

# A Nyelvtechnológiai és Élőnyelvi Osztály jelentése a 2011. évről

## Kiemelkedő kutatási és más jellegű eredmények

A Nyelvtechnológiai Kutatócsoport koordinálásával és 9 vezető európai gépi fordító partner részvételével zajló kétéves *iTranslate4.eu* projektben jelentős eredmények születtek a 2011-es év során. Április folyamán megnyílt a [www.iTranslate4.eu](http://www.iTranslate4.eu) gépi fordító portál, mely a résztvevő partner cégek szolgáltatását integrálja innovatív módon: több mint 40 nyelvet kezel, és fontos európai nyelvpárok esetén - mint például az angol-német, a francia-olasz vagy az angol-magyar - versenytársainál jobb minőségű automatikus fordítást nyújt. A szöveg- és honlapfordítás mellett az internetes keresés, valamint a chat-üzenetek valós idejű fordítása is megoldott. A különféle kiegészítő termékek segítségével a portál funkcionalitását tetszőleges honlapba illetve különböző szoftvertermékekbe lehet beépíteni. A gépi fordító motor térítés ellenében programozói interfészen (API) keresztül is elérhető. Számos nyelvpárnál alternatív fordítások között lehet választani, ami megkönnyíti az eredeti szöveg megértését. A kutatócsoport kifejlesztett és publikált egy megbízható és soknyelvpáros környezetben is hatékony kiértékelő módszert, mely a fordítóportál működése során minőségi sorrendbe rendezi a fordítómotorok által szolgáltatott fordítási alternatívákat.

Sikeresen elindult a szintén EU által támogatott *CESAR* projekt. A projekt célja a közép- és délkelet-európai országok nyelvi erőforrásainak és eszközeinek számba vétele, egységesítése, összekapcsolása és egységes felületen való közzététele. A kutatócsoport összeállította az általa támogatott, magyar nyelvi erőforrásoknak és eszközöknek az első csoportját, melyeket a pályázat követelményeinek megfelelően dokumentált, annotált és metaadatokkal látott el, rögzítette a hozzáférésre vonatkozó licenzeket. A 2011 végén közzétett felhívás nyomán további eszközök és erőforrások bevonása várható a 2012-es év folyamán. A *CESAR* projekt keretei között elkészült a magyar nyelvről és nyelvtechnológiáról szóló fehér könyv *Nyelvek az európai információs társadalomban: magyar* címmel. Az osztály munkatársai a könyv szerzőiként és szerkesztőiként működtek közre. A META-NET által kiadásra kerülő fehér könyv sorozatnak célja, hogy felhívja a figyelmet a nyelvtechnológiára és az abban rejlő lehetőségekre. A kötet megjelenése 2012-re várható.

Teljes egészében elkészült a BUSZI-2 (Budapesti Szociolingvisztikai Interjú) adatbázis társalgási moduljainak az anonimizálása. Az 50 interjú több mint 240000 szavas hanganyaga és lejegyzése anonimizálásra került az adatvédelmi ombudsman rendelkezései alapján, megnyitva az utat az adatbázis szélesebb körben való kutathatóvá tétele előtt.

Eredményesen lezárult a Debreceni Egyetem által vezetett *Az Ember-gép kommunikáció technológiájának elméleti alapjai* című TÁMOP projektum, melynek keretében a kutatócsoport aktív közreműködésével elkészült egy itthon egyedülálló, többszintű annotációval ellátott multimodális adatbázis. Ennek alapján elemzés készült a kommunikatív aktusok automatikus felismeréséről gépi tanulási algoritmusok segítségével, elkészült a fordulóváltás detekciójához/predikciójához felhasználható multimodális markerek specifikációja, és egy szabály alapú multimodális osztályozó rendszerterve.

Az EFNILEX projekt keretében zajló az automatikus szótárkészítés lehetőségeit vizsgáló kutatásnak két fő iránya volt 2011-ben. Mindkét területen publikáció is született. Az egyik irány az egyszavas kifejezéseket tartalmazó magyar-litván, litván-magyar, francia-holland illetve holland-francia szótárak online közzététele volt. A munkálatok eredményeképpen előállt, <http://efnilex.efnil.org> címen elérhető újfajta szótárak (1) személyre szabhatóak, (2) a bennük szereplő tételek a valós

nyelvhasználatot tükrözik, a legtöbbet használt fordítások jelennek meg a először, (3) a természetes nyelvi példák segítik a fordítások értelmezését, (4) a forrásnyelvi és célnyelvi szó közötti szemantikai viszony is megjelenik. A másik irány a szótárakban szereplő tételek körének kiterjesztése az egyszavas kifejezésekről a többszavas igei szerkezetekre. A "Magyar igei szerkezetek" szótár alapját képező automatikus szerkezet-kinyerő eljárást alkalmazva holland-francia illetve francia-holland igeszerkezet-párokat nyertünk ki teljesen automatikusan párhuzamos korpuszt felhasználva. Eredményként olyan összetett igei szerkezeteket kapunk, mint a francia *mettre accent sur* vagy holland megfelelője, a *leggen nadruk op* (magyarul mindkettő: *hangsúlyt helyez vmire*).

A Kutatócsoport munkatársai aktívan közreműködtek a Magyar Generatív Történeti Szintaxis nevű OTKA projektben. A projekt keretében 2011-ben is folytatódott az ómagyar szövegmélekek digitalizálása. Az adatbázis újabb 27 kódex betűhű szövegével bővült, ezek közül 3-nak a normalizálása is elkészült. Jelenleg a korpuszlekérdező felületen (<http://corpus.nytud.hu/rmk>) keresztül 31 kódex és 18 rövidebb nyelvemlék, összesen több mint 700000 szónyi betűhű szövege érhető el. Folyik a mai magyar morfológiai elemző ómagyar szövegekre való adaptációja, és elkezdődött a középmagyar anyagok összegyűjtése.

Egy néhány hónapos kutatási projekt keretében áttekintést végeztünk a jelenleg rendelkezésre álló magyar nyelvű főnévcsoport-felismerőkről. A főnévi csoportok (NP-k) határainak felismerése a magyar szövegek mondataiban fontos állomás a komplexebb nyelvtechnológiai elemzési feladatok számára. A felmérés során összehasonlítottuk (1) az Intézetben készült, NooJ keretrendszerben megvalósított magyar mondatelemzőt, (2) a MorphoLogic Kft. Metamorpho magyar-angol fordítóprogramjának magyar mondatelemzőjét, valamint (3) a hunchunk statisztikai gépi tanulással működő elemzőt. Az eszközök kiértékeléséhez a Szeged Treebank 2.0 mondatszerkezeti elemzéseket tartalmazó adatbázist felhasználva meghatároztuk a programoknak a teljesítményük alapján felállított sorrendjét.

A CLARIN projektum keretében a társadalomtudományokkal (pszichológia, filozófia) foglalkozó szakemberek számára tanácsadás és nyelvi feldolgozómodulok fejlesztése valósult meg.

Az osztály közreműködött az AMICUS nemzetközi projekt munkacsoportjában magyar és német nyelvű népmesék és "kanonikus" tematikus indexek nyelvi elemzésében (NooJ eszközzel), szabványos és interoperábilis reprezentációjában, és a szemantikus technológia elveivel megegyező feldolgozásában.

Elkészült egy beszélt gyermeknyelvi adatbázis kutatási terve és módszertana, melynek célja 4 éves budapesti gyermekek beszédének vizsgálata szociolingvisztikai szempontból.

Megtörtént a MetaMorpho elemző magyar igei kereteinek a NooJ rendszerben felhasználható formájúvá alakítása.

A Szegedi Egyetemmel együttműködve a kutatócsoport részletesen elemzett magyar nyelvi adatbázist fejlesztett ki a mannheimi Institut für Deutsche Sprache által koordinált Eurogr@mm projekt számára.

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

Nyelvművelő és Nyelv Tanácsadó Kutatócsoport munkatársai előadásokat tartottak és tanulmányokat készítettek az alábbi témakörökben: (1) a nyelvhelyességi minősítési rendszerek,

ezen belül a helyes-helytelen kérdéskör története, illetve ennek kapcsolata a nyelvi tanácsadói gyakorlattal; (2) stilisztika; (3) szövegértés, anyanyelvi nevelés, retorika; (4) névtan, (5) helyesírás.

A nyelvi tanácsadó szolgálat a 2011-es évben kb. 1700 emailben (tanacs@nytud.hu) érkezett, kb. 2000 telefonos és 20 postai úton érkezett helyesírási, nyelvhasználati kérdésre adott választ. A partnerek között állami hivatalok és gazdálkodó szervezetek éppúgy megtalálhatók, mint hazai és multinacionális cégek, könyvkiadók vagy szerkesztők, fordítók, tanárok, újságírók, magánemberek. A szolgálat 2011 tavaszán a Magyar Országgyűlés Hivatala felkérése nyomán szakvéleményezte a Magyarország Alaptörvényéről szóló T/2627. számú törvényjavaslatot, az ehhez kapcsolódó módosító javaslatokat és az egységes javaslatot nyelvi, stilisztikai, nyelvhelyességi, helyesírási szempontból. A szolgálat foglalkozott minisztériumi illetve akadémiai felkérésre szövegek lektorálásával, a történelmi nevek cégbíróági bejegyezhetőségével kapcsolatos szakvélemények készítésével a 2006. évi V. tv. 4.§, 5. bek. alapján, valamint emléktáblák, feliratok szakvéleményezésével (25 db) önkormányzatok részére.

A Nyelvtechnológiai és a Nyelvművelő és Nyelv Tanácsadó Kutatócsoport közösen munkálkodik a helyesiras.hu Nyelvi Tanácsadó Portál létrejöttén, mely egy automatikus és interaktív helyesírási tanácsadó programnak fog helyet adni. Elkészült a portál általános rendszerterve, és megkezdődött az egyes kezelőmodulok, nyelvi elemzők programkódjának a kifejlesztése. Elindult a már működő célmodulok, szótárak testüzeme. A kutatócsoport részt vesz az MTA Magyar Nyelvi Bizottságának a Magyar Helyesírás Szabályai 11. kiadását felülvizsgáló, a 12. kiadást előkészítő munkájában.

A kutatócsoport munkájához tartozik a minisztériumi felkérésre készülő hivatalos szakvélemények (480 db) elkészítésén túl, a nyelvi tanácsadó szolgálathoz a nevtanacs@nytud.hu emailcímre (810 db) valamint telefonon (kb. 250 db) érkező utónevekkel kapcsolatos kérdések megválaszolása, ezen kívül magyarigazolványok kiadásához a névátírásokról szóló szakértői vélemények elkészítése (10 db) is.

## **Párbeszéd a tudomány és a társadalom között**

Egyértelműen a nagyközönség, a lehető legszélesebb felhasználói kör számára készült az iTranslate4.eu projekt fejlesztéseinek eredményeként megnyílt gépi fordítási portál (www.itranslate4.eu), amely az EU összes hivatalos nyelve között ingyenesen biztosít magas minőségű fordítást. A portál szolgáltatásaira (automatikus szöveg- és honlapfordítás, lefordított keresés, chat-fordítás) jelentős társadalmi igény mutatkozik.

Az osztály munkatársai aktív szerepet vállaltak a tudománynépszerűsítést célzó *Kutatók éjszakája* rendezvény 2011-es programjának megszervezésében és létrehozásában. A *Nyelvében él a gép* elnevezésű programcsomagban különböző nyelv- és beszédtechnológiai alkalmazásokat mutattak be az érdeklődő közönségnek. A Nyelv és Tudomány és az intézet közös standján elsősorban nyelvész diákok beszéltek kutatásaikról. Egyéb standokon bemutatkozott a SZTAKI, az ELTE Fonetikai Tanszék, a BME TMIT, a Debreceni Egyetem, a MorphoLogic Kft. és a Nyelvtudományi Intézet munkatársai (CESAR projekt és EFNILEX projekt). A nyelvtechnológia népszerűsítését szolgáló előadások hangzottak el a nagyközönség részére szervezett egyéb rendezvényeken is (például a *Magyar tudomány napja* rendezvénysorozatban). Az eredmények ismertetése a felsőoktatásban is helyet kap a Károli Gáspár Református Egyetem magyar terminológia mesterképzésében.

Az osztály rendezte júniusban az európai szinten is jelentős, az egész földrész nyelvtechnológus közösségét megmozgató META-FORUM konferenciát, melynek fő célja a nyelvtechnológia szándékainak és trendjeinek pozicionálása volt az európai szinten. A META-NET által kiadásra kerülő fehér könyv sorozat magyar tagjának is az a célja, hogy felhívja a figyelmet a

nyelvtudományra és az abban rejlő lehetőségekre.

Az osztály egy munkatársa rendszeresen publikál a Nyelv és Tudomány (<http://nyest.hu>) portálon. Tudományos ismeretterjesztő cikkekben mutatja be a nyelv- és beszédtechnológia fejlődési irányait, aktuális kutatási témáit. Az osztály egy munkatársa szakszerkesztőként részt vett a Mindentudás Egyeteme 2.0 nyelvtechnológiai témájú előadásának, magazinműsorának és kerekasztal-beszélgetésének szervezésében. A magazinműsorban és a kerekasztal-beszélgetés során szakértőként megszólaltak a Nyelvtudományi Intézet munkatársai is. Könyvbemutató keretében a szerzőtársak közösen bemutatták a nagyközönségnek a 2010-ben megjelent *Magyar igei szerkezetek* szótárt júniusban, az MTA Nyelvtudományi Intézetben.

Hiánypótló az EFNILEX projekt keretében folytatott kutatás, mely olyan eszközök, szótárak költséghatékony kifejlesztését célozza, amelyek a kevésbé használt nyelvek elsajátítását és használatát is lehetővé teszik a magyar anyanyelvi beszélők számára. Ez a kezdeményezés megfelel az EU-nak az európa nyelvi sokszínűséget támogató illetve az idegennyelv-ismeret népszerűsítő irányelvének. Az idegennyelv-ismeret éppen az olyan nyelvileg egyedülálló ország esetében a legfontosabb, mint Magyarország.

Az osztályon fejlesztett és készülő nyelvi adatbázisok, többek között az anyanyelv iránti általános érdeklődés következtében is, széles körű társadalmi figyelemre tartanak számot, 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 6000 regisztrált felhasználóval bíró Magyar Nemzeti Szövegtár mellett kiemelendő a Magyar Generatív Történeti Szintaxis projekt keretében készülő, a teljes ómagyar nyelvemléki anyagot tartalmazó nyelvi adatbázis, mely már mostani készültségi állapotában (több mint 700000 szót tartalmaz) is egyedülállóan jelentős egységes erőforrás az ómagyar nyelvi örökség megőrzésében.

A Nyelvművelő és Nyelvi Tanácsadó Kutatócsoport tevékenységkörei hozzájárultak a társadalmat, a laikusokat foglalkoztató nyelvi kérdések megválaszolásához. A csoport tevékenységköre jelentős mértékben a tudomány és a társadalom közti párbeszédet is szolgálta, a nyelvi tanácsadás, nyelvi ismeretterjesztés, a helyesírás, az utónév-szakvéleményezés, a névválasztással, névdivattal kapcsolatos ismeretterjesztés területén.

A nyelvi tanácsadás a nyelvhasználók folyamatos igényéből adódóan működik. Tanácsadó szolgálatunk iránti növekvő igény mutatja, hogy a társadalomban ma is fontos az Akadémiának a magyar nyelv gondozására, fejlesztésére, művelésére tett elkötelezettsége, tekintélye. A nyelvhasználat változásai, az új nyelvi jelenségek, a normákban bekövetkező változások (pl. ikes ragozás, új szavak, szleng, nyelvi változatok, a média nyelve, a helyesírás aktuális helyzete stb.) közérdeklődésre számot tartó témák. A tanácsadói gyakorlatban lehetőség volt ezeken a területeken az alapvető tudományos ismeretek átadására, a nyelvi babonák, sztereotípiák, a nyelvi stigmatizáció negatív hatásának csökkentésére. A szolgálat telefonon (kedden és csütörtökön 10--14 óráig) és elektronikus levélben ([tanacs@nytud.hu](mailto:tanacs@nytud.hu) illetve [nevtanacs@nytud.hu](mailto:nevtanacs@nytud.hu) címen) fogadja a kérdéseket. Az akadémiai helyesírás, az erre épülő helyesírási norma folyamatos gondozása ugyancsak jelentős 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 a médiában valósul meg. A kutatócsoport munkatársai egy tucat előadást tartottak a *Magyar nyelv hete* rendezvényein, anyanyelvi versenyeken, szakmai napokon az anyanyelvi nevelés, retorika, kommunikáció, helyesírás, névhasználat, nyelvi változások, neologizmusok, nyelvi babonák témakörökben.

A helyesírás.hu Nyelvi Tanácsadó Portál a helyesírás és nyelvhelyesség kérdéseivel foglalkozik, azonban elsősorban a tág nyilvánosság számára készül. Szűkebb értelemben a nyelvvel foglalkozók, nyelv iránt érdeklődők, tágabb értelemben minden felhasználó, akinek nyelvi-nyelvhelyességi segítségre, eligazításra van szüksége (például egy hivatalos levél megfogalmazásakor). Célzottabb terület az oktatás, hivatalos ügyintézés írásos területei (pl. bankok és egyéb cégek), valamint a magyar nyelvvel hivatásszerűen foglalkozók megszólítása (pl. újságírók, tanárok, hivatali alkalmazottak).

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: 2011-ben 125 új névvel (73 női, 52 férfi) egészült ki a lista. Ide tartoznak a névváltoztatással, a történelmi nevek cégbejegyzésével és a gazdasági reklámokkal és üzletfeliratokkal kapcsolatos feladatok is. A névadással, névdivattal kapcsolatos évi akár 10-20 interjú, tájékoztatás iránt folyamatosan nagy az érdeklődés a nyomtatott és az online médiában egyaránt.

## **A kutatóhely hazai és nemzetközi kapcsolatai 2011-ben**

### **Hazai kapcsolatok**

A kutatócsoport két munkatársa a Debreceni Tudományegyetem vezetésével működő "Az ember-gép kommunikáció technológiájának elméleti alapjai" című TÁMOP 4.2.2. támogatási konstrukcióban létrejött "innovatív kutatói team" tagjaként aktívan közreműködött a projekt munkájában a rendszeres szakmai megbeszéléseken, előadásokon, konferenciákon és hozzájárult a projekt eredményes lezárásához.

A kutatócsoport munkatársai részt vettek és előadtak a VIII. Magyar Számítógépes Nyelvészeti Konferencián.

Egy munkatárs oktat a Károli Gáspár Egyetem Bölcsészettudományi Karán működő Magyar Nyelvtudományi Tanszéken, illetve részt vesz Magyar Terminológia mesterképzés akkreditációjában.

A Nyelvművelő és Nyelvi Tanácsadó kutatócsoport egy munkatársa az Édes Anyanyelvünk című nyelvművelő, ismeretterjesztő folyóirat szerkesztőbizottságának tagja; rovatvezető a Magyartanítás című folyóiratban; valamint a 2011-ben újjáalakult MTA Nyelvhelyességi Tanácsadó Testületének tagja.

Szakmai kapcsolat alakult ki a kutatócsoport és a Nemzeti Erőforrás Minisztérium között. A NEFMI Kultúrpolitikai Államtitkársága megkeresésére a kutatócsoport részt vesz az országgyűlési határozathoz kapcsolódó magyar nyelvi helyzetkép és intézkedési terv összeállításában: a nyelvművelő mozgalmak illetve az állam és a magyar nyelv témakörében. Szintén a NEFMI felkérésére a NAT magyar nyelvi követelményeinek kidolgozásában, valamint a kerettanterv szakértői tanácsában is részt vesznek munkatársak.

### **Nemzetközi kapcsolatok**

A Nyelvtechnológiai Kutatócsoport aktív szerepet játszott a Kínai Társadalomtudományi Intézet Nyelvtudományi Intézetével való kapcsolatépítésben. A kutatócsoport előadássorozattal mutatta be

fontosabb tevékenységeit és eredményeit a Kínai Nyelvtudományi Intézet igazgatóhelyettese, Prof. Liu Danqing intézeti látogatásán, 2011. február 8-án. Várad Tamás osztályvezető szeptember 19 és október 1 között látogatást tett a kínai fél meghívására, és előadást tartott valamint tárgyalást folytatott a két Intézet közötti együttműködésről. A tárgyalások eredményeképpen az év végére megszületett az együttműködési megállapodás tervezete.

A kutatócsoport és a Helsinki Egyetem közös szervezésében az intézet adott otthont a CLARA (Common Language Resources and their Applications) projekt *Thematic Training Course on Processing Morphologically Rich Languages* intenzív kurzusának, ahol öt hazai és egy külföldi előadó tartott előadásokat a főként külföldi diákokból álló hallgatóságnak (<http://www.ling.helsinki.fi/events/CLARA-FSM/>).

Az AMICUS projekt tematikájának kapcsán az osztály munkatársai közös kutatásokat folytattak a Boras-i Egyetemen (Swedish School of Library and Information Studies, University of Boras, Svédország) népmesei motívumok automatikus azonosításában, a DFKI-val (Deutsche Forschungszentrum für Künstliche Intelligenz, Saarbrücken, Németország) és az Osztrák Tudományos Akadémia Korpusznyelvészeti és Szövegtechnológiai Intézetével (Institut für Corpuslinguistik und Texttechnologie, ICLTT) kulturális örökségi szövegek szemantikus feldolgozásában és reprezentációjában.

A kutatócsoport 3 munkatársa vett részt az első ízben megrendezett lisszaboni Machine Learning School nyári egyetemen (2011. július 20-25.).

A CESAR pályázat révén szoros kapcsolat alakult ki a németországi székhelyű META-NET szövetséggel, melynek az Intézet tagja lett, Várad Tamást pedig a META-NET Vezetőségi Bizottságának tagjává választották. Az osztály helyi szervezője volt a magyar EU-elnökség hivatalos programjaként Budapesten megrendezett META-FORUM 2011 konferenciának, amelynek 250 résztvevője között több európai politikus és az európai nyelvtechnológiai szakma élvonalának java képviselte magát.

Folytatódott az osztály és a Szlovák Tudományos Akadémia pozsonyi Ludovít Štúr Nyelvtudományi Intézete, valamint az osztály és a szlovákiai Gramma Nyelvi Iroda közti együttműködés. Utóbbi keretében elkészült egy, a szlovákiai magyar nyelvet érintő (*A magyar nyelv szlovákiai változatainak adatbázisai* című) jelentős tanulmány.

Egy munkatárs a csehországi *Studies in Applied Linguistics* és *Romano Dzaniben* nevű folyóiratok szerkesztőbizottsági tagja lett.

Az osztály fogadta Anyiszima Darja vendégkutatót (Orosz Tudományos Akadémia), aki a magyar és szlovák nyelv grammatikájának kontrasztív vizsgálatával foglalkozik.

## **Felsőoktatási tevékenység**

Heltainé Nagy Erzsébet, további jogviszonyú egyetemi docens: stilisztika, szöveg-és stíluselemzés, írói nyelvbölcseletek, retorikai gyakorlatok, KRE BTK Magyar Nyelvtudományi Tanszék

Kálmán László, Oravec Csaba, Vásárhelyi Dániel: Bevezetés a számítógépes nyelvészetbe I-II. ELTE Elméleti nyelvészet szak. 2011. tavasz, 2011. ős.

Kardos Tamás: óraadás ELTE BTK Mai Magyar Nyelvtudományi Tanszéké, 2011. tavasz

Mátyus Kinga: Általános nyelvészet 2010/2011 első félév, Szociolingvisztika 2010/2011 második félév, ELTE Bárczi Gusztáv Gyógypedagógiai Kar (BGGYK)

Mittelholcz Iván: Logika és érveléstechnika, ELTE TáTK Közgazdaságtudományi Tanszék

Pajzs Júlia: előadás a korpusznyelvészeti kurzus hallgatóinak *Korpuszok és szótárak* címmel 2011. áprilisában valamint *Szótár és nyelvtechnológia* címmel 2011. novemberében, ELTE BTK.

Pintér Tibor: A természettudományok, a társadalom és a nyelvészet összefüggései - rendszertani alapok. Károli Gáspár Református Egyetem, Magyar Nyelvtudomány Tanszék, Terminológia mesterképzés. 2011. ősz.

Raátz Judit, ELTE Mai Magyar Nyelvi Tanszékének főállású docense: számos elméleti és gyakorlati kurzus oktatója, BA, MA és doktori szinten

összes PhD opponensi feladat: 4; PhD bizottsági tagság: 3; PhD témavezetés: 7; MA témavezetés: 9; BA témavezetés: 4

## Intézményi és pályázati együttműködések

intézeti kutató(k) neve	együttműködő			pályázati keret?
	intézmény neve	tanszék neve	vezető kutató neve	
Várad Tamás	Kinai Társadalomtudományi Intézet Nyelvtudományi Intézete		Liu Danqing	
Várad Tamás, Oravecz Csaba	Debreceni Tudományegyetem	Általános és Alkalmazott Nyelvészeti Tanszék	Hunyadi László	TÁMOP 4.2.2.
Pintér Tibor	Gramma		Szabómihály Gizella	
Pintér Tibor	Studies in Applied linguistics		Petr Kaderka	
Pintér Tibor	Romano Dzaniben		Eva Lehecková	
Pintér Tibor	Szlovák Tudományos Akadémia, Nyelvtudományi Intézet		Slavomír Ondrejovic	
Lendvai Piroska	University College Boras (Svédország)	Swedish School of Library and Information Science	Darányi, Sándor	AMICUS (Hollandia)
Lendvai Piroska	DFKI (Németország)	Language Technology Lab	Declerck, Thierry	AMICUS (Hollandia)
Lendvai Piroska	Institut für Corpuslinguistik und Texttechnologie (Ausztria)	ICLTT	Budín, Gerhard	AMICUS (Hollandia)
Simon Eszter	SZTAKI	Nyelvtechnológiai Kutatócsoport	Kornai András	CESAR
Héja Enikő,	EFNIL		František Čermák,	EFNILEX

Várad Tamás			Johan Van Hoorde, John Simpson, Jolanta Zabarskaite	
Várad Tamás, Heltainé Nagy Erzsébet, Pintér Tibor	KRE BTK	Magyar Nyelvtudományi Tanszék	Fóris Ágota	

### Részvétel nemzetközi konferenciákon

előadó	előadás címe	konferencia elnevezése	helye (város)	ideje
Lendvai Piroska	Annotation of Proppian Categories in Folk Tales	NooJ konferencia 2011	Dubrovnik (Horvátország)	2011. június 13-15.
Lendvai Piroska	Lingusitic and Semantic Enrichment of Stith Thompson's Motif Index of Folk Literature	AMICUS Workshop	Nijmegen (Hollandia)	2011. október 20.
Oravecz Csaba	Feature selection and feature representation for multimodal end-of-turn detection and prediction	The Theoretical Fundamentals of Human-Computer Interaction Technologies	Debrecen	2011. május 25-27.
Oravecz Csaba	Dialogue act identification in the HuComTech multimodal corpus	The Theoretical Fundamentals of Human-Computer Interaction Technologies	Debrecen	2011. október 9-11.
Pintér Tibor	Viacjazykovosť cez jazykovú technológiu. Výskumy jazykovej technológie v Jazykovednom ústave Maďarskej Akadémie Vied	Jazyk a diskurz v kultúrnom a politickom kontexte	Smolenice (Szlovákia)	2011. szeptember 21-23.

### Részvétel hazai konferenciákon

előadó	előadás címe	konferencia elnevezése	helye (város)	ideje
Héja Enikő	Igei bővítménykeretek fordítási ekvivalenseinek kinyerése mélyen elemzett párhuzamos korpuszból	VIII. Magyar Számítógépes Nyelvészeti Konferencia	Szeged	2011. december 1-2.
Heltainé Nagy Erzsébet	A népi konceptualizáció mint stíluskohéziós eszköz a sinkai költészetben	A stíluskohézió eszközei a modern és a posztmodern szövegekben	Debrecen, DAB (Debreceni Akadémiai Bizottság)	2011. március 9.



Heltainé Nagy Erzsébet	A Szenci-zsoltárok revíziójának nyelvi kérdései (felkért hozzászóló)	Magyarországi Református Egyház Doktorok Kollégiuma	Budapest	2011. november 4.
Heltainé Nagy Erzsébet	Tények és mítoszok a nyelvi értéktulajdonításokban	Implom József Helyesírási Verseny Kárpát-medencei döntő. Szakmai Nap	Gyula	2011. február 24-26.
Lendvai Piroska	Szemantikus szerepek vizsgálata magyar nyelvű szövegek narratív pszichológiai elemzésében	VIII. Magyar Számítógépes Nyelvészeti konferencia	Szeged	2011. december 1-2.
Miháltz Márton	Magyar NP-felismerők összehasonlítása	VIII. Magyar Számítógépes Nyelvészeti Konferencia	Szeged	2011. december 1-2.
Oravecz Csaba, Sass Bálint, Tihanyi László	Soknyelvpáros gépi fordítás hatékony és megbízható kiértékelése	VIII. Magyar Számítógépes Nyelvészeti Konferencia	Szeged	2011. december 1-2.
Oravecz Csaba, Váradai Tamás	Multimodal end-of-turn detection using machine learning methods	HuComTech Workshop	Síkfőkút	2011. március 25.
Pintér Tibor	Korpuszmunkálatok	A Termini Magyar Nyelvi Kutatóhálózat 10 éve	Budapest	2011. november 24.
Raátz Judit	Az olvasás össztantárgyi feladat	HUNRA 2011	Szombathely	2011. május 13-14.
Sass Bálint	Primáciahatás fordítások értékelésekor	XIX. MAKOG Konferencia	Kaposvár	2011. január 27-29.
Simon Eszter	Korpuszépítés ómagyar kódexekből	VIII. Magyar Számítógépes Nyelvészeti konferencia	Szeged	2011. december 1-2.

## Rendezett konferenciák

szervező	konferencia elnevezése	helye (város)	ideje	(társszervező, ha volt)
Nyelvtechnológiai Kutatócsoport	META-FORUM 2011 - Solutions for Multilingual Europe	Budapest	2011. június 27-28.	META-NET (Németország)
Nyelvtechnológiai Kutatócsoport	CLARA Thematic Training Course on Processing	Budapest	2011. április 11-	Helsinki Egyetem

	Morphologically Rich Languages		15.	
Lendvai Piroska	Fifth ACL-HLT Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities (LaTeCH-2011)	Portland (Oregon, USA)	2011. június 24.	
Lendvai Piroska	2nd AMICUS Workshop	Nijmegen (Hollandia)	2011. október 20.	
Alkalmazott Nyelvészeti Munkabizottság és MTA Nyelvtudományi Intézet, Mátyus Kinga, Váradi Tamás	V. Alkalmazott Nyelvészeti Doktoranduszkonferencia	Budapest	2011. február 4.	
ELTE BTK Magyar Nyelvtudományi Tanszék, Raázt Judit	13. Országos Kossuth Lajos Szónokverseny és Retorikai Konferencia	Budapest	2011. november 18-19.	

## A 2011-ben elnyert fontosabb hazai és nemzetközi pályázatok rövid bemutatása

Az Osztály a Finnugor és Nyelvtörténeti Osztállyal közösen közreműködik a 2011 októberében indult, EU-s támogatású INNET (Innovative Networking in Infrastructure for Endangered Languages) projektben.

## A 2011-ben megjelent jelentősebb tudományos publikációk

- Declerck, T. and P. Lendvai 2011. Linguistic and Semantic Representation of the Thompson's Motif-Index of Folk-Literature. In: S. Gradmann, F. Borri, C. Meghini, H. Schuldt (eds.) Research and Advanced Technology for Digital Libraries - International Conference on Theory and Practice of Digital Libraries. Lecture Notes in Computer Science number 6966, Berlin, Germany, Springer, 9/2011
- Heltainé Nagy Erzsébet 2011. A népi konceptualizáció mint stíluskohéziós eszköz a sinkai költészetben. In: A stíluskohézió eszközei a modern és a posztmodern szövegekben. Szerk. Szikszainé Nagy Irma. Debreceni Egyetemi Kiadó, Debrecen, 63--79.
- Miháltz Márton 2011. Magyar NP-felismerők összehasonlítása. VIII. Magyar Számítógépes Nyelvészeti Konferencia kiadványa. 333-335.
- Pintér Tibor 2011. A magyar nyelv főbb lexikonjai és határon túli vonatkozásuk. In: Szoták Szilvia (szerk.) Magyar nyelv és kultúra a Kárpát-medencében. Gramma Nyelvi Iroda-UMIZ-Imre Samu Nyelvi Intézet: Dunaszerdahely-Alsóór, 33-42.
- Raázt Judit 2011. Nick as self attributed name. In: Nouvelle Revue d'Onomastique. 53: 181-207.
- Sass Bálint 2011. Igei szerkezetek gyakorisági szótára -- egy automatikus lexikai kinyerő eljárás és alkalmazása. Ph.D. dolgozat, PPKE ITK.

- Simon Eszter, Sass Bálint, Mittelholcz Iván 2011. Korpuszépítés ómagyar kódexekből. In: Tanács Attila, Vincze Veronika (szerk.): VIII. Magyar Számítógépes Nyelvészeti Konferencia. SzTE, Szeged, 81-89.