

A Nyelvtechnológiai és Alkalmazott Nyelvészeti Osztály jelentése a 2015. évről

Kiemelkedő kutatási és más jellegű eredmények

Nyelvtechnológiai Kutatócsoport

A *Nyílt, integrált magyar nyelvtechnológiai kutatási infrastruktúra fejlesztése (INFRA2)* projekt célul tűzte ki egy magyar nyelvtechnológiai elemzőkészlet keretrendszer kifejlesztését, eszközök és erőforrások sztenderdizációját. Ennek "Morfológiai elemző és kapcsolódó alkalmazások" modulja 2015-ben előkészítő munkálatokat és az elemző keretrendszer felépítésére vonatkozó fejlesztéseket végzett, melynek eredményeként elkészült egy nyílt alkalmazásokra (HFST és Giellatekno keretrendszer) épülő nagy lefedettségű és nagy sebességű morfológiai elemző prototípus. Elkezdődött egy unicode-os szövegek tokenizálását végző program fejlesztése. Elkészült a program mondatrabontó modulja, a tesztelőkörnyezet és a unicode-os karakterosztályokat tartalmazó modul. A projekt "GATE integráció" moduljában előrehaladtak a magyar nyelvtechnológiai eszközök közös keretbe foglalásának munkálatai is. A kiválasztott közös keretrendszer a Sheffieldi Egyetem gondozásában fejlesztett GATE Developer nevű szoftver, melyet világszerte sokan használnak és számos nyelvtechnológiai kiegészítés elérhető hozzá. A munkálatok célja a magyar nyelvhez jelenleg elérhető elemző eszközök (szó- és mondatfelismerők, szófaji egyértelműsítők, morfológiai és szintaktikai elemzők stb.) beillesztése és feldolgozási láncokba szervezése a GATE rendszerében. Ebben az évben sikerült beilleszteni a Magyarlánc nevű rendszer 7 különböző eszközét. Elkészült több tanulmányprojekt, melyek célja a nem Java nyelven írt magyar nyelvtechnológiai eszközök hatékony befoglalásához szükséges módszerek megismerése volt. Vizsgálatok folytak arról, hogy hogyan lehet a GATE-ben kézzel, illetve automatikusan korpuszt annotálni, valamint hogyan lehet GATE eszközöket korpuszon kiértékelni (pontosságot/fedést meghatározni).

Lezárult az *ECOINNO* projekt, ennek keretében elkészült a nyelvtechnológiai támogatással működő, ökoinnovációs szolgáltatások és igények dinamikus illesztésére szolgáló portál (<http://www.ecoinno.hu>). A portál az érdeklődő vállalkozások számára automatikusan közvetít ki a tevékenységükhöz illeszkedő olyan korszerű zöldtechnológiákat, melyek segítségével energiatakarékosabbá és környezettudatosabbá tehetik tevékenységüket. A portál mögött működő nyelvi elemzőrendszer és a kifejlesztett lexikális erőforrások és algoritmusok teszik lehetővé a vállalkozásokról elérhető szöveges adatok feldolgozását és ezekből a releváns információ kinyerését. A nyelvtechnológiai fejlesztést részletesen a The International Conference on Semantic Web Business and Innovation (SWBI2015) konferencia kötetében megjelent tanulmány írja le.

A *Finn-OTKA* projekt keretein belül párhuzamos és összevethető korpuszokat, valamint prototípusokat hoztunk létre a következő nyelvpárokra: {komi-zürjén, komi-permják, hegyi mari, mezei mari, udmurt, északi számi} - {angol, finn, magyar, orosz}. Ezek jól felhasználhatók további kutatásokban, alkalmazások fejlesztésében, illetve a párhuzamos és összevethető korpuszok szótár bővítésben is alkalmazásra kerülnek. Emellett elkészültek nagy szószámú egynyelvű korpuszok a fent említett veszélyeztetett nyelvekre, melyek nyelvtechnológiai eszközök fejlesztésében, létrehozásában nyújtanak segítséget. A projekt munkálatai során szerzett tapasztalatok és az eddig létrehozott szótárak és korpuszok 2015-ben több nemzetközi és hazai konferencián ismertetésre kerültek.

A *Magyar Generatív Szintaxis 2 (MGTSZ2)* című OTKA-projektben az év folyamán megtörtént a teljes Münchener kódex normalizálása és morfológiai egyértelműsítése. A kódex a keresőfelületen

keresztül hozzáférhető. Elkezdődött a Jordánszky-kódex normalizálása és morfológiai egyértelműsítése, valamint Sylvester János bibliafordításának OCR-ezése a betűhű szöveg kinyerése céljából. Elkezdődött a morfológiailag elemzett ómagyar szövegek konverziója a Universal Dependencies formátumára, amely egy nemzetközi sztenderd, melyet már 38 nyelvre alkalmaztak.

Lezárult a *Magyar WordNet* erre az évre tervezett továbbfejlesztése. Ez egy olyan elektronikus szótár, amely a szavakhoz nem csak definíciókat, hanem szinonimákat és egyéb kapcsolódó fogalmakat is társít (hipernímák, hiponímák, antonímák) és az egész egy hálózatba szerveződik. A munka során átfogalmaztuk a definíciókat, illetve kialakítottunk kapcsolatokat a Magyar WordNet és más, nemzetközileg elismert ontológiák, illetve lexikai erőforrások között, hogy ezzel a Magyar WordNet bekerülhessen a linked open data közösség aktív tagjai közé. A Magyar WordNet legfrissebb változatának teljes anyaga szabadon hozzáférhető a <https://github.com/dlt-rilmta/huwn> címen. Megtörtént a Magyar WordNetben található információk átalakítása a NooJ program számára kezelhető formátumba és az eredmény összefésülése a morfológiai információt tartalmazó korábbi NooJ szótárral. Így lehetővé vált tartalmi keresések végzése korpuszokban.

2015 novemberében lezárult a *slova.ee* projekt. Elkészült a portál új dizájnya és az új tartalmak magyar fordítása. Az osztály feladata volt még a tesztelés, valamint november elején Budapesten megrendezett zárómegbeszélés megszervezése. A felhasználói visszajelzések alapján folyamatosan történik a *helyesiras.mta.hu helyesírási tanácsadó portál* fejlesztése, a hibák javítása. 2015 szeptemberében megjelent az akadémiai helyesírási szabályzat 12. kiadása, s ezzel együtt rövid határidővel megtörtént az új szabályok integrálása a portálba. Mivel a 11. kiadás 2016 szeptemberéig még érvényben lesz, a portál mind a 11., mind a 12. kiadás szerinti írásmódot feltünteti a kiadás feltüntetésével. Az osztály megkezdte a kutatási munkákat a *A multimodális kommunikáció időszervezete* című OTKA projektumban, a "kommunikatív aktusok felismerése gépi tanulási módszerekkel" modul keretében. A kutatócsoport részt vett a *Magyar nyelv nagyszótára* informatikai támogatásában, a Nagyszótárhoz használt "Magyar történeti szövegtár" korpusz bővítésének előkészítésében, valamint megkezdődött a korpusz új lekérdezőfelületének kialakítása. Kutatás folyt az igei szerkezetek algebrai modelljének kidolgozására.

Az *ELRC* (Európai Nyelvi Erőforrások Koordinációja) projekt keretében a Nyelvtechnológiai Kutatócsoport egynapos workshop-ot szervezett, amelyen olyan közigazgatási és kormányzati szereplők vettek részt, akik két- vagy többnyelvű szövegeikkel hozzá tudnak járulni az Európai Bizottság által újjá indított automatizált fordítóplatform (CEF.AT) hatékonyságának növeléséhez és a későbbiekben felhasználói lesznek a CEF.AT-nek. Elindult a *Nemzeti Korpuszportál (NKP)* kezdeményezés, melynek célja minél több magyar nyelvű, online lekérdezővel bíró korpusz összegyűjtése és közös ponton elérhetővé tétele. A NKP-ről többek között a Magyar Tudomány Ünnepe alkalmából rendezett tudományos ülésen esett szó. Elkészült egy külföldi kutatási infrastruktúrához való kapcsolódásról szóló pályázat, mely megteremti annak lehetőségét, hogy Magyarország csatlakozhasson digitális bölcsészet technológiai támogatását célul kitűző európai CLARIN ERIC együttműködéshez.

Az osztály továbbra is zavartalanul üzemelteti az *itranslate4.eu* portált. Látogatottsága 2015-ben: több mint 390 ezer felhasználó és több mint 4 millió oldalmegtekintés, a látogatók 53%-a visszatérő látogató. Az *EFNIL* (European Federation of National Institutions for Language) szervezet portálját is az osztály gondozza és üzemelteti, ennek látogatottsága 2015-ben több mint 6100 felhasználó mintegy 26000 oldalmegtekintését eredményezte. Az csoport a szervezet mindennapi működésében is aktívan részt vesz a titkári munkák elvégzésével.

Matematikai Nyelvészeti Kutatócsoport

Az *INFRA2* projekt "Hangadatbázisok" modulja keretében egy olyan webszerver készül a munkacsoportban, amely elsősorban magyar nyelvű hanganyagok nyilvános, illetve kutatási célú

hozzáférését teszi lehetővé. Ennek a nyílt beszédarchívumnak célja mind a szűkebb értelemben vett szakmai (nyelvészeti, antropológia, etnográfiai stb.) kutatások kiszolgálása, mind a modern beszédtechnológia (hosszú távon az automatikus beszéd felismerés, rövidebb távon a formátumkonverzió, beszélőazonosítás, beszéd/zenedetekció stb.) infrastrukturális támogatása. Elkezdődött a szoftver-infrastruktúra kialakítása, részint a Réger Zita hagyatékából készült archívum, részint több mint ezer magyar film hanganyagának letöltésével. Mind az adatbázis felépítéséhez, mind annak hatékony használatához szükség van az alábbi, folyamatban levő alprojektekre: (1) beszéd/nem-beszéd döntés (voice activity detection, VAD) kétféle lehetséges módszer összehasonlításával; (2) automatikus nyelvazonosítás akusztikus jegyek alapján; (3) beszélő-detekció (kísérletek történtek a hmmlearn csomag tesztelésére felügyelt és felügyeletlen (GMM-HMM) módszerrel egyaránt, és a Praat szoftver segítségével beszélő szerinti annotáció készült Réger Zita gyűjteményének, illetve régi magyar filmek hanganyagain); (4) nyelvmodellezés (n-gram nyelvmodellek készítését az SRILM eszközzel).

Angol és magyar nyelvre egyaránt elkészültek a *4lang* rendszer fő moduljai: a természetes nyelvi bemenethez automatikusan jelentésreprezentációt rendelő `text_to_4lang` és az egynyelvű szótárak definícióból definíciós gráfokat előállító `dict_to_4lang`. Utóbbi eszköz két magyar és három angol egynyelvű szótár felhasználását teszi lehetővé, a szó- és mondat szintű elemzéshez mindkét nyelvre elterjedt és nyílt forráskódú eszközöket vesz igénybe (magyarlanc ill. Stanford CoreNLP). A *4lang* rendszert is használó, mondatok szemantikai hasonlóságát mérő nyílt forráskódú eszközünk sikeresen szerepelt a nemzetközi Semeval versenyen, legjobb rendszerünk a 74 beküldött eszközből a 11. eredményt érte el - a 29 résztvevő csapatból csak 4-en előzték meg a MathLingBudapest csapatot. 2015-ben 80 százalékban elkészült a *Semantics* c. monográfia.

Az *EFNILEX* projekt a számítógépes lexikográfia újabb eredményeit alkalmazta európai nyelvekre. A projekt idején, befejező szakaszában a szavak többértelműségének problémája állt a kutatás középpontjában. Közvetítő nyelvvél generált fordításoknál lett megmérve, hogy a közvetítő nyelv szavainak többértelműsége mennyire rontja a fordítás megbízhatóságát.

Egy szeptemberben indult kutatási projekt középpontjában a *szójelentés vektoros modelljei* (embedding) állnak a következő témák köré csoportosítva: egy szó jelentéseinek elkülönítése több nyelvet figyelembe véve, a szóelemekből épülő kompozicionális és nem-kompozicionális szerkezetek elkülönítése és mindkettő kezelése (többszavas kifejezések, képzett szavak), valamint szavak szótári jelentése közötti viszonyok (ellentét, ok-cél, rész-egész, eszköz-rendeltetés) automatikus kigyűjtése.

Nyelvművelő és Nyelvi Tanácsadó Kutatócsoport

Több éven át tartó munkálatok eredményeképpen 2015 szeptemberében megjelent *A magyar helyesírás szabályai* 12. kiadása (AkH. 12.). A csoport egy munkatársa az új helyesírási szabályzat egyik szerkesztőjeként és az AkH. 12.-vel foglalkozó MTA munkabizottság tagjaként vett részt a munkában. A 2015. szeptember 3-án tartott sajtóbemutatón a szabályzatot az MTA elnöke, az Akadémiai Kiadó vezérigazgatója, az Emberi Erőforrások Minisztériumának helyettes államtitkára, a Magyar Nyelvi Osztályközi Állandó Bizottság jelenlegi és korábbi elnöke mutatták be, nagy érdeklődés mellett. A nyelvi tanácsadó szolgálat keretein belül történik az új szabályzat folyamatos utánkövetése.

A csoport munkatársai előadásokat tartottak és tanulmányokat készítettek az alábbi témakörökben: (1) nyelvi tanácsadás, helyesírás, stílusztika; (2) névtan, névkultúra, névjog.

A nyelvi tanácsadó szolgálat 2015-ben kb. 3000 emailben (tanacs@nytud.mta.hu) érkezett, és kb. 1300 telefonos helyesírási, nyelvhasználati kérdésre adott választ, 4 nyelvi szakvéleményt készített. A partnerek között megtalálhatók állami és európai uniós hivatalok, cégek, más szervezetek,

fordítók, tanárok, újságírók, magánszemélyek. A szolgálat emléktáblák (25 db) szakvéleményezésével is foglalkozott önkormányzatok és magánszemélyek kérésére. A csoport részt vesz a helyesiras.mta.hu Helyesírási tanácsadó portál működtetésében, folyamatosan egyedi válaszokat ad az érkező kérdésekre.

2015 februárjában a csoport munkatársa a magyar helyesírás gondozásában és tanításában végzett eredményes kutatói, tanári munkája elismeréseként Implom József-díjat vehetett át.

A kutatócsoport speciális feladatként az illetékes állami szervvel és az Intézetben erre a célra létrehozott Utónévbizottsággal együttműködve folyamatosan ellátja a hatósági anyakönyvezésre beterjesztett új utónevek nyelvi szakvéleményezését (a 2010. évi I. törvény 44.§-a alapján) és a bejegyzésre alkalmasnak minősített nevek gondozását. A bejegyzésre alkalmasnak minősített nevek listája az intézet honlapján hozzáférhető (<http://www.nytud.hu/oszt/nyelvmuvelo/utonevek>), havonta frissül. 2015-ben a lista 96 új névvel bővült (60 női, 36 férfi) így 2016. január 1-jén 3764 bejegyzésre alkalmasnak minősített utónév (2144 női és 1620 férfi) alkotja a jegyzéket. A nyelvi tanácsadó szolgálat 2015-ben kb. 2500 emailben (nevtanacs@nytud.mta.hu) érkezett, és kb. 500 telefonos névadással kapcsolatos, névhasználati kérdésre adott választ, 15 névhasználati szakvéleményt készített, valamint BÁH egyszerűsített honosítási eljárás során kért 50-60 keresztnévet véleményezett szakértőként.

Párbeszéd a tudomány és a társadalom között

Nyelvtechnológiai Kutatócsoport

A helyesiras.mta.hu Helyesírási tanácsadó portál célja, hogy új lehetőséget teremtsen a széles közvélemény számára a magyar helyesírás területein való tájékozódáshoz. A portál népszerűsége 2015-ben tovább növekedett: 860000 látogató (60%-uk rendszeres látogató) több mint 2 millió tanácsot kért itt. A lekérdezések száma munkanapokon rendszeresen eléri a 7000-t. Az MTA Nyelvtudományi Intézet egy olyan tudományos alapokon álló szolgáltatást üzemeltet immár két és fél éve, mely valóban sokakhoz eljut, sikeresen működik. A helyesiras.mta.hu portál munkatársai több csatornán végeznek tudományos ismeretterjesztő tevékenységet. A hivatalos Facebook-oldalon (és Twitteren) napi két poszt jelenik meg rajta különböző helyesírási, nyelvhelyességi, illetve egyéb érdekes, a magyar nyelvvel kapcsolatos témában. A Helyes blog immár több mint két éve sikeresen folytat ismeretterjesztő tevékenységet. Az egyik legnépszerűbb blogposztot több mint 2500-an ajánlották, több mint 22 ezer embert ért el a közösségi portálon és több hírportál is átvette. A blog idén részt vett a Kut@tók a Neten által szervezett Kutatói Blog, Web- és Facebook-oldal Találkozó, és harmadik helyezést ért el. Idén a Helyesírási portál és blog az MTÜ "Kutatóhelyek tárt kapukkal" programsorozaton is bemutatkozott a MTA Nyelvtudományi Intézetében megrendezett nyilvános eseményen, az előadást követően a közönség egy interaktív online kvízt tölthetett ki.

A *Nyílt, integrált magyar nyelvtechnológiai kutatási infrastruktúra fejlesztése* projekt egy olyan nyílt forrású, szabadon hozzáférhető nyelvtechnológiai infrastruktúra fejlesztését tűzte ki célul, melynek elemei a magyar nyelv gépi elemzésének alapvető eszközeit tartalmazzák egy integrált, szabványos keretben. Az elkészült infrastruktúra megfelelően szolgálja ki a magyar nyelvű szövegekkel vagy beszéddel foglalkozó kutatókat, érdeklődőket és a nyelvtechnológiát használó ipari alkalmazások széles körét.

Megnyílt a <http://www.ecoinno.hu> portál, mely dokumentumok automatikus nyelvi feldolgozására építve vállalkozások számára automatikusan közvetít ki a tevékenységükhöz illeszkedő olyan korszerű zöldtechnológiákat, melyek segítségével energiatakarékosabbá és környezet tudatosabbá tehetik tevékenységüket.

A csoport által fejlesztett nyelvi *adatbázisok* társadalmi szempontból is jelentősek. Ezen adatbázisok az anyanyelvi kulturális örökség digitális formában őrzött részei, melyek referenciapontként szolgálnak nemcsak a tudományos kutatásban, hanem a közgondolkodásban, az érdeklődő laikusok körében is. A több mint 10000 regisztrált felhasználóval bíró Magyar Nemzeti Szövegtár, a Ómagyar Korpusz, a BUSZI és egyéb online lekérdezővel bíró szabadon elérhető korpuszok 2015-től egy közös ponton, a Nemzeti Korpuszportálon érhetőek el a szakma és a nagyközönség számára. A Magyar Nemzeti Szövegtárban 2015-ben 95000 lekérdezést futtattak.

A csoport több munkatársa szerepelt a Kossuth Rádió *A Hely* című adásában, ahol mind a "hagyományos", mind az online nyelvi-helyesírási tanácsadásról, így a helyesiras.mta.hu portálról is szó esett. A csoport munkatársa állandó vendég a Lánchíd Rádió nyelvi, nyelvészeti ismeretterjesztéssel foglalkozó műsorában, nyelvhelyességi kurzust tart a Mária Rádió szerkesztőinek. A csoport munkatársa rendszeresen publikál tudományos ismeretterjesztő írásokat helyesírás témában egy orvosoknak szóló szaklapban. Ezenkívül a csoport munkatársai rendszeresen tartanak előadásokat a magyar NLP meetupon az általuk fejlesztett nyelvtechnológiai eszközökről és erőforrásokról. A meetup előadásainak célja, hogy a nyelvtechnológiával foglalkozó érdeklődőket, illetve a kutatásban és az iparban dolgozókat összehozza, a későbbi együttműködés érdekében közös fórumot biztosítson nekik.

A csoport a beérkező igényeknek megfelelően folyamatosan számítógépes nyelvészeti támogatást ad különböző nyelvészeti kutatásoknak, nyelvi adatokat szolgáltat, egyéni kéréseket teljesít, a Magyar Nemzeti Szövegtár és az ebből származtatott nyelvi erőforrások (pl. gyakorisági adatbázis) tekintetében, illetve informális nemzetközi együttműködés keretében (internetes korpuszok elemzése szlovák kutatóműhelyek számára).

Matematikai Nyelvészeti Kutatócsoport

A Magyar Tudomány Ünnepeén tartott előadáson a szófordítás példáján láthatta a nagyközönség, hogy az egyre nagyobb szöveggyűjtemények és az egyre gyorsabb számításra képes számítógépek hogyan váltják valóra a nyelv modellezésére vonatkozó statisztikai és mesterséges ideghálókkal dolgozó korai elképzeléseket.

Nyelvművelő és Nyelvi Tanácsadó Kutatócsoport

A csoport tevékenysége 2015-ben is hozzájárult a társadalmat foglalkoztató nyelvi kérdések megválaszolásához, nyelvi problémák megoldásához. Aktív szerepet vállalt a nyelvi ismeretek terjesztésében, a nyelvtudomány eredményeinek népszerűsítésében. Közvetlenül a *nyelvi tanácsadás*, a nyelvi közönségszolgálat, közvetve pedig az egyéni nyelvi ismeretterjesztő tevékenység jóvoltából.

Az azonnali válaszadásra lehetőséget adó telefonos közönségszolgálat heti két napon működött, az elektronikus levelezés (tanacs@nytud.mta.hu) pedig lehetővé tette a nyelvi produktumok véleményezését is. 2015-ben a helyesírás volt a leginkább közérdeklődésre tartó terület: a legújabb nyelvi változások és az új szóátvételek írásmódja, az *A magyar helyesírás szabályai* új 12. kiadása iránti várakozás, érdeklődés, majd a megjelenést követően a változásokra, az érvényességi időre irányuló kérdések uralták a tanácsadó szolgálatot. Emellett nyelvhelyességi és egyéb témájú kérdések is előfordultak. Az akadémiai helyesírás érvényesítése, különösen a nyilvánosság elé kerülő szövegek esetében továbbra is jelentős társadalmi igény: 2015-ben az információkéréseken túl közel 3000 írásbeli (e-mail) kérdés irányult a jelenlegi írásgyakorlatot érintő nyelvi, helyesírási problémákra.

A nyelvi, nyelvhasználati ismeretterjesztés különféle rendezvényeken, publikációkban, ismeretterjesztő kiadványok (pl.: *Édes Anyanyelvünk*) szerkesztésében valamint a médiában valósul

meg. A kutatócsoport munkatársai számos előadást tartottak különböző rendezvényeken, továbbképzéseken nyelvhelyesség, helyesírás, névadás témakörökben. E területek számos alkalommal megjelentek a médiában is, rádió- és tv-szereplések, nyomtatott sajtóban megjelenő interjúk formájában. Kiemelt megjelenés A magyar nyelv napja alkalmából az Európa Rádióban (<http://www.refradio.eu/radio/euradio>) készített többrészes interjú a mai magyar nyelv állapotáról, a nyelvhasználat és a stílus változási tendenciáiról, az új helyesírási szabályzatról, az irodalmi nyelvről. A csoport egy munkatársa részt vesz az MTA Nyelvhelyességi Tanácsadó Testületének munkájában és gondozza az MTA Nyelvtudományi Intézete és az Anyanyelvápolók Szövetségének szakmai kapcsolatát.

Törvényi kötelezettségének megfelelően a bejegyzésre alkalmasnak ítélt utónevek (férfi és női keresztnemek) listáját a kutatócsoport kezeli, illetve az újonnan engedélyezett nevekkel havi rendszerességgel bővíti, és az Intézet honlapján közzéteszi. A társadalom, a szülők részéről folyamatosan jelentkező igényt a lista bővülése is mutatja: 2015-ben 96 új névvel (60 női, 36 férfi) egészült ki a lista. Az utónév-szakvéleményezés szakértői közreműködést jelentett a hatóságnak, szakértői segítséget a szülőknek és a névváltoztatást kezdeményezőknél. Az Intézet utónevekkel foglalkozó munkáját bemutató januári előadás nagy érdeklődés mellett zajlott, róla tv-felvétel is készült. Az utóneveket bemutató Utónévportál 2015-ben több mint 40000 látogató 120000 lekérdezését szolgálta ki.

A kutatóhely hazai és nemzetközi kapcsolatai 2015-ben

Hazai kapcsolatok

Nyelvtechnológiai Kutatócsoport

A csoport különböző projektekben együtt dolgozik az MTA-PPKE Magyar Nyelvtechnológiai Kutatócsoporttal, a MTA-SZTE Mesterséges Intelligencia Kutatócsoport nyelvtechnológiai csoportjával, az SZTE Angol-Amerikai Intézetével, a KRE Magyar Nyelvtudományi Tanszékével, a DPH Diamond Project Holding Kft-vel és az AITIA Zrt-vel.

A csoport egy tagja az ELTE Alkalmazott Nyelvészeti doktori iskolájának hallgatója.

Várad Tamás: MTA Alkalmazott Nyelvészeti Munkabizottság -- tag

Pajzs Júlia: MTA Szótári Munkabizottság -- tag

Ludányi Zsófia: MTA Magyar Nyelvi Osztályközi Állandó Bizottság -- tag, MTA Orvosi Nyelvi Munkabizottság -- tag, MTA Nyelvhelyességi Tanácsadó Testület -- titkár

Matematikai Nyelvészeti Kutatócsoport

A csoport egyes tagjai doktoranduszok az ELTE-MTA Elméleti nyelvészeti kihelyezett tanszékén, valamint BME Villamosmérnöki karán, illetve Természettudományi Karán.

Nyelvművelő és Nyelvi Tanácsadó Kutatócsoport

Együttműködés a KRE BTK Nyelvtudományi Tanszékével a Terminológia MA- képzés és a szakmai gyakorlatok támogatásában. Együttműködés az Anyanyelvápolók Szövetségével.

Heltainé Nagy Erzsébet: MTA Magyar Nyelvi Osztályközi Állandó Bizottság -- tag, MTA Magyar Nyelvészeti Munkabizottsága -- tag, Magyar Nyelvtudományi Társaság -- tag, MTA Nyelvhelyességi Tanácsadó Testület -- tag, MANYE -- tag, Termini Egyesület -- alapító tag,

Anyanyelvapolók Szövetsége -- a választmány tagja, Magyar Nyelvőr - a szerkesztőbizottság tagja, Édes Anyanyelvünk - a szerkesztőbizottság tagja

Raátz Judit: MTA Magyar Nyelvi Osztályközi Állandó Bizottság -- tag, Anyanyelvapolók Szövetsége -- választmányi tag, Magyar Nyelvtudományi Társaság -- tag, Magyar Nyelvtudományi Társaság Névtani Tagozat -- tag, Magyar Nyelvtudományi Társaság Magyartanári Tagozat -- tag, Szemere Gyula anyanyelv-pedagógiai kutatócsoport -- tag, HUNRA Magyar Olvasástársaság -- tag

Nemzetközi kapcsolatok

Nyelvtechnológiai Kutatócsoport

A csoport Finn-OTKA projekt keretében együtt dolgozik a University of Helsinki, Institute of Behavioural Sciences intézettel.

Az INFRA2 projektben készülő magyar morfológiai elemző kapcsán szoros kapcsolat alakult ki a kutatócsoport és a Dublin City University valamint a University of Tromsø morfológiával foglalkozó kutatói között. 2015-ben Budapest adott helyet egy közös workshopnak.

A 30 ország részvételével zajló Európai Nyelvi Erőforrások Koordinációja (ELRC) projekt keretében Budapesten került sor egy kormányzati szereplők részére megrendezett workshopra.

Folytatódott a Cseh Akadémiával közös "Számítógépes lexikográfia és dialóguskutatás" együttműködési projektum. Szakmai látogatás történt a külföldi partnernél, konzultációval a készülő cseh szótár informatikai hátteréről.

A kutatócsoport munkatársai több külföldi továbbképzésen vettek részt: egy kutató a Prágában megtartott többszavas kifejezésekről szóló első PARSEME training school-on, többen pedig a Nagyszébenben megtartott EUROLAN 2015 nyári egyetemen.

Matematikai Nyelvészeti Kutatócsoport

A csoport két egymástól független német partnercégekkel közös munkában is részt vett 2015-ben. A TenBetterPages GmbH illetve a BuildingRadar GmbH részére végzett különböző nyelvtechnológiai fejlesztéseket.

Kornai András: ACL SIGFSMNLN tanácsadó testület -- tag, ACL SIGMORPHON tanácsadó testület -- tag, ACL SIGMOL tanácsadó testület -- tag

Nyelvművelő és Nyelvi Tanácsadó Kutatócsoport

A nyelvi tanácsadó szolgálatban nőtt a határon túlról érkező kérdések száma. 2015-ben mintegy 50 helyesírási, tájnyelvi, nyelvhasználati kérdés érkezett külföldi magyaroktól.

Domus magyarországi 2015. évi ösztöndíjpályázat, Misadová Katarína (Pozsonyi Egyetem) témavezetése. A pályázat címe: A magyar oktatási terminológia vizsgálata

Együttműködés a Termini kutatóhálózattal és egyesülettel. Részvétel a Termini (határon túli magyar nyelvi kutatóhálózat) Anyanyelvi nevelés és nyelvi revitalizáció c. konferenciáján, Szabadka, 2015. október 16.

Raátz Judit: International Council of Onomastic Sciences (ICOS) -- tag

Felsőoktatási tevékenység

Nyelvtechnológiai Kutatócsoport

Oravecz Csaba: Statisztikai módszerek anyelvechnológiában (előadás). Pázmány Péter Katolikus Egyetem, 2015. tavasz.

Oravecz Csaba, Mittelholcz Iván: Python a számítógépes nyelvészetben (szeminárium). ELTE Elméleti Nyelvészet szak, 2015. tavasz.

Ludányi Zsófia: Adatbázisok (előadás). Károli Gáspár Református Egyetem, 2015. tavasz.
Adatbázisok (szeminárium), Helyesírási gyakorlatok (szeminárium), Alaktan, szófajtan (szeminárium). Károli Gáspár Református Egyetem, 2015. ősz.

Munkatársak előadást tartottak az ELTE BTK korpusznyelvészeti mesterképzésében és a PPKE ITK nyelvtechnológiai mesterképzésében.

összes PhD opponensi feladat: 4; PhD bizottsági tagság: 4; MA/MSc témavezetés: 2; BA/BSc témavezetés: 1

Matematikai Nyelvészeti Kutatócsoport

Makrai Márton: Jelentésrepresentációk. MTA-ELTE elméleti nyelvészet kihelyezett tanszék.

Ács Judit: Programozás alapjai 1-2. (gyakorlat- és laborvezetés). BMGE.

doktori iskola törzstag: 1; összes PhD opponensi feladat: 1; PhD témavezetés: 7; MSc bíráló: 1; BSc bíráló: 3

Nyelvművelő és Nyelvi Tanácsadó Kutatócsoport

Heltainé Nagy Erzsébet, további jogviszonyú egyetemi docens KRE BTK Magyar Nyelvtudományi Tanszék: stilisztika, szöveg- és stílselemzés, retorikai gyakorlatok, szóalkotás, terminusalkotás. 2015. tavasz: 9 kurzus; 2015. ősz: 9 kurzus.

Kardos Tamás: Helyesírási gyakorlatok -- 2 kurzus. KRE BTK.

Raázt Judit, az ELTE Mai Magyar Nyelvi Tanszékének főállású docense: 12 kurzus oktatója.

összes PhD opponensi feladat: 7; PhD témavezetés: 5; MA/MSc témavezetés: 12; BA/BSc témavezetés: 2; BA/BSc bíráló: 1 -- valamint: tanári portfóliók, OTDK-dolgozatok és MTA Pedagógus Kutatói Pályázat pályamunkák bírálata.

Intézményi és pályázati együttműködések

intézeti kutató(k) neve	a pályázat címe	együttműködő			pályázati keret?
		intézmény neve	tanszék neve	vezető kutató neve	
Ács Judit, Kornai András, Makrai Márton, Pajkossy	Keresőszó-normalizáció a Cliqz webkereső számára	TenBetterPages GmbH	--	--	--

Katalin, Recski Gábor					
Benyeda Ivett, Koczka Péter, Ludányi Zsófia, Simon Eszter, Váradi Tamás	Finnugor nyelvű közösségek nyelvtechnológiai támogatása online tartalmak létrehozásában	University of Helsinki	Institute of Behavioural Sciences	Kristiina Jokinen	Finn-OTKA
Benyeda Ivett, Koczka Péter, Ludányi Zsófia, Simon Eszter, Váradi Tamás	Finnugor nyelvű közösségek nyelvtechnológiai támogatása online tartalmak létrehozásában	Szegedi Tudományegyetem	Angol-Amerikai Intézet	Fenyvesi Anna	Finn-OTKA
Kornai András, Miháltz Márton, Oravecz Csaba, Recski Gábor	Információkinyerő eszközök fejlesztése a Building Radar GmbH- nak	Building Radar GmbH	--	--	--
Ludányi Zsófia, Váradi Tamás	--	Károli Gáspár Református Egyetem	Magyar Nyelvtudományi Tanszék	Fóris Ágota	--
Mittelholcz Iván, Oravecz Csaba, Sass Bálint	ECOINNO	DPH Diamond Project Holding Kft.	--	Sárközy Csongor	ECOINNO
Simon Eszter	INFRA2	Dublin City University	School of Computing	Tommi Pirinen	INFRA2
Simon Eszter	INFRA2	University of Tromsø	Department of Language and Linguistics	Francis Tyers	INFRA2
Váradi Tamás	INFRA2	AITIA Zrt.	--	Tatai Gábor	INFRA2

Váradí Tamás	INFRA2	MTA-PPKE Magyar Nyelvtechnológiai Kutatócsoport	--	Prószéky Gábor	INFRA2
Váradí Tamás	INFRA2	Szegedi Tudományegyetem	Informatikai Tanszékcsoport	Csirik János	INFRA2

Az intézettel szerződéses kapcsolatban álló cégek

Az osztály 2015-ben a következő cégekkel állt szerződéses kapcsolatban: DPH Diamond Project Holding Kft., AITIA Zrt., TenBetterPages GmbH, BuildingRadar GmbH

Részvétel nemzetközi konferenciákon

Nyelvtechnológiai Kutatócsoport

előadó	előadás címe	konferencia elnevezése	helye (város)	ideje (hónap)
Benyeda Ivett	How can language technology fight against language death?	Digital Humanities Conference Benelux (DHBenelux)	Antwerpen, Belgium	június
Miháltz Márton	TrendMiner: Large-Scale Analysis of Political Attitudes in Public Facebook Messages	6th IEEE International Conference on Cognitive Infocommunications	Győr	október
Miháltz Márton	Beyond Sentiment: Social Psychological Analysis of Political Facebook Comments in Hungary	6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA 2015)	Lisszabon, Portugália	szeptember
Oravecz Csaba	Ontology-based Recommendation from Natural Language User Descriptions: Recommending Eco-Innovations to Companies	The International Conference on Semantic Web Business and Innovation (SWBI2015)	Sierre, Svájc	október
Simon Eszter, Koczka Péter, Benyeda Ivett	Automatic creation of bilingual dictionaries for Finno-Ugric Languages	First International Workshop on Computational Linguistics for Uralic Languages (IWCLUL)	Tromsø, Norvégia	január
Simon Eszter	Language technology support for Finno-Ugric digital communities	XII. International Congress for Finno-Ugric Studies	Oulu, Finnország	augusztus
Váradí Tamás	When will ITS Speak Your Language?	22nd ITS World congress	Bordeaux, Franciaország	október

			g	
Váradai Tamás	Language Technology in the Service of Intelligent Transport Systems	22nd ITS World congress	Bordeaux, Franciaország	október

Matematikai Nyelvészeti Kutatócsoport

előadó	előadás címe	konferencia elnevezése	helye (város)	ideje (hónap)
Makrai Márton	Disambiguated Linear Word Translation in Medium European Languages	CogInfoCom 2015	Győr	október
Kornai András	Basic vocabulary for language revitalization	Ugra Global Expert Meeting on Multilingualism in Cyberspace	Hanti-Manszijszk, Oroszország	július
Kornai András	Lexical Semantics and Model Theory: Together at Last?	Mathematics of Language 14	Chicago, USA	július
Kornai András	Realistic hidden structure	The Foundations of Linguistics: Languages as Abstract Objects Workshop	Braunschweig, Németország	június
Kornai András, Ács Judit, Makrai Márton, Nemeskey Dávid, Pajkossy Katalin, Recski Gábor	Competence in lexical semantics	4th Joint Conference on Lexical and Computational Semantics (*SEM 2015)	Denver, Colorado, USA	június
Recski Gábor, Ács Judit	MathLingBudapest: Concept Networks for Semantic Similarity	9th International Workshop on Semantic Evaluation (SemEval 2015)	Denver, Colorado, USA	június

Nyelvművelő és Nyelvi Tanácsadó Kutatócsoport

előadó	előadás címe	konferencia elnevezése	helye (város)	ideje (hónap)
Heltainé Nagy Erzsébet	Kerekasztal-beszélgetés vezetője a következő témákban: A magyar mint idegen nyelv tanításának kialakult gyakorlatáról; A magyar mint idegen nyelv felnőttoktatási kérdéseiről; A rokon népek magyarnyelv-tanítási gyakorlatáról	Magyartanárok találkozója és Hungarobox konferencia	Budapest	január
Raácz Judit	Jó gyakorlatok!	ELTE Magyar Nyelvtudományi és	Budapest	szeptember

Részvétel hazai konferenciákon**Nyelvtechnológiai Kutatócsoport**

előadó	előadás címe	konferencia elnevezése	helye (város)	ideje (hónap)
Benyeda Ivett, Jani Mátyás, Lukács Gergely	Beszéd-zene lejátszási listák nyelvtechnológiai vonatkozása	XI. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2015)	Szeged	január
Benyeda Ivett, Koczka Péter, Ludányi Zsófia, Simon Eszter	Automatikus szótárgenerálás finnugor nyelvekre	25. Magyar Alkalmazott Nyelvészeti Kongresszus (MANyE)	Budapest	március
Benyeda Ivett, Koczka Péter, Ludányi Zsófia, Simon Eszter, Váradi Tamás	Finnugor nyelvű közösségek nyelvtechnológiai támogatása online tartalmak létrehozásában	XI. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2015)	Szeged	január
Kuna Ágnes, Ludányi Zsófia	A magyar orvosi nyelv korpuszának tervezete	25. Magyar Alkalmazott Nyelvészeti Kongresszus (MANyE)	Budapest	március
Miháltz Márton	Trendminer: politikai témájú közösségimédia-üzenetek feldolgozása és szociálpszichológiai elemzése	XI. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2015)	Szeged	január
Oravecz Csaba	Az adatok és a nyelvi erőforrások megosztása: technikai és gyakorlati szempontok	CEF.AT Program - ELRC Training Workshop	Budapest	november
Oravecz Csaba, Sass Bálint, Váradi Tamás	Mennyiségből minőséget: Nyelvtechnológiai kihívások és tanulások az MNSz új változatának elkészítésében	XI. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2015)	Szeged	január
Pajzs Júlia	A Magyar WordNetből nyert szemantikai jegyekkel kombinált keresés a NooJ program segítségével	főmunkatársi előadás, MTA Nyelvtudományi Intézet	Budapest	szeptember
Sass Bálint	Egy kereslet-kínálat elvű elemző működése és a koordináció kezelésének módszere	XI. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2015)	Szeged	január
Sass Bálint	Korpuszok, lekérdezők,	DHU2015 „Számítógép az	Budapest	november

	Nemzeti Korpuszportál	irodalomtudományban” workshop		
Sass Bálint	28 millió szintaktikailag elemzett mondat és 500000 igei szerkezet	XI. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2015)	Szeged	január

Matematikai Nyelvészeti Kutatócsoport

előadó	előadás címe	konferencia elnevezése	helye (város)	ideje (hónap)
Ács Judit	Synonym acquisition from translation graph	XI. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2015)	Szeged	január
Makrai Márton	Elosztott nyelvmodellek összehasonlítása közepes nyelveken	XI. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2015)	Szeged	január

Nyelvművelő és Nyelvi Tanácsadó Kutatócsoport

előadó	előadás címe	konferencia elnevezése	helye (város)	ideje (hónap)
Heltainé Nagy Erzsébet	Sólyom Réka: Mai neologizmusok szemantikája (könyvbemutató)	KRE BTK, szakmai tanszéki nap	Budapest	május
Heltainé Nagy Erzsébet	Nyelvi változások - megújult nyelvi tanácsadás	Implom József Középiskolai Helyesírási Verseny Kárpát- medencei döntő, Szakmai program	Gyula	február
Heltainé Nagy Erzsébet	A nyelvtechnológia szerepe a nyelvi tanácsadásban	25. Magyar Alkalmazott Nyelvészeti Kongresszus (MANyE)	Budapest	március
Heltainé Nagy Erzsébet	Önök szerint hogyan helyes? - Kérdések és válaszok a Nyelvtudományi Intézet tanácsadó szolgálatában	Simonyi Zsigmond helyesírási verseny Kárpát-medencei döntő, tanári továbbképzés	Budapest	május
Heltainé Nagy Erzsébet	Szathmári István köszöntése	Ünnepi konferencia a 90 éves Szathmári István tiszteletére	Budapest	február
Raátz Judit	A keresztnévadás kérdései	Névtan és terminológia 4. "Személynevek és államigazgatás"	Budapest	június
Raátz Judit	Kommunikációelmélet és gyakorlat a kerettantervben és a tankönyvekben	Tanítói és magyartanári konferencia a kommunikációs nevelésről: VIII. "Tudatosság a kommunikációban"	Eger	május
Raátz Judit	Szabályok és trendek a hazai utónévadásban	Családnevek és utónevek a névjog és névtan	Budapest	december

Rendezett konferenciák**Nyelvtechnológiai Kutatócsoport**

szervező	konferencia elnevezése	helye (város)	ideje (napra!)	(társszervező, ha volt)
Falyuna Nóra, Kovács Réka, Ludányi Zsófia, Váradi Tamás	IX. Alkalmazott Nyelvészeti Doktoranduszkonferencia	Budapest	2015. február 6.	--
Kovács Réka, Váradi Tamás	Budapesti ELRC workshop	Budapest	2015. november 23.	Európai Bizottság Fordítási Főigazgatósága
Simon Eszter	HFST workshop	Budapest	2015. június 2.	--

Nyelvművelő és Nyelvi Tanácsadó Kutatócsoport

szervező	konferencia elnevezése	helye (város)	ideje (napra!)	(társszervező, ha volt)
Raátz Judit - TÁMOP 4.1.2.B.2-13/1-2013-0007 "Országos koordinációval a pedagógusképzés megújításáért"	Színház - dráma - iskola	Budapest	2015. február 12-13.	Maladype Színház
Raátz Judit - TÁMOP 4.1.2.B.2-13/1-2013-0007 "Országos koordinációval a pedagógusképzés megújításáért"	A magyar nyelv és irodalom érettségi vizsga 10 éve (2005-2015)	Budapest	2015. március 27.	--
Raátz Judit - TÁMOP 4.1.2.B.2-13/1-2013-0007 "Országos koordinációval a pedagógusképzés megújításáért"	VIII. "Tudatosság a kommunikációban" konferencia és IV. Vályi András anyanyelv-pedagógiai konferencia	Eger	2015. május 9.	EKF
Raátz Judit, Farkas Tamás, Slíz Mariann - ELTE Magyar Nyelvtudományi és Finnugor Intézete	Névtan és terminológia 4. "Személynevek és államigazgatás"	Budapest	2015. június 17.	Károli Gáspár Református Egyetem Magyar Nyelvtudományi Tanszéke és Terminológiai Kutatócsoportja, a Magyar Nyelv Terminológiai Tanácsa és a Magyar

A 2015-ben elnyert fontosabb hazai és nemzetközi pályázatok rövid bemutatása

Az osztály 2015-be nyerte el a *Nyílt, integrált magyar nyelvtechnológiai kutatási infrastruktúra fejlesztése (INFRA2)* című jelentős pályázatot. E kutatásiinfrastruktúra-fejlesztési pályázat célja egy olyan nyílt forrású, szabadon hozzáférhető nyelvtechnológiai infrastruktúra kiépítése, melynek elemei a magyar nyelv gépi elemzésének alapvető eszközeit tartalmazzák egy integrált, szabványos keretben. Az elkészült infrastruktúra megfelelően szolgálja ki a magyar nyelvű szövegekkel vagy beszéddel foglalkozó kutatókat és a nyelvtechnológiát használó ipari alkalmazások széles körét. A kifejlesztett eszközök, erőforrások és keretrendszer kapcsolódnak a stratégiai kutatási infrastruktúra státuszt elért HunCLARIN-on keresztül a CLARIN ERIC infrastruktúrába.

Az osztály NKFIH pályázatot nyert el *A multimodális kommunikáció időszerkezete* címmel, melynek keretében a Debreceni Egyetemmel együttműködve a kommunikatív aktusok gépi tanulási módszerekkel történő felismerését kutatják.

pályázó vezető kutató neve	pályázat címe	pályázat kiírója	pály. azonosító száma	a projekt futamideje (től-ig)	teljes támogatási összeg	együttműködő partnerek	a projekt fő célja egy mondatban megfogalmazva
Várad Tamás	Infrastruktúra fejlesztés 2015 ("Pályázat az akadémiai kutatási infrastruktúra fejlesztésére" tárgyú pályázat 2. kategória) -- Támogatott tevékenység: Nyílt, integrált magyar nyelvtechnológiai kutatási infrastruktúra építése	Magyar Tudományos Akadémia Titkársága	IF-023/2015.	2015. március 6 -- 2016. június 30.	86,2 MFt	--	Nyílt, integrált magyar nyelvtechnológiai kutatási infrastruktúra építése.
Hunyadi László, Várad Tamás	A multimodális kommunikáció időszerkezete	Nemzeti Kutatási Fejlesztési és Innovációs Hivatal	K 116402	2015-2028	16,28 MFt	Debreceni Egyetem	A multimodális kommunikáció időszerkezetének vizsgálata T-minta elemzés segítségével, a kommunikatív aktusok felismerése gépi tanulási módszerekkel.

A 2015-ben megjelent jelentősebb tudományos publikációk

Nyelvtechnológiai Kutatócsoport

Benyeda Ivett, Koczka Péter, Ludányi Zsófia, Simon Eszter, Váradi Tamás. Finnugor nyelvű közösségek nyelvtechnológiai támogatása online tartalmak létrehozásában. In: Tanács Attila, Varga Viktor, Vincze Veronika (szerk.): XI. Magyar Számítógépes Nyelvészeti Konferencia, SzTE, Szeged, 2015. p. 133-144.

http://rgai.inf.u-szeged.hu/mszny2015/files/MSZNY2015_press_B5_PQ.pdf

Duray Zsuzsa – Georgievna, Ekaterina – F. Gulyás Nikolett – Koczka Péter – Mus Nikolett – Németh Szilvia) XII. Nemzetközi Finnugor Kongresszus – Beszámoló. [XII International Congress for Finno-Ugric Studies]. Nyelvtudományi Közlemények 111. 215–228.

Ludányi Zsófia 2015. Kétéves a helyesiras.mta.hu tanácsadó portál. Édes Anyanyelvünk 37(2): 14. https://docs.google.com/viewerng/viewer?url=http://edesanyanyelvunk.anyanyelvapolo.hu/pdf/%C3%89des_Anyanyelv%C3%BCnk-2015-04-02.pdf

Márton Miháltz, Tamás Váradi. TrendMiner: Large-Scale Analysis of Political Attitudes in Public Facebook Messages. In Proceedings of The 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom 2015), Győr, Hungary, pp. 265-266.

<https://docs.google.com/viewer?>

[a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbmxtbWloYWx0enxnDo3MDE2OWYwZjcyNmU0MDM4](https://docs.google.com/viewer?)

Oravecz, Csaba, Bálint Sass. Csongor Sárközy Ontology-based Recommendation from Natural Language User Descriptions: Recommending Eco-Innovations to Companies In: Proceedings of the International Conference on Semantic Web Business and Innovation (SWBI2015), 44-56, Sierre, Switzerland, 2015. <http://sdiwc.us/digitlib/request.php?article=a007b7c9c934e3965514afbd5ca4b1f2>

Pajzs Júlia: Szótár és enciklopédia az online világban In: Fábián Zsuzsanna, Szöllősy Éva (szerkesztő): Szótár, lexikon, enciklopédia. Kérdések és feladatok Tinta Könyvkiadó Budapest, 2015 p. 120-126.

Sass Bálint 2015. 28 millió szintaktikailag elemzett mondat és 500000 igei szerkezet. In: Tanács Attila -- Varga Viktor -- Vincze Veronika (szerk.) MSZNY 2015, XI. Magyar Számítógépes Nyelvészeti Konferencia. Szeged: JATEPress. 303--308.

http://rgai.inf.u-szeged.hu/mszny2015/files/MSZNY2015_press_B5_PQ.pdf

Simon, Eszter , Ivett Zs. Benyeda, Péter Koczka, Zsófia Ludányi. Automatic creation of bilingual dictionaries for Finno-Ugric languages. In: Proceedings of the First International Workshop on Computational Linguistics for Uralic Languages, Tromsø, 2015. pp. 119-131.

<http://septentrio.uit.no/index.php/SCS/article/view/3474>

Váradi Tamás (szerk.) 2015. Technikai szerkesztés (Ludányi Zsófia, Kovács Réka, Falyuna Nóra): Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2015 IX. Alkalmazott Nyelvészeti Doktoranduszkonferencia Budapest, 2015. 02. 06. MTA Nyelvtudományi Intézet: Budapest. ISBN 978-963-9074-62-0 <http://www.nytud.hu/alknyelvdok15/proceedings/anyd9.pdf>

Váradi, Tamás, Marko Tadić, András Gulyás, Mihai Niculescu. Language Technology in the Service of Intelligent Transport Systems. In Proceedings of the 22nd ITS World Congress, Bordeaux,

France. Paper number ITS-2878.

Matematikai Nyelvészeti Kutatócsoport

Ács, Judit: Synonym acquisition from translation graph (2015)
Magyar Számítógépes Nyelvészeti Konferencia 11, 14-21
http://rgai.inf.u-szeged.hu/mszny2015/files/MSZNY2015_press_B5_PQ.pdf

Kornai, A.; Kracht, M: Lexical Semantics and Model Theory: Together at Last? In M. Kuhlmann, M. Kanazawa and G. Kobele (eds) Proc. MOL 2015 51-61
[.http://www.aclweb.org/anthology/W/W15/W15-2305.pdf](http://www.aclweb.org/anthology/W/W15/W15-2305.pdf)

Makrai Márton. Disambiguated Linear Word Translation in Medium European Languages. In: Proceedings of CogInfoCom 2015, Győr, 2015 október. <http://corpus.nytud.hu/efnilex-vect/pub/makrai15-disambig-abs.pdf>

Recski, G. and J. Ács (2015): MathLingBudapest: Concept Networks for Semantic Similarity. In: Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015). pp. 543-547. Denver, CO, U.S.A., Association for Computational Linguistics.
<http://www.aclweb.org/anthology/S15-1019>

Nyelvművelő és Nyelvi Tanácsadó Kutatócsoport

Heltainé Nagy Erzsébet; Sólyom Réka: A mai magyar neologizmusok szemantikája. Magyar Nyelvőr 139. évf. 2015/1. szám, 123–126. <http://nyelvor.c3.hu/period/1391/139110.pdf>

Raátz Judit 2015. Alkalmazott névtan. In: Farkas Tamás – Slíz Mariann (szerk.) Magyar névkutatás a 21. század elején. Budapest: Magyar Nyelvtudományi Társaság – ELTE Magyar Nyelvtudományi és Finnugor Intézet. 227–47. <http://mek.oszk.hu/14700/14775/14775.pdf>