

A Nyelvtechnológiai és Alkalmazott Nyelvészeti Osztály jelentése a 2013. évről

Kiemelkedő kutatási és más jellegű eredmények

Nyelvtechnológiai Kutatócsoport

2013. április 30-án megnyílt a *helyesiras.mta.hu Helyesírási tanácsadó portál*. Az intelligens portál a Nyelvtechnológiai Osztály önálló fejlesztése, melynek célja a széles közönség helyesírással és nyelvhasználattal kapcsolatos leggyakoribb kérdéseinek automatizált kiszolgálása. A tanácsadás a következő területekre terjed ki: egybeírás-különírás, általános helyesírás-ellenőrzés, tulajdonnevek helyesírása, keltezés és számnevek helyesírása, elválasztás, abc-rendbe sorolás. A belső tesztek, illetve a felhasználói visszajelzések alapján folyamatosan történik a fejlesztés, a hibák javítása. Elkészült a portál beágyazható alkalmazása, amelyet bárki beépíthet saját weboldalába.

Sikeresen zárult az intézet által koordinált CESAR projekt. A projekt keretében létrehozott META-SHARE disztribúciós rendszer tárolja és teszi közzé a projektben érintett 6 nyelv 251 nyelvi erőforrását és nyelvfeldolgozó eszközét. Az intézet a META-NET kiválósági hálózat vezetőbizottságának tagjaként részt vesz a META-SHARE hálózat fenntartásában a magyar csomópont üzemeltetése által. 2013-ban a kutatócsoport összesen négy európai pályázat elkészítésében vett részt, köztük egynek a koordinátora volt (DuaMed). Ebből a TrendMiner projekt volt sikeres, a munkálatok novemberben meg is kezdődtek.

Az ECOINNO pályázat keretében elkészült a nyelvtechnológiai támogatással működő, ökoinnovációs szolgáltatások és igények dinamikus illesztésére szolgáló portál rendszerterve. A nyelvtechnológiai elemzőlánc támogatására lexikális erőforrások (pl. WordNet kiterjesztés) és szakontológiák készültek. Továbbfejlesztésre került az MNSz elemzése során alkalmazott nyelvi elemzőlánc.

Befejeződött a *Magyar Generatív Történeti Szintaxis* OTKA-projekt, amelynek keretei között elkészült az Ómagyar Korpusz. A korpusz tartalmazza az összes ómagyar kódex (47 darab), 24 rövidebb ómagyar szöveg és 244 misszilis (elküldött levél) betűhű változatát, összesen több mint 2,2 millió szövegszót. 11 kódex és a rövidebb szövegek normalizált változata, valamint 4 kódex morfológiai elemzése is elkészült. A projekt eredményei és a korpuszkereső felület a <http://omagyarkorpusz.nytud.hu> oldalon érhető el.

Elkészült a *MATRICA* projektben az a technológiai lánc, amely interneten át lehetővé teszi a társadalomtudományi folyóiratok bibliográfiai adatainak nyelvtechnológiai feldolgozását és a feldolgozás szakaszainak hálózaton át történő kézi ellenőrzését. A rendszer az alábbi funkciókat nyújtja: (1) nyers bibliográfiai adatok félautomatikus kinyerése pdf formátumú cikkekből; (2) a kinyert hivatkozások statisztikai alapon történő automatikus elemzése és adatbázisba töltése; (3) az adatok felvitelét és a feldolgozás nyomkövetését, javítását több felhasználó számára párhuzamosan lehetővé tevő webes felület. A projekt célja a kutatókra és könyvtárosokra háruló, publikációs adatok feldolgozásával kapcsolatos adminisztrációs teher csökkentése, illetve esetleges későbbi scientometriai kutatásokhoz adatok biztosítása.

Az MVGYOSZ-szel együttműködésben elkészült a magyar *Braille-rövidírás* megújítására vonatkozó javaslat, mely az eddigi rendszert bővítve rövidítési képességét 30%-kal növeli. A javaslat a tömörség és a kényelmes használhatóság közötti helyes kompromisszum jegyében korpuszgyakoróságok alapján félautomatikus módon állt elő. Vakok általi tesztelést követően alkalmas a bevezetésre.

Idén kezdődött a *Finnugor nyelvű közösségek nyelvtechnológiai támogatása online tartalmak*

létrehozásában című OTKA-projekt (Finn-OTKA). Az egymásnak megfelelő, különböző nyelvű Wikipédia-cikkekből egyre bővülő kétnyelvű szótárakat készülnek különböző módszerekkel, annak érdekében, hogy a kis finnugor nyelvek online jelenléte erősödjön, e nyelvek funkció- és presztizsvesztését megakadályozzuk. A projekt keretein belül elkészült egy adatbázis és az azt előállító szoftvercsomag, amellyel összehasonlítható (comparable) korpuszokat lehet létrehozni a Wikipédiából kisebb finnugor (komi, mari, udmurt stb.) és az ún. "komfort zóna" nyelvek (angol, orosz stb.) között.

2013-ban megtörtént egy *szlovák-magyar* párhuzamos korpusz vázának kialakítása, a korpusz alapváltozatának elkészítése, valamint elkezdődtek a szlovák nyelv oktatását segítő Slovae.eu portál munkálatai.

A csoport a beérkező igényeknek megfelelően folyamatosan számítógépes nyelvészeti támogatást ad különböző nyelvészeti kutatásoknak, nyelvi adatokat szolgáltat, egyéni kéréseket teljesít főként a Magyar Nemzeti Szövegter és az ebből származtatott nyelvi erőforrások (pl. gyakorisági adatbázis) tekintetében.

Alkalmazott Nyelvészeti Kutatócsoport

Alkalmazott Nyelvészeti Kutatócsoport munkatársai előadásokat tartottak és tanulmányokat készítettek az alábbi témakörökben: (1) nyelvművelés, stilisztika, nyelvhelyességi minősítés; (2) helyesírás, nyelvi normativitás; (3) névtan, névkultúra, névjog, névviselet.

A nyelvi tanácsadó szolgálat a 2013-ban a korábbi éveknél több, kb. 4000 (nagyjából felerészben emailben (tanacs@nytud.hu) érkezett, illetve felerészben telefonos) helyesírási, nyelvhasználati kérdésre adott választ, 8 szövegértelmezési szakvéleményt készített. A partnerek között állami és európai uniós hivatalok éppúgy megtalálhatók, mint hazai és multinacionális cégek, szervezetek, oktatási intézmények, könyvkiadók vagy szerkesztők, fordítók, tanárok, újságírók, magánemberek. A szolgálat alkalmankénti nyelvi szakértői munkát végzett a Köztársasági Elnöki Hivatal részére; valamint emléktáblák, feliratok szakvéleményezésével foglalkozott (30 db) önkormányzatok és magánszemélyek számára.

Az Alkalmazott Nyelvészeti Kutatócsoport részt vesz a helyesiras.mta.hu Helyesírási tanácsadó portál működtetésében, folyamatosan választ ad a felmerülő kérdésekben, kétes esetekben. A csoport részt vesz az MTA Magyar Nyelvi Bizottsága Szűkebb Munkabizottságának A magyar helyesírás szabályai 12. kiadását előkészítő munkájában.

A kutatócsoport speciális feladatként ellátja a hatósági anyakönyvezésre beterjesztett új utónevek nyelvi szakvéleményezését 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. 2013-ban a lista 104 új névvel bővült (60 női, 44 férfi) így 2014. január 1-jén 3579 bejegyzésre alkalmasnak minősített utónév (2029 női és 1550 férfi) alkotja a jegyzéket. A minisztériumi felkérésre készülő többszáz hivatalos utónév-szakvélemény elkészítésén túl, a nyelvi tanácsadó szolgálat továbbra is választ ad az emailben (nevtanacs@nytud.hu) és telefonon érkező utónevekkel kapcsolatos kérdésekre.

Az osztály két kutatócsoportjának együttműködésében 2013. novemberében elindult az *Utónévportál* tesztverziója, szabadon elérhető a <http://corpus.nytud.hu/utonevportal> címen. Az utónevek iránt mutakozó állandó érdeklődés kielégítésére az intézet az említett havonta frissülő utónévlisákon túl idéntől ezen a kényelmes online felületen is közzéteszi az adható utóneveket számos releváns adatukkal együtt. A portál használója megismerheti egy-egy név jelentését, eredetét, becéző formáit, rokon neveit. Informálódhat a név névnapjáról, az adott név előfordulásának gyakoriságáról, felfedheti a különböző nevek közötti kapcsolatokat. Megalakult a Közigazgatási és Igazságügyi Minisztérium és az MTA Nyelvtudományi Intézet közös szervezésében a naptári bizottság, melynek tagja a két intézmény képviselőjén kívül egy néprajzos, egy nyelvész-néprajzos és egy egyháztörténész. A bizottság feladata a névnapok esetleges kijelölése

egy-egy névhez.

Elkészült a *Magyar Óvodai Nyelvi Korpusz (MONYEK)* anyagához a morfológiai annotációs séma a HuMor morfológiai elemző és a PurePos egyértelműsítő kombinálásával, a felnőtt írott nyelvre kidolgozott eszközök beszélt gyermeknyelvre történt adaptálásával. Ehhez a korpusz manuálisan annotált része szolgált alapul. Megkezdődtek a MONYEK anyagán alapuló diskurzusszintű vizsgálatok, valamint az eredmények publikálása.

Párbeszéd a tudomány és a társadalom között

Nyelvtechnológiai Kutatócsoport

A helyesiras.mta.hu Helyesírási tanácsadó portál azzal a szándékkal született, hogy új lehetőséget teremtsen egy széles közvéleményt gyakran foglalkoztató kérdéskör, a magyar helyesírás területein való tájékozódáshoz. A megnyitó óta eltelt idő arról tanúskodik, hogy a kezdeményezés beváltotta a hozzá fűzött reményeket. A portál nagy médiavisszhangot váltott ki, igen jelentős látogatottságra és elismertségre tett szert: az elmúlt nyolc hónap alatt több mint egymillió tanácsot adott, a visszatérő látogatók száma meghaladta a kétszázézetet. Az MTA Nyelvtudományi Intézet egy olyan tudományos alapokon álló szolgáltatást üzemeltet, mely valóban sokakat érdekel, sokakhoz el is jut, sikeresen működik. A portál közönségével való kapcsolattartás, a közönség bővítése érdekében, illetve azért, hogy lépést tarthassunk az aktuális igényekkel, elkészült a portál Facebook-oldala, ahol folyamatosan helyesírási érdekességekkel szolgálunk; 2013 októberében elindult a portál blogja, ahol heti két bejegyzés jelenik meg különböző rovatok alatt (A portálról, Hírek, ismertető, Nyelvészet és helyesírás, Közönségszolgálat, Visszajelzések, Helyesírási érdekességek); valamint ismeretterjesztő előadásokon mutattuk be a portált többek között a Magyar Tudomány Ünnepeén.

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 közel 9000 regisztrált felhasználóval bíró Magyar Nemzeti Szövegtár mellett kiemelendő a Magyar Generatív Történeti Szintaxis projekt keretében létrehozott Ómagyar Korpusz. Ez azon túl, hogy fontos és hiánypótló eszköz a nyelvtörténettel és/vagy az elméleti nyelvészet diakrón aspektusaival foglalkozó kutatók számára, kiválóan használható az oktatásban is: egyrészt az egész fennmaradt ómagyar szöveganyag kereshetővé vált, másrészt a normalizált szövegváltozatok lehetővé teszik, hogy a diákok mai magyarul is olvashassák a kódexeket.

Nyelvtechnológiai eszközök segítségével elkészült az új, magyar *Braille-rövidírásra* vonatkozó javaslat. Alkalmas vakok általi tesztelésre, bevezetésre. Ennek az alkalmazott kutatásnak az eredménye így akár már 2014-ben (azaz a projekt indulásától számított két éven belül) közvetlenül hasznosulhat egy fogyatékkal élő csoport, a vakok javára.

A *nyelvi veszélyeztetettség* kérdése és a nyelvi sokszínűség mint érték jelentős hangsúlyt kapott az utóbbi évek nemzetközi kutatásaiban. E vonulatba tartozó cél a rokon kis finnugor nyelvek megőrzése és fennmaradásuk elősegítése. A mai információ-alapú világban azon nyelvek beszélői, akik nyelve nincs képviselve az online felületeken, könnyen kimaradhatnak az új technológiákból, információkból. Ennek megakadályozására törekszik a csoport a Finnugor nyelvű közösségek nyelvtechnológiai támogatása online tartalmak létrehozásában című projektben.

A magyarul is elérhető *EMM NewsExplorer* nemzetközi integrációnkat segíti, a többnyelvű hírösszegző portál segítségével reális kép alakítható ki a magyar közletről a nemzetközi szinten; a szlovák nyelv oktatását segítő leendő Slovake.eu portál pedig a két ország kapcsolatának javításához járulhat hozzá hosszabb távon.

Alkalmazott Nyelvészeti Kutatócsoport

Az Alkalmazott Nyelvészeti Kutatócsoport tevékenysége 2013-ban is hozzájárult a társadalmat, az adófizetőket foglalkoztató nyelvi kérdések megválaszolásához, nyelvi problémák megoldásához. Közvetlenül a *nyelvi tanácsadás*, a nyelvi közönségszolgálat, közvetve pedig az egyéni nyelvi ismeretterjesztő tevékenység jóvoltából. Kiemelten a következő területeken: nyelvművelés, nyelvi tanácsadás, nyelvi ismeretterjesztés, a nyelvtudomány eredményeinek népszerűsítése.

A nyelvi tanácsadó szolgálat 2013-ban is eleget tett az Akadémiának a magyar nyelv gondozására, fejlesztésére, művelésére vállalt elkötelezettségének, hagyományának. A nyelvhasználat változásai, az új nyelvi jelenségek, a normákban bekövetkező változások (pl. ikes ragozás, új szavak, a média nyelve) és az újabb nyelvi alakulatok, átvételek értelmezése és írásmódja közérdeklődésre számot tartó területek voltak. A nyelvi tanácsadás lehetőséget teremtett az alapvető tudományos ismeretek átadása mellett a nyelvi babonák, sztereotípiák és a nyelvi stigmatizáció negatív hatásának csökkentésére. A szolgálat telefonon (kedden és csütörtökön 10--14 óráig) és elektronikus levélben (tanacs@nytud.hu illetve nevtanacs@nytud.hu címen) fogadja a kérdéseket, lehetővé téve a gyors válaszadáson kívül nyelvi produktumok véleményezését is. Az akadémiai helyesírás, az erre épülő helyesírási norma folyamatos gondozása fontos feladat és egyúttal társadalmi igény - különösen a nyilvánosság elé kerülő szövegek esetében. A nyelvi tanácsadó szolgálatba érkező kérdések jelentős része helyesírási vonatkozású.

A nyelvi, nyelvhasználati *ismeretterjesztés* különféle rendezvényeken illetve ismeretterjesztő kiadványok szerkesztésében valósul meg. A kutatócsoport munkatársai számos előadást tartottak különböző rendezvényeken, anyanyelvi versenyeken, lektoroknak, szakértőknek tartott továbbképzéseken az anyanyelvi nevelés, retorika, kommunikáció, helyesírás, nyelvi babonák témakörökben.

Törvényi kötelezettségének megfelelően a bejegyzésre alkalmasnak ítélt *utónevek* (férfi és női keresztnévek) 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: 2013-ban 104 új névvel (60 női, 44 férfi) egészült ki a lista. Ide tartoznak a névváltoztatással, a történelmi nevek cégbejegyzésével és a gazdasági reklámokkal és üzletfeliratokkal kapcsolatos feladatok is, valamint a részvétel az MTA Nyelvhelyességi Tanácsadó Testületének munkájában. A névadással, névkultúrával kapcsolatos számos interjú, tájékoztatás iránt folyamatosan nagy az érdeklődés a médiában.

A korábban is meglévő telefonos illetve emailes közönségszolgálatot egészíti ki 2013-tól a mai kor követelményeinek megfelelően *két új online felület*: egyrészt a már említett helyesiras.mta.hu, másrészt pedig a (tesztverzióban) szintén idén megnyílt *Utónévportál*. E portálon elérhető az adható utónevek teljes és folyamatosan frissített adatbázisa. A célközönség elsősorban a gyermeküknek nevet adni szándékozó szülők csoportja, ennek megfelelően a részletes keresőfelületen olyan szempontok jelennek meg, melyek a névadás során felmerülhetnek: a nevek formai tulajdonságaitól kezdve, a jelentésen, eredeten, névnapon stb. keresztül a konkrét aktuális adatokra támaszkodó gyakorisági keresésig. Az Utónévportál szabadon elérhető a <http://corpus.nytud.hu/utonevportal> címen.

A kutatóhely hazai és nemzetközi kapcsolatai 2013-ban

Hazai kapcsolatok

Nyelvtechnológiai Kutatócsoport

A Nyelvtechnológiai Kutatócsoport januárban "The Hungarian Language in the Digital Age -- A magyar nyelv helyzete a digitális korban" címmel nemzetközi konferenciát rendezett, amelyre köszöntőt küldött Dr. Áder János köztársasági elnök, és amelynek megnyitóján az MTA alelnöke és 4 minisztérium államtitkára is köszöntőt mondott. Az esemény egyben a CESAR projekt záróeseménye is volt.

Az osztály többször adott lehetőséget a Károli Gáspár Református Egyetem terminológus hallgatóinak szakmai gyakorlatuk elvégzésére.

Az osztály különböző projekteken együtt dolgozik a Pécsi Tudományegyetem Egyetemi Könyvtárával, ELTE PPK Addiktológiai Tanszéki Szakcsoportjával, az MVGYOSZ-szel, az Infinity Kft-vel és a GeoX Kft-vel.

A csoport tagjai idén részt vettek a Szegedi Tudományegyetem MTA-SzTE Mesterséges Intelligencia Kutatócsoport által szervezett Natural Language Processing for the Social Media kurzuson. A csoport egyik munkatársa rendszeresen publikál tudományos ismeretterjesztő írásokat helyesírás témában egy orvosoknak szóló szaklapban.

Pajzs Júlia: MTA Szótári Munkabizottság -- tag

Ludányi Zsófia: Nyelvhelyességi Tanácsadó Testület -- titkár

Alkalmazott Nyelvészeti Kutatócsoport

Heltainé Nagy Erzsébet: MTA Magyar Nyelvi Osztályközi Állandó Bizottság -- tag, MTA Magyar Nyelvészeti Munkabizottsága -- tag, MTA Nyelvhelyességi Tanácsadó Testület -- tag, Anyanyelvápolók Szövetsége -- választmányi tag, Magyar Nyelvőr -- a szerkesztőbizottság tagja, Édes Anyanyelvünk -- a szerkesztőbizottság tagja

Raátz Judit: MANYE -- tag, HUNRA Magyar Olvasástársaság -- tag, Magyar Nyelvtudományi Társaság -- tag, Magyar Nyelvtudományi Társaság Névtani Tagozat -- tag, Magyar Nyelvtudományi Társaság Magyarartanári Tagozat -- titkár, Szemere Gyula anyanyelv-pedagógiai kutatócsoport -- tag, Anyanyelvápolók Szövetsége -- választmányi tag, Magyarartanítás -- a szerkesztőbizottság tagja

Nemzetközi kapcsolatok

Nyelvtechnológiai Kutatócsoport

Az osztály különböző projektek keretében együtt dolgozik a következő intézményekkel: University of Helsinki Institute of Behavioural Sciences, DFKI Language Technology Lab, Instituut voor Nederlandse Lexicologie, EC Joint Reseach Centre, Szlovák Tudományos Akadémia Nyelvtudományi Intézete; valamint kapcsolatot tart fenn a Kínai Társadalomtudományi Intézet Nyelvtudományi Intézetével.

2013-ban elindult a Nyelvtudományi Intézet webshopja, mely a nemzetközi konferenciák szervezését is megkönnyíti.

Alkalmazott Nyelvészeti Kutatócsoport

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

Felsőoktatási tevékenység

Nyelvtechnológiai Kutatócsoport

Oravecz Csaba: Bevezetés a nyelvtechnológiába. Szeminárium. ELTE Elméleti Nyelvészet szak. 2013. tavasz.

Pintér Tibor: 6 különböző kurzus. Károli Gáspár Református Egyetem, 2013. tavasz.

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

Alkalmazott Nyelvészeti Kutatócsoport

Heltainé Nagy Erzsébet, további jogviszonyú egyetemi docens, KRE BTK Magyar Nyelvtudományi Tanszék: stilsztika, szöveg-és stílus elemzés, retorikai gyakorlatok, szóalkotás, terminusalkotás. 2013. tavasz: 4 kurzus; 2013. ősz: 9 kurzus.

Kardos Tamás: 2 kurzus. Károli Gáspár Református Egyetem, 2013. ősz.

Mátyus Kinga: 3 kurzus. ELTE Bárczi Gusztáv Gyógy pedagógiai Kar (BGGYK), 2013.

Raátz Judit, az ELTE Mai Magyar Nyelvi Tanszékének főállású docense: 16 kurzus oktatója, 2012. tavasz, 2012. ősz.

összes PhD opponensi feladat: 3; PhD bizottsági tagság: 1; PhD témavezetés: 6; MA/MSc témavezetés: 9; BA/BSc témavezetés: 3

Intézményi és pályázati együttműködések

Nyelvtechnológiai Kutatócsoport

intézeti kutató(k) neve	a pályázat címe	együttműködő			pályázati keret?
		intézmény neve	tanszék neve	vezető kutató neve	
Miháltz Márton	--	ELTE	PPK Addiktológiai Tanszéki Szakcsoport	Németh Dezső	--
Miháltz Márton	--	GeoX Kft.	--	Prajczer Tamás	--
Mittelholcz Iván	MATRICA	PTE	Egyetemi Könyvtár	Markó Tamás	MATRICA
Pajzs Júlia	EMM NewsExplorer	EC Joint Research Centre	--	Ralf Steinberger	--
Pintér Tibor	Szlovák-magyar párhuzamos korpusz	Szlovák Tudományos Akadémia Nyelvtudományi Intézete	--	--	--

Pintér Tibor	Slovake.eu	Szlovák Tudományos Akadémia Nyelvtudományi Intézete; Károli Gáspár Református Egyetem, BTK; Termini Kutatóhálózat	--	--	--
Sass Bálint	együttműködési megállapodás	MVGYOSZ	--	Péter Zsigmond	--
Simon Eszter	--	SZTAKI	Nyelvtechnológiai Kutatócsoport	Kornai András	Finn-OTKA
Simon Eszter	Finn-OTKA	University of Helsinki	Institute of Behavioural Sciences	Kristiina Jokinen	Finn-OTKA
Várad Tamás, Oravecz Csaba	Öko-innovációs piactér fejlesztése szemantikus audit alapon	Infinity Kft.	--	--	KMR_12-1-2012-0036, Piacorientált kutatás-fejlesztési tevékenység támogatása a közép-magyarországi régióban
Várad Tamás, Miháltz Márton	TrendMiner	Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI)	Language Technology Lab	Thierry Declerk	European Union (Grant agreement no: 287863)
Várad Tamás	ENEL: European Network of e-Lexicography	Instituut voor Nederlandse Lexicologie (INL)	--	Martin Everaert	ISCH COST Action IS1305
Várad Tamás	--	Nyelv és Tudomány -- nyest.hu	--	Kincse Szabolcs	--

Alkalmazott Nyelvészeti Kutatócsoport

intézeti kutató(k) neve	a pályázat címe	együttműködő			pályázati keret?
		intézmény neve	tanszék neve	vezető kutató neve	
Várad Tamás, Heltainé Nagy Erzsébet	terminológia MA	KRE	BTK	Fóris Ágota	--

Részvétel nemzetközi konferenciákon

Nyelvtechnológiai Kutatócsoport

előadó	előadás címe	konferencia elnevezése	helye (város)	ideje
Mátyus	MONYEEK -- Magyar Óvodai	The Hungarian Language in	Budapest	2013. január

Kinga	Nyelvi Korpusz	the Digital Age (CESAR roadshow)		18.
Oravecz Csaba	The Hungarian Gigaword Corpus	The Hungarian Language in the Digital Age (CESAR roadshow)	Budapest	2013. január 18.
Sass Bálint	Régi magyar konkordancia	The Hungarian Language in the Digital Age (CESAR roadshow)	Budapest	2013. január 18.
Váradi Tamás	Veni, vidi, vici, - The Language Technology Landscape after CESAR	SLOVKO 2013	Pozsony	2013. november 13.

Részvétel hazai konferenciákon

Nyelvtechnológiai Kutatócsoport

előadó	előadás címe	konferencia elnevezése	helye (város)	ideje
Dékány Éva, Simon Eszter	Az ómagyar -va/-ve és -ván/-vén igenevek használata véges igeiként	Nyelvelmélet és diakronia 2013	Budapest	2013. november 19-20.
Ehmann Bea, Lendvai Piroska, Miháltz Márton, Vincze Orsolya, László János	Szemantikus szerepek a narratív kategoriális elemzés (NARRCAT) rendszerében	MSZNY 2013	Szeged	2013. január 7-8.
Héja Enikő, Takács Dávid	Automatikusan generált online szótárak: az EFNILEX projekt eredményei	MSZNY 2013	Szeged	2013. január 7-8.
Ludányi Zsófia	Az orvosi helyesírás furcsaságai	4. Budapesti Amega Fórum (meghívott előadás)	Budapest	2013. szeptember 27-28.
Ludányi Zsófia	Egy intelligens portál -- helyesiras.mta.hu	Simonyi Zsigmond helyesírási verseny Kárpát-medencei döntője (meghívott előadás)	Budapest	2013. május 25.
Ludányi Zsófia, Miháltz Márton, Hussami Péter	Különírás-egybeírás -- automatikusan	VII. Alkalmazott Nyelvészeti Doktoranduszkonferencia	Budapest	2013. február 1.
Ludányi Zsófia, Miháltz Márton, M. Pintér Tibor, Takács Dávid	Helyesiras.hu -- az intelligens helyesíró portál	XXIII. MANYE Kongresszus	Budapest	2013. március 26-28.
Mátyus Kinga	Anonimizálási gyakorlat? - Egy magyar korpusz anonimizálásának tanulságai	MSZNY 2013	Szeged	2013. január 7-8.

Mátyus Kinga	Stílus a gyermeknyelvben	XXIII. MANYE Kongresszus	Budapest	2013. március 26-28.
Mátyus Kinga, Orosz György	MONYEEK -- Morfológiailag egyértelműsített óvodai nyelvi korpusz	Beszédkutatás Konferencia	Budapest	2013. november 14-15.
Miháltz Márton	OpinHuBank: szabadon hozzáférhető annotált korpusz magyar nyelvű véleményelemzéshez	MSZNY 2013	Szeged	2013. január 7-8.
Miháltz Márton, Hussami Péter, Ludányi Zsófia, Mittelholcz Iván, Nagy Ágoston, Oravecz Csaba, Pintér Tibor, Takács Dávid	Helyesiras.hu: Nyelvtechnológiai megoldások automatikus helyesírási tanácsadó rendszerben	MSZNY 2013	Szeged	2013. január 7-8.
Miháltz Márton, Ludányi Zsófia	helyesírás.mta.hu: automatizált helyesírási tanácsadás nyelvtechnológiai módszerekkel	Velünk élő nyelvtudomány, Magyar Tudomány Ünnepe	Budapest	2013. november 11.
Nemeskey Dávid, Simon Eszter	Automatikus korpuszépítés tulajdonnév-felismerés céljára	MSZNY 2013	Szeged	2013. január 7-8.
Pajzs Júlia	Az Európai Médiafigyelő magyar verziója	Velünk élő nyelvtudomány, Magyar Tudomány Ünnepe	Budapest	2013. november 11.
Pajzs Júlia	Szótár, lexikon, enciklopédia az on-line világban	Szótár, lexikon, enciklopédia	Budapest	2013. október 16.
Sass Bálint	Az új magyar Braille-rövidírás korpuszvezérelt kialakításának lehetőségei	MSZNY 2013	Szeged	2013. január 7-8.
Sass Bálint	Az új magyar Braille-rövidírás kialakítása	Velünk élő nyelvtudomány, Magyar Tudomány Ünnepe	Budapest	2013. november 11.
Szabó Tamás Péter, Mátyus Kinga	Óvodások és kisiskolások nyelvi ideológiái	XXIII. MANYE Kongresszus	Budapest	2013. március 26-28.
Váradi Tamás	A Magyar Nemzeti Szövegtár új változatáról	Velünk élő nyelvtudomány, Magyar Tudomány Ünnepe	Budapest	2013. november 11.

Alkalmazott Nyelvészeti Kutatócsoport

előadó	előadás címe	konferencia elnevezése	helye (város)	ideje
Heltainé Nagy Erzsébet	A helyesírás.mta.hu bemutatása	Nyelv-művészet-ismeretterjesztés	Budapest	2013. november 13.
Heltainé Nagy Erzsébet	Magyarul így! A nyelvész Gárdonyi Géza mai szemmel	Emlékkonferencia a 150 éve született Gárdonyi Géza és Bródy Sándor tiszteletére	Eger	2013. szeptember 26-28.
Oszkó Bea, Raátz Judit	Nyelvtörténeti ismeretek a gyakorlatban	Találkozások az anyanyelvi nevelésben 1.	Pécs	2013. április 22.
Raátz Judit	Globalizációs és identifikációs jelenségek a mai magyar keresztnevéadásban	XXIII. MANYE Kongresszus	Budapest	2013. március 26-28.
Raátz Judit	Szövegértés -- adatok és gyakorlatok	III. Vályi András Anyanyelv-pedagógiai konferencia	Budapest	2013. december 7.
Raátz Judit, Sass Bálint	Az Utónévportál bemutatása	Velünk élő nyelvtudomány, Magyar Tudomány Ünnepe	Budapest	2013. november 11.

Rendezett konferenciák

Nyelvtechnológiai Kutatócsoport

szervező	konferencia elnevezése	helye (város)	ideje	(társszervező, ha volt)
Nyelvtechnológiai és Alkalmazott Nyelvészeti Osztály	The Hungarian Language in the Digital Age -- A magyar nyelv helyzete a digitális korban (CESAR roadshow)	Budapest	2013. január 18.	BME-TMIT
Nyelvtechnológiai és Alkalmazott Nyelvészeti Osztály	VII. Alkalmazott Nyelvészeti Doktoranduszkonferencia	Budapest	2013. február 1.	Alkalmazott Nyelvészeti Munkabizottság

A 2013-ban elnyert fontosabb hazai és nemzetközi pályázatok rövid bemutatása

Nyelvtechnológiai Kutatócsoport

A 2013-ban elindult *Finnugor nyelvű közösségek nyelvtechnológiai támogatása online tartalmak létrehozásában* című projekt célja, hogy a kis finnugor nyelvek is képviseltessék magukat az online felületeken, ezzel akadályozva azok funkció- és presztizsvesztését, illetve segítve ezen közösségek megjelenését az internetes felületeken.

Az osztály sikerrel pályázott az Európai Unió (FP7) által finanszírozott, 2011. november óta futó, 3

éves *TrendMiner* elnevezésű nemzetközi projekthez való csatlakozásra. A projekt célja közösségi médiumokban (facebook, Twitter stb.) megjelenő, felhasználók által generált üzenetek valós idejű elemzése és monitorozása. Az osztály feladata magyar nyelvű, főként politikai témákban íródott üzenetek (kommentek) véleményelemzése és pszichológiai tartalomelemzése.

A Nyelvtudományi Intézet csatlakozott a 2013-ban induló *European Network of e-Lexicography* elnevezésű COST kezdeményezéshez, mely az elektronikus szótárkészítés területén valósít meg európai együttműködést.

pályázó vezető kutató neve	pályázat címe	pály. azonosító száma	szerződés szerinti kezdődátum	2013-ra eső támogatás	várható főbb eredmények
Várad Tamás	Finnugor nyelvű közösségek nyelvtechnológiai támogatása online tartalmak létrehozásában	FNN 107885	2013. szeptember 1.	3551 eFt	összehasonlítható korpuszok és automatikusan előállított szótárak kisebb finnugor nyelvekre, Wikipedia és egyéb online tartalomszerkesztő eszközök nyelvi támogatása
Várad Tamás	TrendMiner -- Large-scale, Cross-lingual Trend Mining and Summarisation of Real-time Media Streams	FP7 ICT-2011.4.2, 287863	2013. november 1.	2289 eFt	nagy tömegű szövegbányászat, érzelmi viszonyulás kimutatása pénzügyi, politikai, gyógyászati témájú szövegekben
Várad Tamás	ENEL: European Network of e-Lexicography	ISCH COST Action IS1305	2013. október 11.	--	európai együttműködés a modern szótárkészítés területén
--	www.slovake.eu - Extending the offer of the e-learning platform for the Slovak language	543135-LLP-1-2013-1-SK-KA2-KA2MP	2013. november 1.	15000 euró	a szlovák nyelv elsajátításának segítése magyar anyanyelvűek számára

A 2013-ban megjelent jelentősebb tudományos publikációk

Nyelvtechnológiai kutatócsoport:

Kovács Réka 2013. Helyesiras.mta.hu. Anyanyelv-pedagógia. 2013/3. Link: <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=474>

Ludányi Zsófia — Miháltz Márton — Hussami Péter 2013. Különírás-egybeírás -- automatikusan. In: Várad Tamás (szerk.) *AlkNyelvDok7. Doktoranduszok tanulmányai az alkalmazott nyelvészet*

köréből. Budapest: MTA Nyelvtudományi Intézet, 116–130. Link:
<http://www.nytud.hu/alknyelvdok13/proceedings13/ANyD7-Ludanyi-Zsofia-et-al.pdf>

Pintér Tibor 2013. Nyelvtechnológiai segédletek a magyarnyelv-oktatásban. In: Dringó-Horváth Ida (szerk.) Digitális tananyagok – oktatásinformatikai kompetencia a tanárképzésben. 88-90.

Prószéky Gábor, Miháltz Márton, Kuti Judit: Lexikális szemantika: a számítógépes nyelvészet és a pszicholingvisztika határán. In: Kenesei István, Pléh Csaba (szerk.): Általános Nyelvészeti Tanulmányok XXV. Akadémiai Kiadó, Budapest, 143--172, 2013.

Simon Eszter: Approaches to Hungarian Named Entity Recognition. PhD disszertáció. Budapesti Műszaki és Gazdaságtudományi Egyetem, 2013.
http://www.omikk.bme.hu/collections/phd/Termesztudomanyi_Kar/2013/Simon_Eszter/ertekezes.pdf

Steinberger, Ralf - Maud Ehrmann - Júlia Pajzs - Mohamed Ebrahim - Josef Steinberger - Marco Turchi

Multilingual Media Monitoring and Text Analysis – Challenges for Highly Inflected Languages. In: Habernal, Ivan; Matousek, Vaclav (Eds.): Text, Speech, and Dialogue Pilsen, Czech Republic, September 1-5, 2013. Proceedings, Lecture Notes in Computer Science, Volume 8082, Springer Berlin Heidelberg, 2013, p 22-33.
http://link.springer.com/chapter/10.1007%2F978-3-642-40585-3_3#page-1

Váradi, Tamás (2013): Veni Vidi, Vici: The Language Technology Infrastructure Landscape after CESAR. in Katarina Gajdsová & Adriána Zaková (szerk.) SLOVKO 2013. Natural Language Processing, Corpus Linguistics, E-Learning. Slovak National Corpus, L. Stur Institute of Linguistics, Slovak Academy of Sciences, Bratislava, Slovakia 2013, pp. 261-279

Alkalmazott nyelvészeti kutatócsoport:

Heltainé Nagy Erzsébet: Vélemény Albert Nagy Ákos beszédéről. In: Az egyházi retorika. Szerk. Raátz Judit és Tóth Etelka. ELTE Mai magyar nyelvi Tanszék, Budapest, 2013. 194–197.

Mátyus Kinga 2013. The (bVn) variable in the speech of five- and ten-year-old Hungarian children. In: Polczynska, M. et al. (szerk.) Young Linguists' Insights – Taking interdisciplinary approaches to the fore. Poznan: Uniwersytet im. Adama Mickiewicza w Poznaniu. 55-62.

Raátz Judit 2013. Társadalmi változás – névjog – névválasztás. In: Tóth Szergej (szerk.) Társadalmi változások – nyelvi változások. Alkalmazott nyelvészeti kutatások a Kárpát-medencében. MANYE, Vol. 9. MANYE – Szegedi Egyetemi Kiadó Juhász Gyula Felsőoktatási Kiadó. Budapest–Szeged. 396–403.