

XV. Magyar Számítógépes Nyelvészeti Konferencia Szeged, 2019. január 24–25.

Hol ugat a kutya? Örömeiben.

Helyhatározói esetragos névszók pontosabb annotációja

Ligeti-Nagy Noémi^{1,2}, Novák Attila^{2,3}

¹Pázmány Péter Katolikus Egyetem, Bölcsész- és Társadalomtudományi Kar
2087 Piliscsaba, Egyetem u. 1.

²MTA-PPKE Magyar Nyelvtechnológiai Kutatócsoport

³Pázmány Péter Katolikus Egyetem Információs Technológiai és Bionikai Kar
1083 Budapest, Práter u. 50/A

ligeti-nagy.noemi, novak.attila@itk.ppke.hu

Kivonat Tanulmányunkban ismertetjük a helyhatározói esetragos névszók pontosabb annotációját célzó kutatásunkat, melyet egy szövegekkel kapcsolatban releváns kérdéseket megfogalmazni képes elemzőrendszer igényei motiválnak. A *Hol?*, *Honnan?* és *Hová?* kérdésekre felelő három-három-három esetrag egyikét magán viselő névszók kategorizációja, a mondatban betöltött határozói szerepének pontosabb definiálása elkerülhetetlen a határozókra irányuló megfelelő kérdések megfogalmazásához. Cikkünkben a magyar UD-korpusz alapján 30 kategóriát mutatunk be, melyek megfelelőek ahhoz, hogy a velük annotált névszók határozói szerepe felismerhető és kérdezhető legyen.

1. Bevezetés

Jelen tanulmányunkban azt vizsgáljuk, hogyan kategorizálhatóak a névszótövek az alapján, hogy a kilenc hagyományos helyhatározói esetraggal való előfordulásuk során milyen szerepet töltenek be a mondatban, elsősorban szabad határozóként, nem az ige kötelező bővítményeként. Az 1. példában szereplő *bAn* esetragos szóalokról egyértelműen tudjuk, hogy nem egy helyszínt jelöl, hanem valamilyen formát; az *írásban* szóalakra nem a *Hol kérte a vezetőség a különvéleményeket?*, hanem a *Milyen formában kérte a vezetőség a különvéleményeket?* kérdéssel tudunk a legjobban rákérdezni.

(1) A különvéleményeket *írásban* kérte a vezetőség.

Amint azt a példánk magyarázatával már részben szemléltettük, vizsgáltunkat egy olyan elemzőrendszer víziója motiválja, amely képes egy szöveggel kapcsolatos releváns kérdések megfogalmazására [1]. Ennek eléréséhez szükséges egy megfelelően annotált korpusz, amely tanítóanyagként szolgálhat. Ebben olyan annotációnak kell szerepelnie, amely a fentihez hasonló, határozói szerepű névszók (névszói csoportok) pontosabb kategorizálását mutatja.

Ha a kérdésfeltevés és válaszolás irányából közelítjük a kérdést, nyilvánvaló, hogy az esetragok önmagukban nem nyújtanak elegendő információt az adott szóalakra irányuló megfelelő kérdés megfogalmazásához, vagy éppen egy adott kérdésre a megfelelő válasz megtalálásához. Ha csupán a szótőre és a morfológiai elemzésre (*Case:Ine*) támaszkodunk, nincs elég információnk ahhoz, hogy eldöntsük, a 2a és a 2b példákban a dőlttel szedett főnévi csoport jó válasz-e egy *Hol* kezdetű kérdésre - vagy megfordítva a dolgot: a főnévi csoportra a *Hol...?* kérdéssel kérdezzünk-e.

- (2) a. A Péterfy kórház sürgősségi belgyógyászati és klinikai toxikológiai osztálya több szempontból is érintett lehet *a különleges nappal kapcsolatban*.
- b. Országsszerte „nagyüzem” várható *a detoxikálókban*, illetve a mérgezési osztályokon.

Szeretnénk tehát olyan tudást kódolni a korpuszba, amiből megtanulható, hogy mi a különbség aközött, hogy *Ubul öltönyben ment dolgozni* vagy *Ubul decemberben ment dolgozni*. Illetve hogy a *főiskolán* szóalak minden gond nélkül lehet helyhatározói vonzata egy igének, de a *főiskolában* szóalak kevésbé - holott mindkettő a *főiskola* szótő és egy helyhatározói esetrag kombinációja.

Ehhez a Universal Dependencies korpusz [2] magyar alkorpuszában (mely a Szeged Treebank függőségi elemzett változatának [3] UD verziója¹) OBL függőségi viszonyal annotált², a kilenc helyhatározói esetrag egyikét magukon viselő szóalakokat kategorizáltuk.

Az természetesen minden, a magyar nyelv esetrendszerével foglalkozó irodalomban világosan megfogalmazásra került, hogy az esetragok és a mondatbeli szerepek között nem áll fenn egyértelmű megfeleltetés (ld. például [4], [5]). A szótövek részletes csoportosításával, szótő és esetrag beható vizsgálatával azonban nem foglalkoztak.

2. Módszer

Első lépésként a Szeged Dependency Treebank [3] egy részének UD-sémára átírt változatából válogattuk ki a függőségi elemzésben OBL éllel az igéhez kapcsolt szóalakokat, amelyeken a *Hol?*, *Hová?* és *Honnan?* kérdésre válaszoló, a belső (inesszívuszi *bAn*, illatívuszi *bA*, elatívuszi *bÓl*), külső (adesszívuszi *nÁl*, allatívuszi *hOz*, ablatívuszi *tÓl*), illetve felületi (szuperesszívuszi *On*, szublatívuszi *rA*, delatívuszi *rÓl*) helyhatározás paradigmáját alkotó 3-3-3 esetrag egyike található.

Ezután vettük a szavaknak ezt a kilenc csoportját, és mindegyik csoportot először „elő-kategorizáltuk” a word2vec modelleken alapuló szóbeágyazási modell [6] segítségével. A modell klaszterezésre hierarchikus klaszterezést alkalmaz.

¹ https://github.com/UniversalDependencies/UD_Hungarian-Szeged

² Ezek egy része szabad határozó, más része vonzat.

Ennek bemenete jelen esetben a csoportosítandó szavakhoz tartozó szemantikai vektor. A klaszterezés részleteit Siklósi és Novák [6]-ban és [7]-ben ismerteti. Mindehhez a vizualizációs felület volt a segítségünkre [8].

Végül a klaszterezéssel kapott, általában 3-8 elemű csoportok listáját manuálisan javítottuk, elsősorban azt a célt tartva szem előtt, hogy az azonos kérdőszóval kérdezhető, tehát azonos típusú határozóként előforduló szavak maradjanak egy csoportban. Fontos hangsúlyozni, hogy ez a csoportosítás ebben az első fázisban nem a szótöveket, hanem a ragozott alakokat érintette.

A feladat megfogalmazható úgy is, hogy határozókat csoportosítunk: vannak természetesen helyhatározók, mint a *sarkon*, vagy a *bankban*, vannak időhatározók, mint a *télen*, *decemberben*. De persze találkozunk időtartam-határozókkal is, mint az *Öt hónapra béreltük a lakást*. mondatban a *hónapra*.

Második lépésként már a szótövekre koncentráltunk. Az eddigi kategorizációt igyekeztünk általánosítani. Ha egy szótő egy adott esetraggal valamilyen határozóként funkcionált, akkor az vajon általános tulajdonsága a szótőnek, vagy csak azzal az egy bizonyos esetraggal együtt jellemző rá? Ha általános tulajdonsága, akkor van olyan esetrag, amivel együtt viszont más határozóként viselkedik?

A manuális kategorizálás végére 1100 szótövet soroltunk be valamilyen alapértelmezett kategóriába, illetve jelöltük meg külön, ha egy esetraggal együtt az alapértelmezettől eltérő viselkedés jellemzi.

3. Eredmények

3.1. A határozóragos névszók főkategóriái

A 2.táblázatban látható az a 30, illetve az alkategóriákkal együtt összesen 50 kategória, amelyekbe a szótöveket soroltuk. A főkategóriák valamilyen szemantikai kategóriának feleltethetők meg; ezen belül az alkategóriák általában az adott kategórián belüli esetrag-preferenciákat jelzik. A kategóriák részletesen a következők:

- *body*: testrészek nevei; egy részük (*bAn-On*, pl. *fej*) a *bAn* és *On* esetragokkal, illetve ezek irányhármasság szerinti paradigmájának egyéb tagjaival jelölnek *testrészhatározót*, más részük (*any*, pl. *derék*) bármely esetraggal állhat.
- *build=inst*: olyan köznevek, melyek egyszerre jelölnek egy fizikai épületet és valamely intézményt, például *bank*
- *cause*: okhatározók
- *circumst*: körülményhatározók; a megfelelő esetraggal együtt a *Milyen körülmények között?* kérdésre válaszolnak, pl. *hátrány*
- *curr*: pénznemek, pl. *forint*
- *date*: időhatározók; ezek között is van olyan, amely a *bAn* esetragot (és annak paradigmáját) preferálja (*percben*), és olyan, amelyik az *On-t* (*héten*)
- *dem*: a mutató névmások, illetve az *-é* birtokjellel ellátott szavak, melyek a *Melyikben?*, *Melyikbe?* stb. kérdésekre felelnek, pl. *előbbi*
- *direct*: irányhatározók; ezek között is megfigyelhető a belső helyhatározás esetragjainak (*oldalirány*) és a felületi helyhatározás esetragjainak kiegészítő eloszlása (*délnyugat*)

- *event*: eseményhatározók; különlegességük, hogy a megfelelő esetraggal egyszerűen fejeznek ki helyet és időt. *Találkoztam a barátommal az előadáson. Mikor találkoztál vele?*, vagy *Hol találkoztál vele?* Alkategóriákra osztható a csoport, esetrag-preferencia alapján. Például *háborúban, tüntetésen*.
- *form*: formahatározók; a *Milyen formában?* kérdésre válaszolnak. Például *szóban, papíron*.
- *group*: embercsoportok határozója, pl. *család*. Bizonyos esetragokkal *Hol?* kérdésre felel, helyhatározói szerepben; más esetragokkal viszont nem helyhatározó (pl. *családon*).
- *loc*: sok elemű csoport; vegyes szemantikai kategóriájú szavak, melyek helyhatározói szerepben állnak a megfelelő esetragokkal. Erős az esetrag-preferencia. Az alkategóriákat az 1. táblázatban részletesen bemutatjuk.
- *loc = who*: olyan, elsősorban földrajzi jellegű nevek, melyek bizonyos esetragokkal az *org = who* kategória elemeihez hasonlóan a *Kitől?* stb. kérdésekre felelnek. Pl. *EU-tagállam: Felkérés érkezett négy EU-tagállamtól*. esetében *Kitől érkezett felkérés?* a releváns kérdés. (Természetesen itt nem szabad határozóról, hanem az ige vonzatáról beszélünk.)
- *material*: anyaghatározók; elsősorban *bÓl* esetraggal kapcsolódva, a *Milyen anyagból?* kérdésre válaszolva
- *meas*: mértékegységek nevei; a *Mennyiben?* stb. kérdésekre zömében a saját módosítójukkal együtt válaszolnak.
- *mode*: módhatározók
- *num* és *num2*: számnevek; a különbség a két csoportban a *rA* és *rÓl* esetraggal kapcsolatos: a *négyre, 3500-ra* típusú *num* kategóriájuk a *Mennyire?*, míg a *negyedére, tizenötszörösére* típusú *num2* kategóriájuk inkább a *Mekkorára?* kérdésre felelnek.
- *org*: szervezetek, cégtípusok, vállalatok. Esetrag-preferenciával: *Hol? A Gazpromnál*, de **a Gazpromban*; viszont *a cégben* és *a cégnél*
- *org = who*: olyan szervezetek, hivatalok, melyek bizonyos esetragokkal *Ki?* formájú kérdésre felelnek. Ezekben az esetekben az ige szerepe is jelentős - ha *Jött egy levél a banktól*, akkor *Kitől jött a levél?*; viszont ha *Elindultam a banktól?*, akkor *Honnan indultál el?*. Ebben a kategóriában is megfigyelhető az esetrag-preferencia.
- *part*: részhatározók; bár releváns ezekre irányuló kérdés a *Hol?*, *Honnan?* stb., mindegyik esetraggal kérdezhetőek a *Melyik részében?*, *Melyik részénél?* stb. kérdésekkel is.
- *period*: időtartam-határozók.
- *place*: tárgyak, fizikai helyek nevei. Jellemzőjük, hogy a *nÁl*, *hÓz* és *tÓl* esetragok mindegyiknél helyhatározói funkciójuk érvényesül (tehát a *Hol?*, *Hová?* és *Honnan?* kérdésre felelnek ezekkel), a *bAn* és *On* esetragok (és ezek irányhármasság-beli társai) azonban komplementerei egymásnak. *Árokban*, de *autópályán*.
- *posi*: tisztségek, pozíciók, szerepek nevei. Esetrag-preferenciával. A *nÁl* esetrag, és annak irányhármasság-beli társai egyik alkategóriában sem váltják ki a helyhatározói funkciót.

- *poz*: szemponthatározók - a lemmák a megfelelő esetraggal ellátva a *Milyen szempontból?* kérdésre felelnek.
- *state*: állapothatározók; csak a *bAn*, a *bA* és a *bÓl* esetrag váltja ki a szótő állapothatározói funkcióját.
- *thing*: azoknak a szavaknak a kategóriája, amelyek nem töltenek be speciális határozói szerepet a mondatban, hanem a *Miben?*, *Minél?* stb. kérdésekre válaszolnak (legtöbbször az ige vonzataként).
- *way*: úttal, útvonallal kapcsolatos kifejezések. A belső helyhatározás paradigmájának esetragjai kivételével minden esetraggal betöltik ezt a funkciót.
- *who*: *Kiben?*, *Kinél?* stb. kérdésekre felelő szavak.

3.2. A *loc* kategória alkategóriái

Az alkategóriák szemléltetésére a valóban helyhatározást szolgáló, *loc* kategóriába sorolt töveket mutatjuk az 1. táblázatban.

kategória	példa		esetrag		
	fő	al	bAn	nÁl	On
loc	any	<i>szekrény</i>	hol	hol	hol
loc	bAn	<i>állam</i>	hol	minél	min
loc	nÁl	<i>pék</i>	miben	hol	min
loc	On	<i>címoldal</i>	miben	minél	hol
loc	bAn-On	<i>könyv</i>	hol	minél	hol
loc	city-bAn	<i>Párizs</i>	hol	hol	melyik városon
loc	city-On	<i>Miskolc</i>	melyik városban	hol	hol
loc	country	<i>Afganisztán</i>	hol	hol	melyik országon

1. táblázat. A meghatározott esetragokkal társulva helyhatározói szerepet betöltő szavak alkategóriái. Az alkategória oszlopa a legtöbb esetben valamely esetrag-preferenciát jelöl: a *pék* csak a *nÁl* esetraggal felel a *Hol?* kérdésre, míg az *állam* a *bAn*-nal. A *city-bAn*, *city-On* és *country* példák szemantikai információk alapján is elkülönítik a csoport tagjait. Az országnevek a korpuszban szereplő esetek mindegyikében a belső helyhatározás és a külső, ponthoz kapcsolódó helyhatározás paradigmájának elemeivel töltenek be helyhatározói funkciót.⁴ A városnevek két csoportra oszthatók: az inesszívuszi paradigmát (*Esztergomban*) és a szuperesszívuszi paradigmát (*Szegeden*) preferálókra. Az utóbbi csoportba tartozik a történelmi Magyarország területén található települések többsége, az előbbibe minden egyértelműen külföldi település. Mindkét csoport tagjai tölthetnek be helyhatározói funkciót a külső, ponthoz köthető helyhatározás paradigmájának esetragjaival. A táblázatban szereplő *Melyik városban?*, *Melyik városon?* és *Melyik országon?* kérdést nem úgy kell értelmezni, hogy azok egy helyhatározóra kérdeznek rá (pl. *Melyik városban laksz?*); a *Melyik városban bízol?* - *Budapestben.* esetre utalnak.

⁴ Megjegyzendő, hogy bár a korpuszban nem szerepelnek, vannak olyan országnevek, amelyek helyhatározói szerepükhöz az *On* esetragot igénylik: *Fülöp-szigeteken*, *Izlandon*. Ezek mind szigetek, de nem minden szigetország neve viselkedik így (pl. *Kubában*, *Írországban*).

4. Általánosságok és kivételelességek

Bár a 3. fejezetben bemutatott fő- és alkategóriák rendszere nagyon merev osztályozásnak tűnhet, részletesen megnézve az adatokat, azaz az 1100 szótó kategorizációját, számtalan példát találunk kivételekre, az alapértelmezettől eltérő viselkedésre.

Nagyon gyakori természetesen az is, amikor egy adott szótó mind a kilenc esetraggal a *default* kategóriájának megfelelő viselkedést produkálja. Ezek a szótók tulajdonképpen a kategóriáik prototipikus példányai. Néhány példa ezekre:

- *body_any*: *derék*
- *build_inst*: *ügyészség*
- *date_bAn*: *1987*
- *group*: *család*
- *org_bAn*: *cég*
- *period*: *félév*
- *place_On*: *célállomás*
- *posi_On*: *hatalom*
- *who*: *árus*

Szintén nagyon gyakori, hogy a szó az esetek nagy részében (kilencből 6-7-szer) az alapértelmezett kategóriájánál leírt módon funkcionál a mondatban, de 1-2 esetben ettől eltérő viselkedést produkál. Néhány példa ezekre:

- *alkalom*: alapértelmezett kategóriája az *event_On*; de a *bÓl* esetraggal együtt okhatározó (*cause*).
- *egy*: alapértelmezett kategóriája a *date*; de bizonyos esetragokkal módhatározó (*egyben* vagy *egyből*).
- *eleje*: alapértelmezett kategóriája a *part*; bizonyos esetragokkal időhatározói szerepben állhat (*elején*, *elejétől*).
- *előadás*: ez egy esemény (*event_On*), de a *bAn* esetraggal *Mikor?* kérdésre felelő időhatározó (*date*) vagy *Milyen formában?* kérdésre felelő formahatározó (*form*).
- *semmi*: alapértelmezetten a *thing* kategóriába soroltuk. A *semmi*be szóalak azonban a mennyiséghatározók *Mennyibe?* kérdésére is felel; a *mennyiből* szóalak pedig a helyhatározók *Honnan?* kérdésére is.

A fenti pár szó csak néhány példa a sok közül, de jól szemléltetik, hogy számos lemmánál a merev kategorizáció helyett rugalmasan kell eljárni, néhol az alapértelmezett kategória funkcióját teljesen felülírva, néhol csak kiegészítve azt más lehetőségekkel.

Sok esetben pedig nem minősül elégnek a korábban bemutatott 50 kategória: néhol az adott tőre és esetragra specifikus kérdéseket kell definiálni a kategóriák helyett. Néhány példa erre:

- A *többség* szótó a *thing* címkét kapta alapértelmezetten, de bizonyos esetragokkal a *Milyen arányban?*, vagy *Az érintettek mekkora részénél?* jellegű releváns kérdéseket lehet feltenni vele kapcsolatban.

- A *méter*, *kilométer* szavak mértékegységek. Kategóriájuk (*meas*) kérdései megfelelnek minden esetragnál. A *rA* esetraggal együtt azonban a *Milyen messze?*, *Milyen messzire?* kérdésekre felelő határozók is lehetnek. (Valamint természetesen *Hol?* kérdésre felelő helyhatározók is.)

A fentiekén kívül meg kell említenünk, hogy néhány kategória elemeire kifejezetten jellemző, hogy nem rájuk, hanem velük kérdezzük, az őket módosító elemre.

- Jellemző ez az eljárás a pénznemek kategóriájára: a *rA* esetraggal együtt például mindnek releváns kérdése a *Hány <pénznem neve>-rA?* A *400 forintra emelkedett a benzin ára.* esetében a *Mennyire?* és a *Hány forintra?* is megfelelő kérdések.
- Hasonló a helyzet a mértékegységek neveinél is. *Hány kilométerre?*, *Hány százalékon?*, *Hány tonnától?* stb.: ezek mind megfelelő kérdések, és ez mindegyik esetragnál valid kérdésfeltevési forma.

Végül nem hallgathatjuk el, hogy van egy olyan kategóriánk, mely nem alapértelmezett címkéje egyetlen szótőnek sem. Ez a *cause*, az okhatározók kategóriája. A korpusz alapján azt találtuk, hogy bár bizonyos szótövek bizonyos esetragokkal okhatározói funkciót töltenek be, alapértelmezett kategóriaként egyetlen szóra sem illik az okhatározói címke. Néhány példa:

- *apropó: thing*, de *apropóján*, *apropójából*
- *cél: loc_any*, de *célből*
- *megfontolás: thing*, de *megfontolásból*
- *nyomás: thing*, de *nyomásra*

5. Konklúzió

Tanulmányunkban bemutattuk, hogy egy olyan elemzőrendszert szem előtt tartva, amely szövegekről releváns kérdést képes feltenni, a határozók esetében milyen annotációt tartunk megfelelőnek egy tanítókorpusz létrehozásánál. Elemzésünkben a függőségi elemzett korpusznak azokkal az OBL viszonytal annotált elemeivel foglalkoztunk, amelyek az irányhármasság paradigmáinak valamely esetragját viselik magukon. Ezek nagy része szabad határozói szerepet tölt be, más részük vonzat. 30, az alkategóriákkal együtt összesen 50 kategóriát állapítottunk meg, amelyekbe a korpusz fenti kritériumnak megfelelő szavai besorolhatóak. A legtöbb szótőnél nem elégséges egy alapértelmezett kategória jelölése, bizonyos esetragoknál az alapértelmezettől eltérő funkciót tölthetnek be a szavak, melynek címkézése szintén feladat. A bemutatott kategorizáció megfelelő jegyeket biztosít egy tanítókorpuszban egy az előzőekben említett kérdezőrendszer létrehozásához. Maradtak azonban nyitott kérdések: azokban az esetekben, mikor egy szótő és egy adott esetrag kombinációja több lehetséges határozói funkciót is fedhet, más támpontok kellenek a biztos kérdés megtalálásához. Az *órára* esetében szükséges a névszó módosítóit is ismernünk a döntésünkhöz: ha *az órára*, akkor

Hova?, de ha *néhány órára*, akkor *Mennyi időre?*. Néhány kategória elemeinél tehát elkerülhetetlen az elemek módosítóinak ismerete, és a módosító és a névszó együttesének felcímkézése is. Kutatásunkat ebbe az irányba visszük tovább, kiegészítve mindezt azokkal a határozókkal, melyek nem az eddig vizsgált kilenc esetrag egyikét, hanem más esetragot, például az eszközhatározó esetragját viselik magukon. A következő lépés természetesen az itt bemutatott kategorizációnak és annotációs javaslatnak a mérése, elsősorban a kérdezőrendszer működése során tesztelve és kiértékelve.

Köszönetnyilvánítás

Jelen kutatás az FK 125217 számú projekt keretében az FK 17 pályázati program finanszírozásában a Nemzeti Kutatási Fejlesztési és Innovációs Alap által biztosított támogatással és az Emberi Erőforrások Minisztériuma ÚNKP-18-3-III-PPKE-26 kódszámú Új Nemzeti Kiválóság Programjának támogatásával valósult meg.

Hivatkozások

1. Novák, A., Laki, L.J., Novák, B., Dömötör, A., Ligeti-Nagy, N., Kalivoda, Á.: Egy magyar nyelvű kérdezőrendszer. In: XV. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2019), Szeged, SZTE (2019)
2. Nivre, J., de Marneffe, M.C., Ginter, F., Goldberg, Y., Hajic, J., Manning, C.D., McDonald, R., Petrov, S., Pyysalo, S., Silveira, N., Tsarfaty, R., Zeman, D.: Universal dependencies v1: A multilingual treebank collection. In Chair), N.C.C., 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), Paris, France, European Language Resources Association (ELRA) (2016)
3. Vincze, V., Szauter, D., Almási, A., Móra, G., Alexin, Z., Csirik, J.: Hungarian dependency treebank. In Chair), N.C.C., Choukri, K., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S., Rosner, M., Tapias, D., eds.: Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10), Valletta, Malta, European Language Resources Association (ELRA) (2010)
4. Antal, L.: A magyar esetrendszer. In Bottyán, G., Kis, Á., eds.: A formális nyelvi elemzés. A magyar esetrendszer. SZAK Kiadó Kft. (2005)
5. Kiefer, F.: A ragozás. In Kiefer, F., ed.: Strukturális magyar nyelvtan 3. Morfológia. Akadémiai Kiadó (2000) 569–618
6. Siklósi, B., Novák, A.: Beágyazási modellek alkalmazása lexikai kategorizációs feladatokra. In Tanács, A., Varga, V., Vincze, V., eds.: XII. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2016). (2016) 3–14
7. Siklósi, B., Novák, A.: Közeli rokonunk, az autó. In Tanács, A., Varga, V., Vincze, V., eds.: XII. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2016). (2016) 27–36
8. Novák, A., Siklósi, B., Wenszky, N.: Szóbeágyazási modellek vizualizációjára és böngészésére szolgáló webes felület. In: XIII. Magyar Számítógépes Nyelvészeti Konferencia. (2017) 355–362