatoknál az összes gyakoriság nagy részét néhány nagyon gyakori elem teszi ki, az összes előfordulás 90,5%-át lefedi a 15 leggyakoribb szótó (*köz, értelem, baj, akadály, kedv, ok, esély, jel, nyom, keresnivaló, kifogás, dolog, szükség, gond, jelentőség*). Ezzel szemben az egzisztenciálisok első 1500 leggyakoribb szótóve (amelyeknek jelentős része már a legritkább, 1 gyakoriságúak közül kerül ki) is csak az összes eset 89,9%-ának felel meg. Úgy tűnik tehát, hogy az egzisztenciális *semmi* + főnév szerkezet jóval produktívabb a birtokosnál. Mindemellett mindkét esetről elmondható, hogy ezeket a szerkezeteket (vagy legalábbis közülük a gyakoriakat) valószínűleg érdemesebb lenne a szófajcímekre alapozott szabályok helyett lexikálisan, kollokációként kezelni.

6 Következtetések

A korpuszadatok szerint az *itt, ott, hol* névmások mellett a létige elhagyása gyakori, de a hagyományos kopulával ellentétben csak egyes számban lehetséges. Amint [3] mutatja, a testes létige nélküli lokatív mondatokban az alany kötelezően névelős, és az adatok szerint a névelő az esetek kb. 93%-ában határozott.

- [1] a. Itt *(van) Németország.
 b. Itt *(van) János kutyája.
 c. Itt a pénzed.

Bár a korpuszvizsgálat a szórendekre nem tért ki, a mutató névmást tartalmazó lokatív mondatokról általánosságban elmondható az is, hogy a létige csak semleges szórend mellett lehet zéró:

- [2] a. Ott (van) a kedvenc képem.
 b. A kedvenc képem ott *(van).

Érdekes (az adatoktól független) észrevétel még, hogy a mutató névmás beillesztésével más helyhatározós, létige nélküli mondatok is „megjavíthatók”:

- [3] a. Az asztalon?(van) a kávé.
 b. Az asztalon ott (van) a kávé.

Sőt, az is előfordul, hogy az *ott* valójában nem is helyre utal a mondatban, hanem a létige egzisztenciális jelentését erősíti. Egy korpuszbeli példa:

- [4] De nagyobb fejtörésnek ott volt nekik a ló.

Az egzisztenciális jelentés és a lokatív névmások kapcsolata más nyelvekben is ismert jelenség, például az angol *there* és a kínai *you* grammatikalizálódott formái is létezését fejeznek ki. (Hoekstra–Mulder 1990) Az adatok alapján, úgy tűnik, a magyarban is megfigyelhető ez az összefüggés. Ennek részletesebb vizsgálata érdekes elméleti kutatási téma lehet.

Az egzisztenciális és birtokos létigék jellemzően tagadó mondatokban lehetnek zéró alakúak. Ehhez szükséges valamilyen tagadó névmás (pl. *semmi*, *sehol*) jelenléte. A testes létigét nem tartalmazó birtokos mondatokban alanyként megjelenhető névszók köre lexikálisan behatároltnak látszik; úgy tűnik, a zérós szerkesztésmód ennél a típusnál nem annyira produktív.

7 Alkalmazás

A kutatás eredményei alapján visszatérhetünk a 2. fejezetben feltett kérdésekre: mi lehet a leghatékonyabb módja a zéró létigék számítógépes elemzésének. Ha a tartalmasfej-megközelítésnél maradunk, akkor azt kell újragondolni, hogy tartalmaznak-e a lokatív mondatok létigéjét. Ha nem, mi lesz ezeknek a mondatoknak a feje? Ha igen, mit kezdünk a testes létigét nem tartalmazó lokatív mondatokkal? Még nagyobb probléma az egzisztenciális és birtokos létigék kezelése, hiszen ezek tartalmassága aligha vitatható, mégis láthattuk, hogy sok esetben lehetnek zéró hangalakúak. A dolgozatban vizsgált jelenségeket figyelembe véve tehát egyértelműsége miatt előnyösebbnek tűnik a funkciófej-elemzés.

Ebben a keretben az állítmány, azaz a mondat fejének keresése egy pszicholingvisztikai indíttatású elemzőben, tehát balról jobbra történő beolvasással, a következőképpen képzelhető el az *itt/ott/hol* + NP típusú mondatokban:

Input: ott/itt/hol + (egyres számú) NP
 Ha előtte nem volt „alkalmasabb” állítmány → illessz be egy
 potenciális zérót: ott/itt/hol + (van)? + NP
 Csatold hozzá a szükséges éleket (határozók, alany)!

Olvasd végig a mondatot, és folytasd az elemzést!
 Ha találtál másik állítmányt → az addig a zéróhoz kapcsolt éleket
 (alany, határozók) ő kapja meg! A zéró törlődik.

Példa egy elemzésre:

Input eleje: „Ott egy hangya...”
 Kezdeti elemzés: Ott (van)? egy hangya...
 Input vége a.: „... a fűben.”
 Elemzés a.: a mondat feje a (van).
 Input vége b.: „... mászik.”
 Elemzés b.: a mondat feje a mászik.

A zéró létige „előre jóslásának” jelentősége abban áll, hogy a kutatás motivációjául szolgáló, a 2. fejezetben ismertetett elemzőeszköz valós időben, szavanként elemez, így számára nagyon hasznos az ige esetleges hiányának minél korábbi jelzése, még ha ez később „backtrack”-kel is járhat. Az adatok szerint azonban a kérdéses jelenség elég gyakori ahhoz, hogy az elemző jó okkal feltételezzen potenciális zéró létigét az elemzés előbbiekben látott korai pontján.

A vizsgált egzisztenciális és birtokos típusok automatikus elemzésében a *semmi* + főnév szerkezeteket érdemes egy egységként kezelni, és a leggyakoribb ilyen egységeket akár lexikálisan betanítani a gépi eszköznek mint olyan egzisztenciális vagy birtokos jelentésű kollokációkat, amelyek nem feltétlenül kívánják meg a *nincs* létige jelenlétét.

8 Összegzés

A zéró létige a magyarban nem csak a hagyományos értelemben vett névszói állítmányt érintő jelenség. Különböző feltételek mellett bármelyik típusú létige opcionálisan elhagyható jelen idő harmadik személyben, azaz a bevezetőben tett hipotézissel ellenkezően az egzisztenciális jelentésű mondatok létigéje is. A Clark-hierarchiával egybehangzóan azonban elmondható, hogy az egzisztenciális és birtokos mondatok esetén a legmegszorítottabb az ígétlen szerkesztésmód lehetősége: ezekben a létige többnyire csak akkor lehet zéró, ha a mondatban tagadó névmás szerepel.

A lokatív mondatok esetén különösen gyakori jelenség a létige elhagyása, ha a mondatban mutató névmás található. Ez arra utal, hogy van valamilyen szemantikai összefüggés a névmások és a létezés jelentéstartalma között.

Mivel a zéró alakot felvehető létigék esetén úgy tűnik, nem mindig lehet egyértelműen állást foglalni a tartalmasság kérdésében, a függőségi elemzésekben érdemesebb lehet a funkciófej-megközelítést követni. Mind az *itt/ott/hol* + NP, mind a *semi* + főnév szerkezetek előfordulása testes létige nélküli mondatokban elég gyakorinak bizonyult ahhoz, hogy ezeket egy elemző programnak valamilyen módon tanítsuk. Erre a 7. fejezetben egy pszicholingvisztikai indíttatású elemző működési elveit követve láttunk példát.

Hivatkozások

- Clark, E. V. 1978. Locational: Existential, Locative and Possessive Constructions. In: Greenberg, J. H. (szerk.) *Universals of Human Language IV. Syntax*. Stanford: Stanford University Press. 85–126.
- Heggie, L. 1988. *The Syntax of Copular Structures*. PhD-értekezés. Los Angeles: University of Southern California.
- Hoekstra, T., Mulder, R. 1990. Unergatives as copular verbs: Locational and existential predication. *The Linguistic Review*, 7: 1–79.
- Kádár, E. 2011. *A kopula és a nominális mondatok a magyarban*. Budapest: Akadémiai Kiadó.
- Maienborn, C. 2005. A Discourse-based Account of Spanish *ser/estar*. *Linguistics*, 43(1): 155–180.
- Oravecz, Cs., Váradi, T., Sass, B. 2014. The Hungarian Gigaword Corpus. In: Calzolari, N., Choukri K., Declerck T., Loftsson H., Maegaard B., Mariani J., Moreno A., Odijk J., Piperidis S. (szerk.) *Proceedings of Ninth International Conference on Language Resources and Evaluation (LREC 2014)*. Elérhető: http://www.lrec-conf.org/proceedings/lrec2014/pdf/681_Paper.pdf. Letöltve: 2017. szeptember 26.
- Prószéky, G., Indig, B. 2015. Magyar szövegek pszicholingvisztikai indíttatású elemzése. *Hungarian Journal of Applied Linguistics*, 2015(1–2): 29–44.
- Simkó, K., Vincze, V. 2016. *Egyszer „van”, hol nem „van”: A létige kezelése függőségi nyelvtanokban*. In: Tanács, A., Varga V., Vincze V. (szerk.) XII. Magyar Számítógépes Nyelvészeti Konferencia (Szeged, 2016. január 21–22.). Szeged: JATEPress. 243–250.
- Stassen, L. 1994. Typology Versus Mythology: The Case of the Zero-Copula. *Nordic Journal of Linguistics*, 17(2): 105–126.