

KUTATÓKÖZPONT NEVE:
NYELVTUDOMÁNYI KUTATÓKÖZPONT
Cím:
1068 Budapest, Benczúr u. 33.
Postafiók címe:
1394 Budapest, Pf. 360
telefonszám:
(1) 342-9372
honlap:
www.nytud.hu
e-mail:
linginst@nytud.hu

I. Közvetlenül a kutatóközpontokhoz tartozó kutatási egység(ek) kutatási eredményei a tématerületek megjelölésével

n/a

II. Tudományszervezési eredmények, események

1. A kutatóhely rövid bemutatása, főbb feladatai

A Nyelvtudományi Kutatóközpont (illetve az őt alkotó négy intézet: az Általános és Magyar Nyelvészeti Intézet, a Lexikológiai Intézet, a Nyelvtechnológiai és Alkalmazott Nyelvészeti Intézet, valamint a Történeti Nyelvészeti és Uralisztikai Intézet) alapfeladata szerint a magyar, az uráli, az általános és alkalmazott nyelvészet, a történeti nyelvészet, továbbá a lexikológia, a fonetika és a pragmatika területén nemzetközi színvonalú felfedező kutatásokat végez, a pszicho- és neurolingvisztika területén orvosi-népegészségügyi relevanciával is bíró kísérletes kutató és fejlesztő munkát folytat. Hosszú távú, közcélú feladatként készíti *A magyar nyelv nagyszótárát*, gondoskodik az elkészülő kötetek megjelentetéséről, az online változat folyamatos elérhetőségéről és frissítéséről, valamint a szótár archív anyagának gondozásáról. Kutatják a magyar nyelv változatait, egyes magyarországi kisebbségi nyelveket és vizsgálat tárgyává teszik az európai integráción belüli nyelvpolitikai kérdéseket. Kiegészítő feladatként nyelvi korpuszokat, adatbázisokat, nyelvi elemző eljárásokat és alkalmazásokat hoznak létre, valamint mindezekhez nyilvános elérést biztosítanak. Közcélú feladatként mindemellett közönségszolgálatot üzemeltetnek, szakértői véleményeket készítenek, részben hatósági megkeresésekre. Az ELTE BTK–NYTK Kihelyezett Elméleti Nyelvészeti Központ (korábbi nevén: ELTE–MTA Elméleti Nyelvészeti Kihelyezett Tanszék) működtetése révén a kutatóközpont közvetlenül részt vesz a felsőoktatásban is, annak minden képzési szintjén.

2. A kutatóhely kiemelt beruházásai 2021-ben

2.1. Neurális NLP Infrastruktúra (Nyelvtechnológiai és Alkalmazott Nyelvészeti Intézet)

Az infrastruktúra az ELKH 2020. évi infrastruktúrafejlesztési pályázatán kapott támogatást, és 2021-ben került beüzemelésre. Az elnyert támogatás összege 121 685 eFt. Az infrastruktúra 3 elemből áll: 1 db mélytanulós NLP kutatásokra tervezett nagy kapacitású GPU klaszter 1; 1 db mélytanulós NLP kutatásokra tervezett nagy kapacitású GPU klaszter 2; nagy kapacitású háttértár. Az infrastruktúra világviszonylatban is élvonalbeli technológiát képvisel, Magyarországra ilyen kapacitású készüléket még nem szállítottak, Európában is párját ritkítja a berendezés, amellyel lehetővé válik a magyar nyelv digitális felzárkóztatása a nagy világnyelvekhez. Ezekkel a berendezésekkel a magyar nyelvre is képesek leszünk előállítani a neurális technológián alapuló olyan nyelvmodelleket, amelyeket eddig a technológiai világcégek leginkább csak angolra fejlesztettek ki.

2.2. Beszédkusztikai és elektrofiziológiai labor (Általános és Magyar Nyelvészeti Intézet)

Hangtompított süketszoba beszédfelvételek és beszédészlelési kísérletek végzésére külső felvétel-irányítással, stúdióminőségű hangfelvevő eszközökkel és a 2021-ben beszerzett EEG-berendezés. Az újonnan létesített EEG-vizsgálóállomás újabb fontos fejlemény a tudományos eszköztár

bővítésében, amely újabb kutatási irányok felé nyit kaput. A mondatfeldolgozással kapcsolatos kutatások egyik célja a magyar mondatok információszerkezetével összefüggő munkamemória- és figyelmi folyamatok vizsgálata elektrofiziológiai mérések alapján. E kutatási irány szervesen folytatható a neurogén nyelvi/kommunikációs zavarok terén, melynek során feltárható, hogy az agyi traumák következtében fennálló nyelvi deficitek (ún. afáziák) esetében tapasztalt mondatfeldolgozási nehézségekhez milyen módon járul hozzá a munkamemória és a figyelmi allokáció esetleges sérülése.

2.3. Lambda munkaállomás-hálózat (Általános és Magyar Nyelvészeti Intézet)

Mélytanulásra, beszédfelismerési kutatásokra optimalizált nagy számítású munkaállomás-hálózat. A meglévő és tervezett beszédatbázisok hanganyagának automatikus lejegyzése mellett alkalmas a különböző gépi tanuló módszerek szisztematikus kutatására és tanulmányozására. A beszerzés 2021 áprilisában történt. A tanuló algoritmusok fejlesztése és kutatása a BME Beszédfelismerés labor és a NYTK Fonetikai kutatócsoport közös munkájával történik.

3. A kutatóközpont közcélú feladatai

3.1. A magyar nyelv nagyszótára

A *magyar nyelv nagyszótára* a magyar nyelvtudomány eddigi legnagyobb szótári vállalkozása, amely a magyar lexikográfiai gyakorlatból még hiányzó, új műfajt képvisel. Történeti jellegű értelmező szótárként legfontosabb feladata a szavak jelentéstartalmának kibontása és e jelentések viszonyrendszerének ábrázolása. Az adatbázisként készülő szótár százezer szócikkben mintegy 110 ezer címszót dolgoz fel, az irodalmi és köznyelv szókincse mellett bőségesen válogatva az archaikussá vált szavakból, a csoport és rétegnyelvek, a nyelvjárások, a szleng stb. szókészletéből is. 2006 óta 8 kötete jelent meg, a 8. kötet 2021 júniusában látott napvilágot. 2017 óta a Nagyszótár elkészült szócikkei online felületen is elérhetők. Ez az évszázados jelentőségű alkotás nem csupán a nyelvészet, az irodalomtudomány és társtudományaik kutatóinak fontos segédeszköze, hanem jogászok és tanárok, újságírók, bölcsészek és természettudósok, könyvtárosok és könyvolvasók, nyelvtanulók vagy a kultúrhistoria iránt érdeklődők egyaránt haszonnal forgathatják.

<http://nagyszotar.nytud.hu/index.html>

3.2. Nyelvi tanácsadó szolgálat

A nyelvi tanácsadó szolgálat működtetése mintegy 70 éves múltra tekint vissza, és fontos szerepe van a nyelvi ismeretterjesztésben, a beszélők nyelvi attitűdjének formálásában. Lehetőséget teremt az alapvető tudományos ismeretek átadása mellett a nyelvi babonák és a nyelvi stigmatizáció negatív hatásának csökkentésére. A nyelvi közönségszolgálat a kutatóközpont egyik legrégebbi, a nagyközönség körében legismertebb és legnépszerűbb szolgáltatása. A közönségszolgálat munkatársai e-mailben és telefonon fogadják a nyelvhasználattal, helyesírással kapcsolatos kérdéseket, s válaszolnak rájuk. A felhalmozódott nyelvi tudásanyagot szabadon hozzáférhetővé teszik.

<http://www.nytud.hu/oszt/nyelvmuvelo/index.html>

3.3. Helyesírási tanácsadó portál

A helyesírási tanácsadás legkorszerűbb, legszélesebb körben használt eszköze a helyesírási tanácsadó portál (<https://helyesiras.mta.hu/>), amely hétféle helyesírási területen (pl. különírás-egybeírás, számok, dátumok helyesírása stb.) képes automatikus javaslatot adni. A felhasználók körében kimagaslóan népszerű, ingyenesen elérhető helyesírási tanácsadó portál évente mintegy 4-5 millió lekérdezést regisztrál. A portál folyamatos fejlesztése, karbantartása társadalmi elvárás, ezért kiemelten fontos célt szolgál.

<http://helyesiras.mta.hu>

3.4. Utónév szakértői munka és Utónévportál

A Nyelvtudományi Kutatóközpont feladata, törvényi kötelezettsége a nem anyakönyvezhető nevek kérelméről, annak bejegyezhetőségéről szakmai állásfoglalás készítése, az Utónévbizottság állásfoglalásainak előkészítése, valamint a honlapon megtalálható névlista folyamatos bővítése. Az anyakönyveztethető utónevek iránti érdeklődés folyamatosan jelen levő társadalmi igény. A hiteles tájékoztatás fontos eszköze az a szabadon letölthető utónévjegyzék, amely az 1997 óta bárki számára

anyakönyvi bejegyzésre alkalmasnak minősített utóneveket tartalmazza. Tekintettel a névkérelmek növekvő számára, a lista havonta frissül.

<http://www.nytud.hu/oszt/nyelvmuvelo/utonevek/index.html>; <http://corpus.nytud.hu/utonevkereso/>

3.5. Többszínűségi és edukációs tudástár és jelnyelvi szótár

A Többszínűségi és Edukációs Nyelvészeti Kutatócsoport korpuszépítési munkálataival párhuzamosan kutatja a magyar jelnyelvet és a jelnyelvhasználó siket közösség, valamint a romani-magyar kétnyelvű cigány közösség helyzetét. Edukációs nyelvészeti kutatásai és fejlesztései eredményeit a Tudástár tartalmazza. Az eredmények a hátrányos helyzetben lévő közösségek oktatásának támogatása, a nyelvi, szociokulturális hátrány pontosabb azonosítása, a magyar nyelvű írás, olvasás és szövegértés fejlesztése és identitásuk erősítése révén növelik e csoportok iskolai és munkai sikerességét. A folyamatosan bővülő Magyar Jelnyelvi Nagyszótár jelenleg 8362 jelnyelvi szócikket tartalmaz (a hozzájuk tartozó videófájlokkal). Funkcióit és tartalmát tekintve különféle felhasználói csoportokat szolgál ki, nemzetközi viszonylatban is innovatív, többszínű kereshetőséget is biztosító, tudományos alapokra épülő szótár.

<http://edulingua.net>; <http://szotar2.jelesely.hu>

4. A kutatóközpont jeles eseményei

4.1. Tudományos rendezvények

XV. Alkalmazott Nyelvészeti Doktoranduszkonferencia, 2021. február 5. Helyszín: NYTK és online. <http://www.nytud.hu/alknyelvdok21/index.html>

15th International Conference on the Structure of Hungarian (ICSH-15), 2021. augusztus 25–26. Helyszín: Pécsi Tudományegyetem BTK (társszervező), Pécs, <https://icsh15.netlify.app/>

Nemzetközi akadémiai emlékülés *Sajnovics János – Demonstratio 250* címmel, az MTA Uráli Nyelvészeti Munkabizottság társszervezésében. 2021. október 21. Helyszín: NYTK és online. <https://mta.hu/i-osztaly/demonstratio-250-111663>

4.2. Egyéb rendezvények, események

2021. május 25.: nemzetközi résztvevőkkel egész napos online megemlékezés Kiefer Ferenc akadémikusról. <http://www.nytud.hu/archiv/kiefer90.html>

2021. július 15.: *A K + K = 120. Papers dedicated to László Kálmán and András Kornai on the occasion of their 60th birthday* c. kötet bemutatója a Nyelvtudományi Kutatóközpontban.

2021. november 9.: *A magyar nyelv nagyszótára 8. (em–ez)* kötetének sajtóbemutatója a kutatóközpontban, a Magyar Tudomány Ünnepe keretében. A szótár 1780 szócikkben 1753 önálló, illetve 307 alcímszót dolgoz fel.

2021. december 16.: *Az Általános Nyelvészeti Tanulmányok XXXIII.* kötetének (*Tanulmányok az Átfogó magyar nyelvtan kutatásainak köréből*) könyvbemutatója a kutatóközpontban. A kötet online elérhető a kiadó MERSZ felületén: <https://mersz.hu/altalanos-nyelvezeti-tanulmanyok-xxxiii>

4.3. Díjak, elismerések

A kutatóközpont egy fiatal nyelvtechnológus munkatársa elnyerte az ELKH Bárány Róbert-díját, „Az igekötők produktív kapcsolódási mintái” című tanulmányáért.

A kutatóközpont által alapított díjakat (Herman József Fiala Kutatói Díj, Szépe György Publikációs Díj) 2021-ben négy kutató kapta meg.

2021-ben a kutatóközpont két munkatársát az Academia Europaea rendes tagjai közé, egy további kollégát pedig a Brit Akadémia tagjává választották.

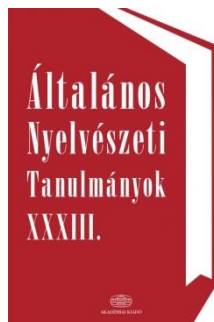
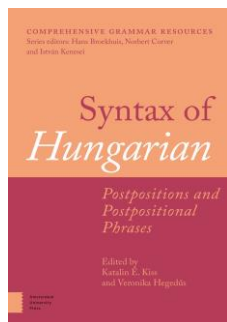
KUTATÓKÖZPONT NEVE:
Nyelvtudományi Kutatóközpont
KUTATÓINTÉZET NEVE:
Általános és Magyar Nyelvészeti Intézet
Cím:
1068 Budapest, Benczúr u. 33.
Postafiók címe:
1394 Budapest, Pf. 360
telefonszám:
(1) 342-9372
honlap:
www.nytud.hu
e-mail:
linginst@nytud.hu

I. Az Általános és Magyar Nyelvészeti Intézet kutatási eredményei a 2021. évre vonatkozóan

Tématerület: 3. Emberi erőforrás

Elméleti és kísérletes szintaxis kutatócsoport és Átfogó magyar nyelvtan kutatócsoport

A leíró és az elméleti nyelvészet területén két jelentős kötet szerkesztési munkálatai fejeződtek be (l. 1. kép). Az Amsterdam University Press gondozásában megjelent a *Syntax of Hungarian* leíró grammatikai könyvsorozat [Postpositions and Postpositional Phrases](#) című kötete, mely a magyar névutók, határozóragok, igekötők és adverbiumok és az ezekre épülő kifejezések eddigi legrészletesebb leírását adja. Megjelent továbbá az [Általános Nyelvészeti Tanulmányok XXXIII. kötet](#) [Tanulmányok az Átfogó magyar nyelvtan kutatásainak köréből](#) címmel, mely tanulmánygyűjtemény a 2012-ben indult *Átfogó magyar nyelvtan* című kutatási projekt leíró természetű munkáiból kiinduló elméleti és kísérletes kutatások eredményeit tartalmazza.



1. kép: *Syntax of Hungarian: Postpositions and Postpositional Phrases*; *ÁNYT XXXIII*

Egy kiadásra elfogadott tanulmány az applikatív vonzatok alapszerkezeti helyzete és a konstrukció által mutatott eszmintázatok közötti összefüggéseket elemzi formális szintaktikai keretben, további megjelenésre már elfogadott tanulmányok pedig az igei kifejezés elhagyását, az ún. ezafé konstrukciót, illetve a szintaxis és fonológia határterületét vizsgálják.

A kutatócsoport több jelentős kísérletes módszerekkel végzett kutatásokon alapuló publikációt is közölt. Kísérletes módszerekkel végzett kutatáson alapul egy [tanulmány](#), amely a fókusz prozódiai jelölését vizsgálja a grúz nyelvben. A grúzban a fókusz-jelölés a mondat prozódiai frázisszerkezetén alapszik, s így olyan nyelvtípusba tartozik, amelyben mindeddig nem vizsgálták a prozódiai fókuszjelölésnek a szintaktikai szerkezettől függő szisztematikus többértelműségének lehetőségeit (az ún. fókusz-projekció jelenségét). A cikk akusztikai bizonyítékok alapján kimutatja, hogy a prozódiai tipológiában elfoglalt helye ellenére a grúzban is megtalálható fókusz-projekció grammatikai rendszere a dallamhangsúly-alapú fókuszjelölést alkalmazó nyelvekben tapasztalt rendszerrel analóg módon

működik. A mondatfókusz a jelentés oldaláról közelíti meg egy olyan [kísérlet](#), amely a fókusz által grammatikalizáltan képviselt különféle implicit jelentések megértését vizsgálta tipikus nyelvfejlődésű, és normál IQ-jú autizmus spektrum zavarral (ASD) élő, illetve szociális pragmatikai kommunikációs zavarral (SPCD) élő 4-9 éves gyermekek csoportjaiban. A konkrét eredmények az *Egészség* téma alatt szerepelnek.

Fonetikai kutatócsoport

Társalgáselemzési kutatásokban a beszélőváltások elemzése azt mutatta, hogy a váltások gyakorisága nőtt a beszélgetések folyamán; gyakrabban mentek végbe egyszerre beszéléssel; a szünettel megvalósuló váltások pedig rövidebb időtartamúak voltak a társalgás vége felé. A társalgások jellemzőit befolyásolja a résztvevők száma is: az egyszerre beszéléses váltások hosszabb idő alatt mentek végbe kétfős beszélgetésekben, mint a háromfősökben. Az önkéntelen vagy szociális nevetések a percepció számára elkülöníthetők. Stresszorok hatásának vizsgálata azt mutatta, hogy a fizikai stressz kondícióban rövidebbek voltak a beszédszakaszok, hosszabbak voltak a néma szünetek, ritkábbak a kitöltött szünetek és a megakadások. Az érzelmi stressz kondícióban volt a legtöbb kitöltött szünet és a legtöbb megakadás. A szünetek elemzése társalgásokban azt mutatta, hogy a szünetek gyakorisága az idő előrehaladtával csökkent, a beszélgetések utolsó szakaszaiban volt a leggyakoribb a közös hallgatás.

A longitudinális kutatások eredményei rámutattak, hogy az alapfrekvencia jellemzően és beszélőn belül konzekvensen, azaz beszédstílustól függetlenül csökken fiatal felnőtt nők beszédében 10 év elteltével, míg a férfiaknál nem volt hasonló csoportszintű változás kimutatható. A hangköz-értékek csökkenése jellemző volt mind a nők, mind a férfiak esetében olvasott beszédben, míg a spontán / félig spontán beszédben nagyobb egyéni variabilitás jelentkezett. Kimutatták, hogy az azonos felvételi időpontban és a 10 év elteltével rögzített beszédminták esetében a beszélőazonosítás eltérő eredményre jut. Míg a biometrikus beszélőazonosítás (Vocalizer) eredménye magas találatlással ismeri fel, hogy az azonos időben rögzített beszédminták akár azonos, akár eltérő beszédstílus esetében azonos beszélőtől származnak-e, addig a 10 év különbséggel rögzített beszédminták esetében a tévesen eltérőnek feltételezett beszélőre utaló pontértékek majdnem minden beszélőnél megemelkednek (egyes beszélőknél akár 80%-ra is a 0–2%-ról).

A beszélőspecifikus jellemzők vizsgálati megállapították, hogy a toldalék nélküli szavak megvalósulása rövidebb, a toldalékosaké hosszabb. A több morfémából álló szavak lexikális hozzáférése több időt vesz igénybe. A beszédben észlelt kellemesség benyomásának vizsgálata azt mutatta, hogy zenészek gyakrabban ítélik meg kedvezőbbben a beszédmintákat, mint nem zenészek, míg a nemek között nem volt szignifikáns különbség.

A dajkanyelv fonetikai jellemzőinek kutatása kimutatta, hogy a dajkanyelvre leheletesebb zöngé jellemző, mint a felnőttekhez szóló beszédre, függetlenül a baba életkorától. A leheletes zöngé a boldog és nyugodt érzelmi állapotok kifejeződése lehet. A gyerekek leheletesebb zöngével beszélnek, mint a felnőttek, így a dajkanyelvi leheletes zöngé megjelenést a beszélgetőpartner produkciójához való igazodás is motiválhatja. Az érzelmek kifejezés dajkanyelvi jelenlétét percepciók eredmények is alátámasztják: az édesanyák beszédében erősebb és pozitívabb töltetű érzelmeket észleltek. Az édesanyák beszéde a babával való kommunikáció során folyamatosan változik ebbe az irányba.

A beszélők közötti variabilitás és a beszélőn belüli variabilitás vizsgálatának keretében bővítették a Longitudinális Korpusz felvételeinek számát. A Beszélt nyelvi adatbázis (BEA) lejegyzési rendszerének egységesítése után megkezdődött az első alkorpusz közreadásának előkészítése, ami 2022 elején várható. Régi magyar filmhíradók hangsávjaiból kezdtek adatbázist fejleszteni az 1931 szeptembere és 1932 decembere közötti időszakból, 140 filmhírből, közel egy órányi fonetikai szempontból feldolgozható beszédanyaggal, beszédszakaszok szintjén annotálva, egy metaadatokat tartalmazó táblázattal kiegészítve.

Fonológia és morfológia kutatócsoport

A magánhangzó-harmónia témájában végzett kutatásokban kísérleteket végeztek és korpusz vizsgálatokat folytattak a magyar előségi harmóniában tapasztalható variációval kapcsolatban. Új [eredmények](#) jelentek meg a fonológiai kondicionált magánhangzó-harmónia és a morfológia interakciójával kapcsolatban, melyek eddig még nem felismert összefüggéseket derítettek fel. Ehhez kapcsolódóan a magyar magánhangzó-harmónia és a kölcsönzésfonológia eddig nem vizsgált találkozási pontjának kísérletes vizsgálata is zajlott. Megvalósult az uráli nyelvek magánhangzó-harmóniáinak tipológiai osztályozása (internetes adatbázis készült, ill. megjelenés előtt állnak tanulmányok), [tanulmány jelent meg](#) az erza harmóniáról, illetve konferencián adtak elő és megjelenés

alatt van egy tanulmány a finn harmónia egy sajátos jelenségéről. Befejezték a munkát két tanulmányon, melyek jelenleg megjelenés előtt állnak a *The Oxford Handbook of Vowel Harmony* című kézikönyvben.

Vizsgálták a magyar paradigmikus rendszer fonológiai jellemzőit, különös tekintettel a tő- és toldalékosztályok szerepére, az itt tapasztalható [graduális mintázatokra](#) (l. 2. kép). Ezzel kapcsolatban beadásra került egy nagydoktori értekezés. Kutatták a fonotaktikai mintázatok gradualitását és univerzális jellemzőit, továbbá részletesen [feltárták](#) a magyar kicsinyítő, illetve gyakorító képzés és a fonotaktika kapcsolatát.

Nemzetközi konferenciára készült előadás, illetve egy angol nyelvű tanulmány a nazális + zár + obstruens kapcsolatokról az angolban és az affrikáták nazális és orális zárhangok utáni gyengüléséről; továbbá egy az [s]+mássalhangzó kapcsolatok CV fonológiai keretben való vizsgálatából készülő tanulmány átdolgozás alatt van.

A cigány nyelvváltozatok kutatásában fontos eredmény a cigányon alapuló kevert nyelvek, a pararomani változatok egyikének, a hungaro-romaninak a vizsgálata, mely változat létének valódi, mai adatokkal történő igazolása és vizsgálata nagy jelentőséggel bír mind a cigány nyelvészet, mind a nyelvi repertoárok kutatása, mind a cigány nyelv identitásformáló szerepe szempontjából. Egy cigány nyelvi, egyelőre kezdeti szakaszban levő adatbázis épül; 18 darab romungró és 3 darab vend nyelvváltozatú hangfelvétel rögzítése és feldolgozása történt meg az ELAN rendszerben.

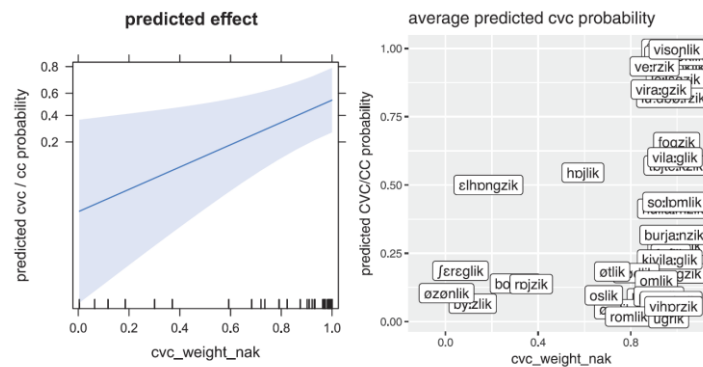


Figure 3: Similarity and CVC/CC probability across stems in the model.

2. kép: Ábra Rác – Rebrus – Törkenczy (2021: 19, Fig. 3) tanulmányából

Kísérleti szemantika és pragmatika kutatócsoport

A figyelemnek a nyelvi feldolgozásban betöltött szerepével kapcsolatos [eredményeket adtak közre](#), melyek hasznosíthatók a logopédiai diagnosztikai, terápiás és intervenciós eljárások kidolgozásában. A pragmatikai jelentések kikövetkeztetésében kimutatható a kontextus feldolgozásának a szerepe már (neurotipikus) óvodáskorú gyermekeknél is. A kontextus feldolgozásának sikere azonban összefügg a munkamemória hatékonyságával, valamint a pragmatikai inferencia az irreleváns információ sikeres gátlásának is függvénye. A mondatfeldolgozás során 18-36 hónapos gyerekeknél kimutatható az élőség hatása a tematikai szerepek mondatrészhez való rendelésében: ha egy tranzitív ige egy élő és egy élettelen tárgyat denotáló főnévvel fordul elő, a gyerekek az élő tárgyat fogják az ágensnek, az élettelen tárgyat pedig a páciensnek tekinteni akkor is, ha ez ellentmond a morfológiai információknak. A hatás különösen a legfiatalabb kohorszban érvényesül.

Nyelvtan és pragmatika kutatócsoport

A kérdő mondatok számos tulajdonságáról tartottak (részben meghívott) előadásokat. Meghívott előadásként prezentáltak a kérdő mondatokban az elfogultság kódolásáról, illetve a tűnődő kérdések indirekt használatának kontrasztív elemzéséről. Előadásokat tartottak az első olyan percepciók kutatás eredményeiről, amely két eldöntendő kérdő mondati alak pragmatikai megfelelőségét vizsgálta három földrajzi régió beszélőivel, a *nem-e-t* tartalmazó kérdő mondatok „elfogultsági profiljáról”, továbbá a kérdő mondatok jelöléséről a német nyelv marzahn (Berlin) dialektusában, valamint az evidenciális-jelölő használatát kutatták udmurt kérdő mondatokban.

Egy [tanulmány jelent meg](#) arról, hogy miért olyan ritka a nyelvekben a promisszív mondatfajta. Konferencia-előadást tartottak a *talán* és az erdélyi dialektusra jellemző *ejsze* partikulák interpretációjáról. Vizsgálták az udmurt evidencialitást és különböző mondatfajtákat, valamint az udmurt létigének a nem kijelentő módokban és modális szerkezetekben megjelenő alakját.

A *Magyar büntetőperes bírósági kihallgatások komplex nyelvészeti elemzése* c. doktori értekezés 2021-es sikeres védeése után készült tanulmány foglalja össze a legfontosabb eredményeket. Olyan kérdésfeltevési stratégiákat és mintázatokat mutatnak be, amelyek a leginkább alkalmasak a kihallgatottak elfogulatlan, pontos és részletes kihallgatásához. Egy másik speciális diskurzustípusként megkezdődött a magyar nyelvű vallási diskurzusok vizsgálata, azon belül is a kérdő mondatok kutatása, melynek keretében adatbázis építését kezdték meg.

Lexikoncikk jelent meg a Montague-Grammatikáról; a *Year's Work in English Studies* c. kritikai bibliográfia (OUP) *English Language* című fejezetében az 2019-ben megjelent szemantikai témájú publikációkat ismertették és értékelték; a *Theoretical Linguistics* c. folyóirat két számát szerkesztették.

Tématerület: Egészség: neurobiológia, pszichológia

Kísérleti szemantika és pragmatika kutatócsoport

Együttműködtek a Szent János Kórház Neurológiai Osztályával egy, a kisagysérült személyek affektív-kognitív profilozását lehetővé tevő tesztcsomag magyar nyelvre történő adaptációjára. Az első elemzések szerint szignifikáns különbségek mutathatók ki a Borderline személyiségzavarban érintett és a kontroll személyek beszédjellemezői között. A klinikai csoport beszédtempója lassabb, több hezitálás jellemzi. Ép intellektusú, autizmusban érintett kisiskolás gyerekek pragmatikai képességei erősebb összefüggést mutatnak tudatelméleti készségekkel, mint a végrehajtó funkciók fejlettségével.

Nyelvi folyamatok agyi reprezentációi kutatócsoport

A sclerosis multiplex (SM) kutatása gépi tanulási szemszögből is vizsgálja az SM nyelvi megjelenési formáit. Az eddigi elemzések eredményei azt mutatják, hogy a vizsgálati protokoll feladatai egy részének automatikus beszédelemzése képes jól elhatárolni az SM-mel élő csoportot a kontroll csoporttól. A fonetikai elemzések eddigi eredményei szerint a megakadások gyakorisága nem mutat különbséget a két csoport között, és mindkét beszélői csoportban hasonló megakadástípusok fordulnak elő. A kognitív és nyelvi összehasonlítás eddigi eredményei szerint az RRMS-betegek beszédét jobban befolyásolja a fokozott kognitív terhelés, mint a kontrollokét.

Az enyhe kognitív zavar (EKZ) nyelvi jellemzőinek kutatása során az S-GAP tesztet farmakológiai hatásmérésre használtuk fel. [Tanulmány jelent meg](#) az EKZ temporális változóiáról. Kutatási együttműködés keretében vizsgálják a Parkinson-kórban megjelenő kognitív, nyelvi- és beszédbeli változásokat. Az első évben feldolgozták a nemzetközi szakirodalmat, kidolgozták a vizsgálati protokollt, és megkezdődtek a vizsgálatok.

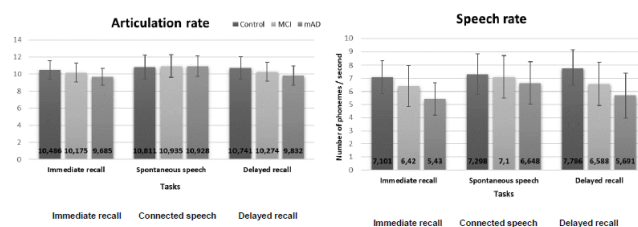


Figure 1. Average values of articulation rate and speech rate markers (number of phonemes / second) of the three speech tasks for the three speaker groups. Groups: MCI = mild cognitive impairment; mAD = mild Alzheimer's disease. Error bars indicate standard deviation. Columns indicate mean values for different groups.

3. kép: Vincze *et al* (2021: Fig 1) tanulmányából

A pragmatikai nyelvi zavarok (PKZ) vizsgálatában autizmussal (ASD) és pragmatikai kommunikációs zavarral (SPCD/PKZ) élő gyermekeknél vizsgálták [implicit jelentéseket](#) (közös eredmény az Elméleti és kísérletes szintaxis kutatócsoporttal), és azt találták, hogy mind az ASD, mind az SPCD csoport szignifikánsabb alacsonyabb szinten teljesített az implicit jelentések megértésében, mint a tipikus fejlődésű gyermekek. Az ASD-vel élő gyermekeknél a tudatelméleti képesség szintje befolyásolta az implicit jelentések megértésének sikerességét, míg a grammatikai képességek nem; ezzel szemben SPCD gyermekeknél a grammatikai képességek befolyásolták a megértést, míg a tudatelméleti képesség nem. Tipikus fejlődésű gyermekeknél sem a tudatelméleti képesség szintje, sem a grammatikai képesség fejlettsége nem mutatott összefüggést a grammatikalizált implicit jelentések megértésével.

Kutatás indult a belső (önmagunkkal némán folytatott) beszéd szintaktikai és szemantikai jegyeinek felderítésére; online felmérésorozatot zajlott le. Folytatódott a magyar rekurzív szerkezetek elsajátításának vizsgálata, melyben a produkció eredményei azt mutatják, hogy a négytől nyolcéves korig terjedő csoportban a rekurzív szerkezeteket előbb sajátítják el, mint az ugyanolyan komplexitású, de nem rekurzív szerkezeteket.

Nyelvfejlődés kutatócsoport

A korai nyelvfejlődés neurokognitív és szociális korrelátumait vizsgálták; zajlott az adatgyűjtés, és megindult az adatok feldolgozása, melyekből több előadás készült és előkészületben vannak publikációk. A kétéves korban megfigyelt anyai szenzitivitásnak a kisgyermek nyelvfejlődésével való összefüggésére irányuló kutatást végeztek, valamint anya-gyermek interakciókat vizsgáltak kihívást jelentő helyzetben, illetve ezek összefüggéseit a nyelvfejlődéssel. Kimutatták, hogy kétéves gyerekek gesztushasználata negatív összefüggésbe hozható megkésett nyelvfejlődésükkel. Korai mondatmegértési heurisztikák (tematikus szerepazonosítás) szemmozgás-követéses vizsgálatában az élőség és a szórend relatív fontosságának alakulását vizsgálták.

A Komplex Beszélte Anyanyelvi Képességteszt (KOBAK) fejlesztésével egy hazánkban hiánypótló, gyermekkorban alkalmazható, az anyanyelvi fejlődés kimeneteit objektíven mérni képes, így többek között a nyelvfejlődési zavarok, a diszlexia és az autizmus spektrumzavar diagnosztikájában használható, átfogó diagnosztikai eljárás kialakítása a cél. A teszt fejlesztése az ELKH-BME Lendület Nyelvsajátítás kutatócsoporttal (vk: Lukács Ágnes) és az ELTE BGGYK Logopédiai szakcsoporttal együttműködésben folyik.

Tématerület: Egészség: 1.1 Covid19: a koronavírus-járványhoz kapcsolódó kutatások

Nyelvfejlődés kutatócsoport

Kutatás indult a járványhelyzeti korlátozások hatásainak felmérésére a korai társas-kommunikatív viselkedésben és szókincsfejlődésben. A kutatásban másfél éves kisgyermekek közös figyelem képességét és szókincsfejlődését vizsgálták a járványhelyzettel összefüggésben. Az eredmények szerint a járványhelyzetet legalább fél éven keresztül megtapasztalt 18 hónapos csecsemőket kevesebb szemkontaktus-jellegű közös figyelem kezdeményezés jellemezte, míg szókincsfejlődésük tipikusnak tekinthető. A családon belüli interakciók a normatív úton tarthatták a korai nyelvfejlődést és gesztushasználatot, ugyanakkor a külvilággal folytatott érintkezések számának jelentős csökkenése állhat annak háttérében, hogy a gyermekek egy ismeretlen felnőttel folytatott interakcióban ritkábban néztek a partnerre közös figyelmi helyzet kialakítása céljából. A járványidőszakban a tipikusan fejlődő gyermekeknél is megjelenhet az atipikus fejlődésre (pl. ASD) jellemző viselkedésmintázat a közös figyelem készségekben, amit a gyakorlati szakembereknek szükséges lehet figyelembe venni.

II. Tudományszervezési eredmények, események

1. Rövid általános ismertető a kutatóhely tudományos munkáját illetően.

Az Általános és Magyar Nyelvészeti Intézetben folyó kutatások nyelvelméleti és általános nyelvészeti kutatásokra, az általános és magyar fonetika területére, a pszicholingvisztikai, neurolingvisztikai és kognitív tudományi alap kutatásokra fókuszálnak, valamint beszédatadabázisok, illetve a nyelvsajátítást, anyanyelvi fejlesztést és rehabilitációt támogató módszerek fejlesztésével foglalkoznak. A kutatócsoportok feladatai közé tartozik többek között a beszédprodukció és beszédfelismerés kutatása; a kommunikációs jellemzők és a dajkanyelv vizsgálata; a mai magyar

nyelv és más nyelvek nyelvi rendszereinek leírása; kutatások a fonológia, a morfológia, a szintaxis, a szemantika és a pragmatika területein; a magyar prozódia vizsgálata; a nyelvi feldolgozás mentálisan reális folyamatainak kutatása, az anyanyelvi fejlődés és zavarainak vizsgálata; különböző neurodegeneratív betegségekkel összefüggő nyelvi funkciózavarok kutatása. Az ELTE BTK – NYTK Kihelyezett Elméleti Nyelvészeti Központ fenntartása az ELTE Bölcsészettudományi Karával közös küldetés; a Központ feladata az ELTE elméleti nyelvészeti alap- és mesterszak, valamint doktori program működtetése.

2. A tudomány és a társadalom kapcsolata, társadalmi hasznosság.

Az intézet kutatási területei között fontos (és egyre növekvő) szerepe van a közvetlen orvosi-népegészségügyi relevanciával bíró pszicho- és neurolingvisztikai kutatásoknak és fejlesztéseknek. A neurolingvisztikai munkálatok fókuszában a sclerosis multiplex, valamint új kutatásként a Parkinson-kór nyelvi és kognitív vonatkozásainak feltárása áll. Kutatják továbbá az enyhe kognitív zavar nyelvi jellemzőit és a pragmatikai nyelvi zavarokkal élő gyermekek nyelvi képességeit, mely kutatások eredményei megfelelő diagnosztikai eljárások kidolgozásához járulnak hozzá. A nyelvelsajátítás kutatásának egyik fontos célja egy gyermekkorban alkalmazható, átfogó diagnosztikai eljárás kidolgozása.

Egy munkatársunk számos ismeretterjesztő tanulmánya a [Qubit](#) tudományos-technikai online folyóiratban jelent meg, valamint rendszeres rádióműsort (Szószátyár) vezetett és egyéb ismeretterjesztő előadásokat tartott. Egy további kollégánk egy internetes [blogon](#) végez szerkesztői és vendégszerzői tevékenységet.

Kollégáink ismeretterjesztő programokat tartottak a Kutatók Éjszakáján és a Hang világnapja rendezvénysorozataiban. Közreműködtek a Magyar Nyelvészeti Diákolimpia szervezésében és a Nemzetközi Nyelvészeti Diákolimpia magyar csapatának összeállításában és felkészítésében. Zsűritagként vettek részt az Eötvös-szónokversenyen.

3. A kutatóhely legfontosabb hazai és nemzetközi eredményes K+F+I együttműködései az elmúlt évben.

a. Vállalati kapcsolatok, az állami költségvetési, a nonprofit és a kormányzati szektorral való kapcsolatok

Hazai kapcsolatok:

- Tudományos együttműködés és kutatási kapcsolatok az Uzsoki Utcai Kórház Neurológiai Osztályával, a Szegedi Tudományegyetem Pszichiátriai Klinikájával.
- MILAB pályázat keretében együttműködés: ELTE-BTK, ELTE-IK Mesterséges Intelligencia Tanszék, SZTE-ELKH Mesterséges Intelligencia Kutatócsoport; SZTE Neurológiai Klinika, SOTE Neurológiai Klinika Parkinson Ambulancia
- Kapcsolat a Devwing Kft.-vel az NKFIH 132460 sz. pályázat online tesztfelületének fejlesztése és karbantartása céljából.
- Együttműködés a Szent János kórház *Cerebellum* kutatócsoportjával.
- Együttműködés a Budapesti Műszaki és Gazdaságtudományi Egyetem Távközlési és Médiainformaticai Tanszékével mélytanuláson alapuló neurális hálók beszédfelismerési feladatra való alkalmazására használható Lambda munkaállomás-hálózat kialakításában.
- Együttműködés a Nemzetbiztonsági Szakszolgálattal olyan vizsgálatokat lehetővé tevő korpusz kidolgozásában, amelyet igazságügyi szakértői feladataik során fel tudnak használni.
- Meixner Alapítványnál óraadói feladat ellátása a Meixner Ildikó által kidolgozott dyslexia prevenció és reedukációs módszer ismertetése c. tanfolyamon részt vevő pedagógusoknak. Kurzus címe: *Nyelvészet: beszédfejlődés, akusztika, szociolingvisztika.*

Nemzetközi kapcsolatok:

- A berlini Humboldt-Universitát (Institut für deutsche Sprache und Linguistik) látja vendégül Humboldt-ösztöndíjas kollégánkat.
- „A másodlagos predikátumok nyelvtani megszorításainak vizsgálata” c. MTA kutatási mobilitási pályázat keretében megrendezték a *Second International Workshop on Secondary*

Predication műhelykonferenciát, a japán partnerek (Kóbe, Tokió) együttműködésével (2021. október 30–31. Online.)

- Közös kutatás folyik a Groningeni Egyetem kutatóival a tagadott diszjunktív mondatok nyelvközi kísérletes vizsgálata témájában.
- Együttműködés kezdődött a berlini ZAS kutatóival a *Realizing Leibniz's Dream: Child Languages as a Mirror of the Mind* című ERC projektben. A magyar nyelv felelős kutatójának feladata elsősorban a kísérletek magyarra adaptálása, illetve később ezeknek a vizsgálatoknak az elvégzése magyar anyanyelvű gyerekek és felnőttek bevonásával.
- Zsűritagként részvétel az Academia Europaea Budapest Knowledge Hub által szervezett nemzetközi mentori rendezvényen a humán- és társadalomtudományi ERC pályázati mentorprogramban.
- Online „Yiddish Reading Group” szervezése a University College London kutatóival közösen.

b. Felsőoktatási intézményekkel való együttműködések: oktatás, utánpótlás-nevelés, doktori iskolák stb.

Több mint három évtizede működik az ELTE-vel közös nyelvészeti képzés, az ELTE BTK–NYTK Kihelyezett Elméleti Nyelvészeti Központ, mely gondoskodik az alap- és mesterszakos, valamint a doktori kurzusok megtartásáról, a tutorizálásról, szakdolgozati és doktori disszertációs témavezetésről.

Emellett az intézet munkatársai számos felsőoktatási intézményben oktattak: ELTE BTK; ELTE BGGYPK; BME; PPKE BTK; PTE BTK; SZTE BTK; KRE BTK.

Részt vesznek doktori képzésekben: témavezetés az ELTE, a PPKE, az SZTE Nyelvtudományi Doktori Iskoláiban; bizottsági tagságok.

A trondheimi Norwegian University of Science and Technology meghívására három napos minikurzust tartottak doktori hallgatóknak a másodlagos predikációról és adjunkcióról (2021. május 18–20.).

c. Egyéb együttműködések, kapcsolatok.

Részvétel hazai tudományos testületek munkájában:

MTA Nyelvtudományi Bizottság, MTA Orientalisztikai Bizottság, MTA Közoktatás Fejlesztési Programtanács, MTA Pedagógus Pályázat Programtanács

Részvétel nemzetközi tudományos testületek munkájában:

European Association of Chinese Linguistics (vezetőségi tag), Európai Kutatási Tanács SH–4 panel *The Human Mind and Its Complexity* ERC Consolidator Grant zsűri tagság, International Society of Phonetic Sciences (alelnöki pozíció), IPA History Subcommittee, International Society of English Phonetics, Romanian National Research Council (Consiliul Național al Cercetării Științifice) zsűri tagság, Studia Anglica Resoviensia, Speech Prosody Special Interest Group

Sorozatszerkesztés:

Comprehensive Grammar Resources (Amsterdam University Press), Studia Grammatica (De Gruyter), *Studies in Natural Language & Linguistic Theory* (Springer)

(Fő)szerkesztés:

Általános Nyelvészeti Tanulmányok, *Beszédtudomány – Speech Science*, *Finno-Ugric Languages and Linguistics*, *Logopédia*, *Theoretical Linguistics*

Szerkesztőbizottsági tagság:

Acta Linguistica Academica, *Alkalmazott Nyelvtudomány*, *English Linguistics*, *Estudios de Fonética Experimental (Journal of Experimental Phonetics)*, *Frontiers in Language Sciences*, *Glossa*, *Gondolat-Jel* (nyelvészeti könyvsorozat, Szépirodalmi Figyelő Alapítvány), *Govor*, *International Journal of Chinese Linguistics*, *Jelentés és Nyelvhasználat*, *Journal of Estonian and Finno-Ugric Linguistics*, *Journal of Linguistics*, *Korean Journal of Linguistics*, *Languages*, *Linguistic Inquiry*, *Linguistic Variation*, *Linguistics Beyond and Within*, *Linguistische Berichte*, *Logopédia*, *Magyar Nyelvőr*, *Natural Language & Linguistic Theory*, *Nyelvtudományi Közlemények*, *Papers in Arts and Humanities*,

4. A kutatóhely hazai vonatkozású (így különösen NKFI Alap, OP) pályázati adatai.

NKFI Alap

NKFIH K_21_139271 – *A paradigmaszerkezet szerepe a magyar fonológiában és alaktanban nyelvtipológiai szempontból*

Vezető kutató: Rebrus Péter

Időtartam: 2021.09.01–2025.08.31.

Támogatás: 39,723 M Ft; Támogatás forrása: NKFI Alap

Téma: A magyar és más nyelvek nem agglutinatív morfológiai morfofonológiai jelenségeinek paradigma-alapú leírása és ennek elméleti következményeinek kutatása. A projekt célja kettős: egyrészt kimerítő, deskriptív szempontból precíz és kvantitatív, azaz a gyakoriságot is figyelembe vevő leírását akarja adni a vizsgálandó jelenségeknek, hangsúlyosan ideértve az e jelenségek körében tapasztalható lexikális variációt és ingadozást, paradigmikus hiányt, illetve templatikus jelenségeket; másrészt olyan formális modelleket, esetlegesen azok számítógépes implementációit akarja alkalmazni és kifejleszteni, melyek a feltárt nem agglutinatív jelenségek elemzését jobban lehetővé teszik.

NKFIH MEC_SZ_21_140739 – *20. Nemzetközi Morfológiai Találkozó szervezésének támogatása*

Pályázó: Rebrus Péter

Időtartam: 2021.12.01–2022.11.30.

Támogatás: 5 040 000 Ft; Támogatás forrása: NKFI Alap

Téma: A 20. Nemzetközi Morfológiai Találkozó megszervezése. A konferencia nagy nemzetközi jelentőséggel bír: ez a legrégebbi morfológiai rendezvény nem csupán Európában, hanem a világon. Kétévente felváltva Bécsben és Budapesten rendezik meg, ami egyrészt fontos nemzeti eseménnyé, másrészt jellegzetesen közép-európaivá is teszi, és nyelvészeti tekintetben előkelő helyet biztosít Magyarország számára.

ELKH

ELKH kiemelt téma kutatásaként (SA-53-2021): *A moldvai magyar nyelv: Adatbázis-építés és grammatikai kutatások*

Vezető kutató: Surányi Balázs

Időtartam: 2022. szeptember 1– 2024. augusztus 31.

Támogatás: 45 940 000 Ft.

Téma: A moldvai magyar (csángó) kitüntetett helyzet foglal el a magyar nyelvváltozatok között sajátos fejlődése, nyelvjárási tagoltsága és súlyosan veszélyeztetett helyzete okán. Az elnyert projekt kettős célzatú: hanganyag-alapú, hanghű lejegyzéssel ellátott és morfológiailag annotált moldvai nyelvi adatbázis létrehozása, és a moldvai magyar nyelv egyes jellegzetes grammatikai vonásainak leírása és korszerű nyelvméleti keretben való elemzése.

5. Tudományos publikációk

Bóna, Judit – Kohári, Anna. 2021. Rate vs. rhythm characteristics of cluttering with data from a “syllable-timed” language. *Journal of Fluency Disorders* 67. 105801. 1–30. <http://real.mtak.hu/116754/>, <http://doi.org/10.1016/j.jfludis.2020.105801>

Borise, Lena. 2021. Focus projection need not be based on pitch accents: evidence from Georgian. *Glossa: a journal of general linguistics* 6(1): 128, 1–46. <https://www.glossa-journal.org/article/5733/galley/18392/download/>

É. Kiss, Katalin – Hegedűs, Veronika (szerk.) 2021. *Syntax of Hungarian. Postpositions and Postpositional Phrases*. Amsterdam: Amsterdam University Press. <https://library.oapen.org/handle/20.500.12657/47855>

Fejes László. Erzya stem-internal vowel-consonant harmony: A new approach, *Acta Linguistica Academica* 68(1-2): 158–174. <http://real.mtak.hu/134589>, <https://doi.org/10.1556/2062.2021.00466>

Gärtner, Hans-Martin. 2021. On the Utility of Commissive Signals and the "Promissive Gap." In: Ömer, Eren; Asimina, Giannoula; Sam, Gray; Chi-Dat, Lam (Daniel); Aurora, Martinez Del Rio (szerk.) *Proceedings of the Fifty-fifth Annual Meeting of the Chicago Linguistic Society*. Chicago Linguistic Society. 123-135. <http://real.mtak.hu/137527/>

Gósy, Mária 2021. On the history of palatography in Hungarian phonetics. *Journal of the International Phonetic Association* 1–12. Published online by Cambridge University Press: 03 November, 2021. <http://real.mtak.hu/138494/>, <https://doi.org/10.1017/S0025100321000293>

Gosztolya, Gábor – Balogh, Réka – Imre, Nóra – Egas-López, José Vicente – Hoffmann, Ildikó – Vincze, Veronika – Tóth, László – Devanand, Davangere P. – Pákáski, Magdolna – Kálmán, János. 2021. Cross-lingual detection of mild cognitive impairment based on temporal parameters of spontaneous speech. *Computer Speech and Language* 69: Paper 101215. 1–15. <http://real.mtak.hu/128678/>, <https://doi.org/10.1016/j.csl.2021.101215>

Harmati-Pap Veronika – Vadász Noémi – Kas Bence – Tóth Ildikó. 2021. Anyai dajkanyelvi narratívák lexikai és szintaktikai jellemzőinek longitudinális vizsgálata, *Beszédtudomány / Speech Science*: 207–242. <https://ojs.mtak.hu/index.php/besztud/issue/view/672/414>

Hegedűs Veronika (szerk.) 2021. *Általános Nyelvészeti Tanulmányok XXXIII. Tanulmányok az Átfogó magyar nyelvtan kutatásainak köréből*. (Főszerkesztő: Surányi Balázs) Budapest: Akadémiai Kiadó. <https://mersz.hu/altalanos-nyelvészeti-tanulmányok-xxxiii>

BENNE: Bánréti Zoltán – Lévai Dániel – Ugrin Bálint: Társas interakciót jelentő nominalizációs szerkezetek kísérletes vizsgálata
Szeteli Anna – Kenesei István: Szintaktikai dialektusok statisztikai elemzése felé: Egy új alárendelő szerkezet elfogadottságának felmérése

Káldi, Tamás – Babarczy, Anna. 2021. Linguistic focus guides attention during the encoding and refreshing of Working Memory content. *Journal of Memory and Language* 116, 104187. <https://www.sciencedirect.com/science/article/pii/S0749596X20301017>

Káldi, Tamás – Szöllösi, Ágnes – Babarczy, Anna. 2021. Hungarian Structural Focus: Accessibility to Focused Elements and Their Alternatives in Working Memory and Delayed Recognition Memory. *Frontiers in Psychology* 12, 1233. <https://www.frontiersin.org/article/10.3389/fpsyg.2021.514886>

Rácz, Péter – Rebrus, Péter – Törkenczy, Miklós. 2021. Attractors of variation in Hungarian inflectional morphology. *Corpus Linguistics and Linguistic Theory* 17(2): 287–317. <http://real.mtak.hu/90748/>, <https://doi.org/10.1515/cllt-2018-0014>

Rebrus, Péter – Törkenczy, Miklós. 2021. Harmonic Uniformity and Hungarian front/back harmony. *Acta Linguistica Academica* 68(1-2): 175–206. <https://doi.org/10.1556/2062.2021.00475>

Svindt, Veronika – Surányi, Balázs. 2021. The comprehension of grammaticalized implicit meanings in SPCD and ASD children: A comparative study. *International Journal of Language & Communication Disorders* 56: 1147-1164. <https://onlinelibrary.wiley.com/doi/epdf/10.1111/1460-6984.12657>

Vincze Veronika – Szatlóczki, Gréta – Tóth László – Gosztolya, Gábor – Pákáski Magdolna – Hoffmann, Ildikó – Kálmán, János. 2021. Telltale silence: temporal speech parameters discriminate between prodromal dementia and mild Alzheimer's disease. *Clinical Linguistics & Phonetics* 35:8: 727–742. <http://real.mtak.hu/121463/>

KUTATÓKÖZPONT NEVE:
Nyelvtudományi Kutatóközpont
KUTATÓINTÉZET NEVE:
Lexikológiai Intézet
Cím:
1068 Budapest, Benczúr u. 33.
Postafiók címe:
1394 Budapest, Pf. 360
telefonszám:
(1) 342-9372
honlap:
www.nytud.hu
e-mail:
linginst@nytud.hu

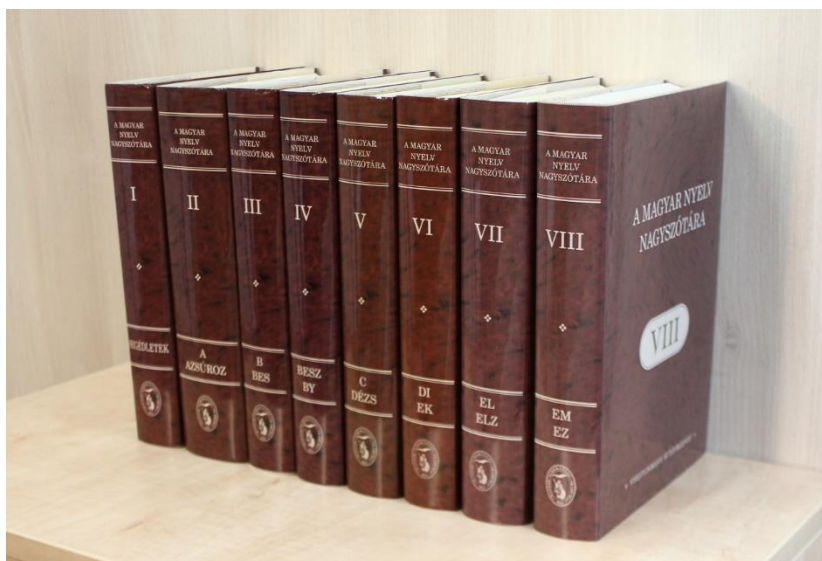
I. A Lexikológiai Intézet kutatási eredményei a 2021. évre vonatkozóan

Tématerület: 3. Emberi erőforrás

A magyar nyelv nagyszótárának munkálataiban 2021. első felében befejeződött a VIII. *em*–ez kötet szócikkeinek főszerkesztése, lektorálása, morfológiai kódolása, és ezzel párhuzamosan megtörtént a kötet új szerzőségi adatainak ellenőrzése, és júniusban megjelent a nyomtatott kötet, amelynek anyaga elérhetővé vált a [nagyszotar.nytud.hu oldalon](http://nagyszotar.nytud.hu/oldalon) is. Az év során folyt a *fa*–*fá* kezdetű anyag szerkesztése és a *fi* kezdetű szavak szócikkeinek írása. November–decemberben a kutatócsoport kb. felének részvételével elkezdődött egy új munkaszervezési rend kipróbálása. Az ebben a munkarendben dolgozók az *f* betű utáni anyagrészből válogatva korábban megírt szócikkekkel szemantikai, illetve fogalomkörü kapcsolatokat mutató mintegy 200 szó jelentős részben formalizált struktúrájú szócikkét írták meg. A két munkarend szerint összesen 862 szócikket (2790 jelentésegység) írtak meg, és megtörtént 697 szócikk (2094 jelentésegység) hagyományos formájú megszerkesztése. A hagyományos munkarendben dolgozók csoportjában ebben az évben is folyt új szerkesztő munkatársak betanítása az ún. előszerkesztői rendszerben, amelyben a betanulók által szerkesztett anyag kontrollszerkesztését a főszerkesztő, illetve egy gyakorlott szerkesztő végezte. A kutatócsoport két félállású új munkatárssal bővült. A lexikográfiai munka segítésére megtörtént az archivális cédulagyűjtemény korábban már szkennelt *f* betűs céduláinak kereshetővé tétele. Ehhez a kutatócsoport munkatársai mintegy 400 000 cédula mintegy 40 800 címszavának rögzítését végezték el, megtörtént a hibás adatok ellenőrzése, javítása, valamint a szótár belső portálján elérhető új webalkalmazásként a keresőfelület kialakítása. Szövegválogatással folytatódott az újabb korpuszbővítés előkészítése. E bővítés során újabb hárommillió szövegszönyi 2011 és 2020 között keletkezett szöveggel, szövegrészlettel bővül majd az elektronikus formában tárolt [Magyar történeti szövegtár](#).

Az **Új magyar etimológiai szótárban (ÚESz.)** a szócikkírás fázisa lezárult 2021-ben, de a szócikkek mintegy harmadának főszerkesztői ellenőrzése, illetve az ebből fakadó javítások, átdolgozások végrehajtása még hátravan. Az ÚESz.-címszavak szótörténeti frissítésében ezúttal a Magyar történeti szövegtár anyaga kapott nagyobb szerepet, ennek segítségével számos címszónál sikerült korábbi adatokkal gazdagítani a szótörténeti részeket; pl. *szaltó*, *szamovár*, *tabletta*, *technikum*, *tendál* stb. Az Arcanum Digitális Tudománytár nagy anyagában az 1780-as évektől kezdődően ezúttal is számos olyan fontos adalékot sikerült találni, amelyek az ÚESz. más forrásaiban nem szerepelnek, például a következő szavak esetében: *szaletli*, *szanitéc*, *szemafor*, *tájfún*, *tarhál*, *táv*-(irat, -köz, -úsás), *tendencia*, *teniszszekik*, *textil* stb.

A szócikkek etimológiai szakaszait tekintve az újabb etimológiai irodalom által érintett címszavak esetében, illetőleg az irodalomtól függetlenül szükségesnek mutató felülvizsgálatot követően az *r*–*t* részben kialakították az eredetmagyarázatok első változatát. Összeállították a szótár bibliográfiáját A magyar nyelv történeti-etimológiai szótára, az *Etymologisches Wörterbuch des Ungarischen* és az Arcanum Digitális Tudománytár alapján. Ezen kívül elkészültek az egységes szófajrövidítések, a szótár mutatója, a szócikkvégi utalórendszer, a rövidítésjegyzék is. Az új források, új első adatok felkutatása, bedolgozása (247 adat) is megtörtént.



1. kép: A magyar nyelv nagyszótára I–VIII.

A Lexikai tudásrepresentáció kutatócsoporton belül az új magyar értelmező szótár munkálataiban a 2021-es év legfontosabb eredménye egy magyar igekötő-kapcsoló eszköz létrehozása volt (az eszköz az [e-magyar nyelvfeldolgozó rendszer](#) emPreverb moduljaként érhető el).

cql "elkészül" 1,153 (65 per million)

Detalleft context KWIC Right context

1	<input type="checkbox"/>	<input type="checkbox"/>	és az Enese közötti szakasz	készülne	el . </s><s> • FEJLESZTÉS </s>
2	<input type="checkbox"/>	<input type="checkbox"/>	ék . </s><s> 1997 júniusában	elkészült	az egymilliomodik Opel motor
3	<input type="checkbox"/>	<input type="checkbox"/>	nillárd forintos beruházással	készült	el 1996-ban . </s><s> A 4740
4	<input type="checkbox"/>	<input type="checkbox"/>	1 megjelenésű Multipiának is	elkészült	egy új , környezetbarát verziój
5	<input type="checkbox"/>	<input type="checkbox"/>	re csatlakozik . </s><s> Nem	készült	el viszont az új , hidraulikus ve
6	<input type="checkbox"/>	<input type="checkbox"/>	en fekvő komplexum 2002-re	készül	el teljesen , maga a beruházás

2. kép: Keresés az emPreverb modullal elemzett, NoSketchEngine rendszerbe feltöltött korpuszban

További előkészítő munkálatok és kísérletek folytak az új magyar értelmező szótár projektben: elkészült a felhasználható magyar korpuszok számbavétele, illetve egy OCR-javító kiértékelésre szolgáló eszköz is. Vizsgáltuk *A magyar nyelv értelmező szótára* definícióinak redundanciáját, megkezdődött egy gyakorisági listákat összevető, valamint egy címszólistákat és gyakorisági listákat összevető programmodul készítése. Ezekon kívül kísérleteztek előtanított nyelvmodellekkel az ékezetelenség-felismerés és a töredezettségmentesítés területén, és egyelőre belső használatra elkészült egy nagyméretű, e-magyar rendszerrel elemzett korpusz NoSketchEngine korpuszkezelő rendszerbe integrált változata, a korpusz annotációjának előállításánál az igekötő-kapcsoló modul is felhasználásra került. Elkészítették az emCompound e-magyar modult, ami az összetett szavak összetételi határát állapítja meg és annotálja az e-magyar rendszerrel elemzett korpuszokban.


Egy további vizsgálat az igekötős szerkezetek korpuszvezérelt vizsgálatára irányult, ennek egyik fontos, gyakorlati eredménye a [PrevDistro adatbázis](#), amely 49 igekötős szerkezet típus 41,5 millió korpuszbeli előfordulását tartalmazza, és bárki számára szabadon hozzáférhető.



June 21, 2021

Dataset Open Access

PrevDistro - Preverb Distributions in Hungarian

 Kalivoda, Ágnes

PrevDistro (Preverb Distributions) is an open-source dataset containing 41.5 million corpus occurrences of 49 preverb-verb construction types. It consists of the following columns:

- 1 *sid*: ID
- 2 *constype*: construction type
- 3 *subtype*: construction subtype
- 4 *prevpos*: preverb position
- 5 *prev*: preverb
- 6 *verb*: verb lemma
- 7 *intervening*: intervening words (as lemmas)
- 8 *actform*: actual form (the same content as in column 10, but this column is lowercase)
- 9 *left*: left context
- 10 *kwic*: keyword in context
- 11 *right*: right context
- 12 *docid*: document ID from the Hungarian Gigaword Corpus
- 13 *title*: document title
- 14 *style*: document style (e.g. official, press, ...)
- 15 *region*: document region (e.g. Transylvania, Subcarpathia, ...)
- 16 *year*: year of publication (sometimes several years can be found in one document)

The first row stands for the header. If a cell's value is unspecified, it is marked with underscore (_).

3. kép: A PrevDistro adatbázis

2021 elején lezárultak a [Szovjet táborok magyar foglyai adatbázis](#) munkálatai. A Magyar Nemzeti Levéltár által vezetett projektben a kutatócsoport feladata a cirill betűkkel írt magyar tulajdonnevek rekonstrukciója volt. 2021. február 25-én megnyílt a nagyközönség számára a keresőfelület. A közel 700.000 magyar fogoly adatait tartalmazó adatbázis jelentős sajtóvisszhangot kapott.

2021 szeptemberében létrejött a Terminológiai Kutatócsoport a leendő Magyar Terminológiai Koordinációs Központ előkészítésére. Ennek a csoportnak a 2021. évre allokkált feladata a központ leendő munkatársaival kialakítandó kapcsolat kiépítése volt. Ennek keretében a kutatócsoport felkeresett számos hazai és EU-s terminológiai hálózatot, valamint kapcsolatot épített a IATE-adatbázisának üzemeltetőjével. Hazai és nemzetközi kapcsolatok építése terén fontos lépés volt a hazai terminológiában kiemelkedő szerepet játszó intézményekkel kapcsolat kialakítása. Ennek kapcsán együttműködési keretmegállapodást kötött többek között az Országos Fordító- és Fordításhitelesítő Iroda Zrt.-vel, valamint a határon túli magyar nyelvhasználatok nyelvi tervezésével, leírásával foglalkozó Termini Magyar Nyelvi Kutatóhálózattal.

2021-ben a hazai és nemzetközi romani nyelvű szótárakkal foglalkozó szakirodalmakban megfogalmazott kritikai észrevételek, ajánlások alapján egy kétnyelvű szótár megalkotását tűztük ki célul, amely elsősorban a hazai lovári nyelvű szótárakra alapozva kívánja a nyelv fejlesztését, művelését, megmaradását szolgálni. A ma, hazánkban fellelhető cigány nyelvvel, nyelvekkel foglalkozó szótárak nem klasszikus értelemben vett szótárak, azokat sokkal inkább szöszedeteknek hívhatjuk, amely elsősorban a szavak denotatív jelentését hivatottak megadni nélkülözve a kontextuális jelentéseket, illetve a példamondatok is, amely nagymértékben szűkíti a nyelv megtanulása iránt érdeklődők lehetőségeit is. A tervezett szótár számtalan példamondattal, kontextuális jelentéssel kívánja bővíteni a már meglévő lexikai elemeket; új, eddig szótárakban még

meg nem jelent szavakkal, szócikkkel kívánja gazdagítani a deskriptív nyelvészet lexikográfiai ágát; nyelvújító szándékkal javaslatokat tesz a romani nyelvben fel nem lelhető szavak ekvivalensének megalkotására az LWB szótárszerkesztő program segítségével. A magyar nyelvű gyakorisági mutató alapján, valamint a romani nyelvű szótárakban fellelhető szavak számát figyelembe véve célunk közel 6000 szócikk megalkotása, melyből 2021-ben 500 szócikk készült el.

2. A tudomány és a társadalom kapcsolata, társadalmi hasznosság

A Lexikológiai Intézetben készült egynyelvű szótáraknak kiemelt jelentőségük van és lesz nyelvünk, azaz egyik legfontosabb kulturális örökségünk jelenének és múltjának feltérképezésében és hozzáférhetővé tételében mind a tudomány, mind a művelt nagyközönség számára.

A *magyar nyelv nagyszótára* a magyar nyelvtudomány eddigi legnagyobb szótári vállalkozása. Történeti jellegű értelmező szótárként legfontosabb feladata a szavak jelentéstartalmának kibontása és e jelentések viszonyrendszerének ábrázolása. A szótár nem csupán a mai nyelvállapot szókészletét elemzi a lexikográfia eszközeivel, hanem címszóállományában, szótárazott jelentéseiben és példamondataiban egyaránt történeti dimenziót érvényesít, amikor az elmúlt 240 esztendő szókincséből merít. A Nagyszótár korpuszalapú, történeti jellegű értelmező szótárként nem csupán a mai nyelvállapot szókészletét tükrözi és elemzi a lexikográfia eszközeivel, hanem történeti szemléletet is érvényesít. A szófaji és egyéb grammatikai információk mellett a szó stílus- és használati körére vonatkozó minősítések is találhatóak a szótárban. Ez a régóta várt, hiánypótló mű nem csupán a nyelvészet – különösen a szótörténet –, az irodalomtudomány és társtudományaik kutatóinak lehet majd fontos segédeszköze, hanem a nyelvtanulóktól a kultúrhistoria iránt érdeklődőig sokak számára lesz számos új információt nyújtó, izgalmas olvasmány.

A mintegy 15 000 címszó alá rendezett *Új magyar etimológiai szótár* befejezése jelentős eredményekkel szolgál a magyar és más nyelvek etimológiai és szótörténeti kutatásai számára, valamint jól tükrözi a mindenkorai magyar társadalom műveltségének változásait.

Jelenleg hiányzik egy korszerű, az aktuális nyelvállapot szókincsét tükröző magyar értelmező szótár. A kortárs, 21. századi magyar írásbeliségre jellemző szövegtípusok reprezentatív gyűjteményének létrehozásával megteremtjük annak a lehetőségét, hogy egy ezen alapuló korszerű, az aktuális nyelvállapot szókincsét tükröző értelmező szótárt hozzunk létre. Egy ilyen szótárt ma már valójában nem megírni kell: épp ezért a Lexikai tudásreprezentáció kutatócsoport azt kutatja, hogy a számítógépes nyelvészet, a nyelvtechnológia különböző szoftveres megoldásait miként foghatók hadra – a humán erőforrások leváltására. Azt gondoljuk, hogy bizonyítást nyer: a szótárírásnak régebben évtizedekre elhúzódó, fáradságos munkája automatikus eszközökkel jelentősen felgyorsítható. A magyar lexikográfia történetében még soha nem született ún. korpuszvezérelt értelmező szótár: a csoport azt vizsgálja, hogy jól használható szótárhoz juthatunk-e egy olyan módszerrel, amelynek fő vezérlő elve a gyakoriság. Az új online értelmező szótár társadalmi hasznosíthatóságának vonatkozásában a közép- és felsőfokú oktatás mellett a magyar nyelvet mint idegen nyelvet tanulók köre emelhető ki, illetve gyakran felmerül az igény egy hivatkozási alapként használható egynyelvű szótárra a jogalkotók részéről is.

Az intézet munkatársai tudományos ismeretterjesztő előadásokkal és rendezvényeken gyarapították a laikus érdeklődők ismereteit is a nyelvészet, különösen a lexikográfia kérdéskörében. 2021. szeptember 24-én a nagyszótári kutatócsoport „Hogyan készül a Nagyszótár?” címmel sikeres rendezvényt tartott a Kutatók Éjszakája programsorozat keretében. Ismeretterjesztő előadást tartott az intézet a Nagyszótár munkálatairól a Károli Gáspár Református Egyetem által rendezett ún. tréninghéten, illetve egy angol nyelvű workshopot szervezett az intézetben folyó lexikográfiai munkát bemutatva a Károli Gáspár Református Egyetem European Master in Lexicography (EMLex) képzésén tanuló nemzetközi hallgatóknak. 2021. november 9-én jeles író- és nyelvész vendég részvételével a Magyar Tudomány Ünnepe keretében került sor A magyar nyelv nagyszótára VIII. kötetének sajtóbemutatójára.

3. A kutatóhely legfontosabb hazai és nemzetközi eredményes K+F+I együttműködései az elmúlt évben

a. az állami költségvetési, a nonprofit és a kormányzati szektorral való kapcsolatok:

Új tudományos kapcsolat a Társadalomtudományi Kutatóközpont Gárdos Judit vezette Kutatási Dokumentációs Központjával, együttműködés a magyar ELSST (European Language Social Science Thesaurus) tezauszus készítésében. A kutatócsoport feladata a tezauszus nyelvészeti lektorálása volt.

Dr. Andréka Tamással, az Agrárminisztérium Jogalkotási Főosztályának főosztályvezetőjével elkezdtek a kutatóközpont és a minisztérium között együttműködési megállapodás szervezését. Ennek keretében a kutatóközpont felkérést kapna a magyarországi védett természeti értékek elnevezéseinek rendszerezésére, egységes szerkezetbe foglalására.

Az OSZK Gyűjteményszervezési Főosztályának vezetőjével, Matolay Katalinnal való együttműködés keretében a kutatóközpont megkapta a Magyar nemzeti bibliográfia 2001 és 2020 közötti, mintegy 340.000 rekord terjedelmű anyagát.

b. felsőoktatási intézményekkel való együttműködések: oktatás, utánpótlás-nevelés, doktori iskolák stb.:

Az intézet 5 munkatársa tanított felsőoktatási intézményben: összesen 7 gyakorlati és 1 elméleti kurzus tartása az ELTE BTK Magyar Nyelvtudományi és Finnugor Intézetében, 2 elméleti kurzus a Pázmány Péter Katolikus Egyetem Információs Technológiai és Bionikai Karán, 1 gyakorlati kurzus tartása az ELTE Eötvös József Collegiumban és 3 gyakorlati kurzus tartása a KRE Bölcsész- és Társadalomtudományi Karának Magyar Nyelvészeti Tanszékén.

Új együttműködési kapcsolatot létesített az intézet a Károli Gáspár Református Egyetemen működő EMLex-kurzus vezetőivel, az együttműködés keretében az EMLex (European Master in Lexicography) program keretein belül tanuló diákok számára lexikográfiai workshopokat tartottak a kollégák.

c. egyéb együttműködések, kapcsolatok.

A terminológiai, terminológiastratégia kutatások folytatásához olyan, a hazai tudományterületen létfonosságú szervezetekkel alakítottunk ki kapcsolatot, mint az Országos Fordító- és Fordításhitelesítő Iroda Zrt., a Magyarságkutató Intézet Nyelvtervezési Kutatóközpontja, az MTA Magyar Nyelvi Osztályközi Állandó Bizottsága, a Termini Magyar Nyelvi Kutatóhálózat, egyes kutatók személyén keresztül a Károli Gáspár Református Egyetem, valamint a Budapesti Műszaki és Gazdaságtudományi Egyetem. A kapcsolatok egyik célkitűzése a közös tudásépítés, amely közvetve társadalmi és gazdasági haszonnal jár – itt elsősorban a terminológiaharmonizáció és terminológiamenedzsment haszna lesz kézzel fogható.

Az intézet munkatársai tagként részt vesznek az MTA Nyelvtudományi Bizottsága mellett működő Szótári Munkabizottság, az MTA Magyar Nyelvi Osztályközi Állandó Bizottság, illetve az MTA Alkalmazott Nyelvészeti Munkabizottság munkájában.

4. A kutatóhely hazai vonatkozású pályázati adatai

ELKH Kiemelt Témák pályázat: A Magyar Terminológiai Koordinációs központ létrehozásának előkészítése: 2021. szeptember 1. – 2023. augusztus 31. 10.8 M Ft. Vezető kutató: Prószyk Gábor

NKFIH K 124127: Új magyar etimológiai szótár. Második ütem. 2017. szeptember 1. – 2022. augusztus 31. 32.594 Ft. Vezető kutató: Gerstner Károly

5. Kiemelkedő tudományos publikációk

Feldmann Ádám, Hajdu Róbert, Indig Balázs, Sass Bálint, Makrai Márton, Mittelholcz Iván, Halász Dávid, Yang Zijian Győző, Váradi Tamás. 2021. HILBERT, magyar nyelvű BERT-large modell tanítása felhő környezetben. In: Berend Gábor, Gosztolya Gábor, Vincze Veronika (szerk.): *MSZNY 2021, XVII. Magyar Számítógépes Nyelvészeti Konferencia*, Szegedi Tudományegyetem, Informatikai Intézet, Szeged. pp29-36. <http://real.mtak.hu/120856>

Horváth László. 2021. A tévedett : tévedt kettősség történetéhez. *Magyar Nyelvőr* 145: 3 pp. 357–363., 7 p.

Ittész Nóra (főszerk.): *A magyar nyelv nagyszótára VIII. Em–ez*. Nyelvtudományi Kutatóközpont, Budapest, 2021. www.nagyszotar.nytud.hu

Kalivoda Ágnes. 2021. Igeközös szerkezetek a magyarban. Doktori disszertáció, PPKE BTK Nyelvtudományi Doktori Iskola. Megjelenés/Fokozatszerzés éve: 2021. <http://real-phd.mtak.hu/1101/>

Kalivoda Ágnes. 2021. Az igeekötők produktív kapcsolódási mintái. *Argumentum* 17: 56–82. <https://doi.org/10.34103/ARGUMENTUM/2021/4>

Kristóf Ibolya. 2021. Forrás- és névjegyzékek: Kiegészítések és javítások a segédletekhez. In: Ittész, Nóra; Csengery, Kinga; Dömötör, Éva; Fiers, Márta; Gyenese, Ilona; Györffy, András; Kiss, Csilla; Simon, László (szerk.) *A magyar nyelv nagyszótára: 8. Em-ez*. Budapest, Magyarország : Nyelvtudományi Kutatóközpont. pp. 1003–1011.

Lipp, Veronika, Simon, László. 2021. Towards a new monolingual Hungarian explanatory dictionary: overview of the Hungarian explanatory dictionaries. *Studia Lexicographica* 15 (29), 83–96. <https://studialexicographica.lzmk.hr/sl/article/view/356>

Martelli, F., Navigli, R., Krek, S., Kallas, J., Gantar, P., Koeva, S., Nimb, S., Sandford P. B., E., Olsen, S., Langemets, M., Koppel, K., Üksik, T., Dobrovoljc, K., Ureña-Ruiz, R., Sancho-Sánchez, J.-L., Lipp V., Váradi T., Györffy A., Simon, L., Quochi, V., Monachini, M., Frontini, F., Tiberius, C., Tempelaars, R., Costa, R., Salgado, A., Cibej, J., Munda, T. 2021. Designing the ELEXIS Parallel Sense-Annotated Dataset in 10 European Languages. In: Kosem, I.; Cukr, M.; Jakubíček, M.; Kallas, J.; Krek, S.; Tiberius, C. (szerk.): *Electronic lexicography in the 21st century: post-editing lexicography*. Brno, Csehország: Lexical Computing. 377–395. <http://real.mtak.hu/129642/>

Pomázi Bence. 2021. A neki az ő élete szerkezet típusa a -nak/-nek rag grammatikálizációjában. *Magyar Nyelvőr* 145: 4 pp. 447–459., 13 p.

Pomázi Bence. 2021. Empirikus vizsgálati mód három névszói esetrag jelentéseinek elkülönítésére. *Jelentés és Nyelvhasználat* 8:1 pp. 31-46. , 16 p. <http://www.jeny.szte.hu/JENY-2021-PomaziB>

Pomázi Bence. 2021. Grammatikálizációs ösvények a -nak/-nek rag jelentéshálózatában. *Magyar Nyelv* 117:4. pp 446-465. <http://real.mtak.hu/136594/>

Sass Bálint, Mittelholcz Iván, Halász Dávid, Lipp Veronika, Kalivoda Ágnes. 2021. Magyar hadifoglyok adatainak orosz-magyar átírása és helyreállítása, és a szabadszöveges adatbázisok tulajdonságai. In: Berend Gábor, Gosztolya Gábor, Vincze Veronika (szerk.): *MSZNY 2021, XVII. Magyar Számítógépes Nyelvészeti Konferencia*, Szegedi Tudományegyetem, Informatikai Intézet, Szeged, 2021, 39–51. <http://real.mtak.hu/121986/>

KUTATÓKÖZPONT NEVE:
Nyelvtudományi Kutatóközpont
KUTATÓINTÉZET NEVE:
Nyelvtechnológiai és Alkalmazott Nyelvészeti Intézet
Cím:
1068 Budapest, Benczúr u. 33.
Postafiók címe:
1394 Budapest, Pf. 360
telefonszám:
(1) 342-9372
honlap:
www.nytud.hu
e-mail:
linginst@nytud.hu

I. A Nyelvtechnológiai és Alkalmazott Nyelvészeti Intézet kutatási eredményei a 2021. évre vonatkozóan

Nyelvtechnológiai Kutatócsoport

Tématerület: Digitalizáció: informatika
Emberi erőforrás: nyelv

Sikeresen zárult a Nyelvtechnológiai Kutatócsoport vezetésével folyt [MARCELL projekt](#), melynek keretében elkészült egy többnyelvű, jogi szövegeket tartalmazó elemzett korpusz magyar nyelvű alkorpuszának végleges változata. Az alkorpusz szövegeinek feldolgozására szolgáló alkalmazás ugyancsak elkészült, Dockeren keresztül futtatható.

Sikeresen megtörtént az Infra2020 pályázaton elnyert támogatásból egy világszínvonalú szuperszámítógép nemzetközi közbeszerzési eljárás keretében történő beszerzése és üzembe helyezése. A Magyarországon mindmáig egyedi berendezés új távlatot nyit a kutatócsoport számára a neurális technológián alapuló természetes nyelvfeldolgozás és a mesterséges intelligencia kutatás terén.

Általános célú neurális nyelvmodellek építése és kísérletek a nyelvmodellekkel

2021-es céljainkkal összhangban számos nyelvmodell készült, és ezekkel különféle kísérleteket végeztünk. Kis méretű magyar nyelvű ELECTRA és Electric modellek szolgálták különböző osztályozásos modellek alapjául, többek között szentimentanalízis, névelemfelismerés, főnévi csoport felismerés és extraktív összefoglalás céljából.

Különböző méretű magyar nyelvű BART modellek is készültek, amelyek többfajta osztályozásos és szöveggeneráló feladatokra lettek finomhangolva: szentimentanalízis, absztraktív összefoglaló és versgeneráló modelleket tanítottunk, összefoglaló generálás területén „state of the art” eredménnyel a magyar nyelvre.

Továbbá elkészült egy kisméretű magyar nyelvű GPT-2 modell is, szintén elsősorban osztályozásos és szöveggeneráló feladatok megoldására. A következő lépésben szentimentanalízis, névelemfelismerő, főnévi csoport felismerő, absztraktív összefoglaló, versgeneráló és hírgeneráló modellek készültek. Létrejött a magyar nyelvre az első GPT alapú hírgeneráló modell. Némelyik feladathoz demó alkalmazás is készült.

Nagyobb neurális hálók tanítása komolyabb számítási kapacitást igényel, ezért megtörtént az NVIDIA DGX-100 szuperszámítógép infrastruktúrájának kiépítése és beüzemelése, valamint a SLURM feladatkezelő telepítésének előkészületei is.

Benchmark teszt-adatbázis készítése

A HuLU névre keresztelt *benchmark* adatbázis (Hungarian Language Understanding and Evaluation Benchmark Kit) létrehozását jelentő első lépésben 4 korpusz készült el. A korpuszok röviden a következők: (1) A HuCOLA egy majdnem 10 000 magyar nyelvű mondatot, és azok elfogadhatósági

ítéletének címkéjét tartalmazó korpusz. (2) A HuSST 11 000 mondatot és azok szentimentcímkéjét tartalmazza. (3) A HuCOPA 1000 egységében egy premisszához két mondat közül az van rendelve, amelyikkel valószínűbb az ok-okozati kapcsolat. (4) A HuRC korpuszban 88 000 hírszöveg található, melyeknél az utolsó bekezdés egy eltakart névelemének az azonosítása a feladat.

A fentiekén kívül előállt a Winograd-sémák magyar fordítása is. Az elkészült, tanító-, validációs és tesztalmazra osztott adatbázisok elérhetőek a githubon (<https://github.com/nytud>) és a huggingface-en (<https://huggingface.co/NYTK>). Ugyanitt elérhetőek a magyarra fordított Winograd-sémák is.

Adatinfrastruktúra építése gondozott szövegek OCR hibáinak javításával

A nemzetközi szakirodalom áttekintése alapján kiválasztásra került a megfelelő élvonalbeli módszer, amely egy BERT alapú detekciós és egy NMT alapú encoder-decoder korrektúrázó neurális hálózat kombinálásából áll össze. Ezeknek a modelleknek a tanításához olyan párhuzamos korpuszra van szükség, amely egyfelől az OCR hibákat tartalmazó szövegből, másfelől ugyanennek a szövegnek az OCR hibák nélküli változatából áll össze. Ilyen azonban magyar nyelven nem volt megfelelő mennyiségben, ezért ezt létre kellett hozni. A korpuszra vonatkozó előzetesen megszabott kritériumoknak a Jókai és Mikszáth életművek feleltek meg, ezekből találtunk mind OCR-ezett, mind begépelte verziót az Arcanum adatbázisban. Sajnos sok esetben a regények a begépelte, illetve OCR-ezett változatai kiadásaikban eltértek, így bár a szöveg alapvetően megfelelt a kritériumoknak, további utómunkát igényelt az eltérő kiadások közötti különbségek kiszűrése. Az ilyen módon illesztett anyagokból előálló *kézzel ellenőrzött* párhuzamos tanítóanyag már alkalmas volt a modell fejlesztésére: segítségével betanításra került egy mély-neurális hálózatokon alapuló javító-rendszer. Az OCR utófeldolgozás két fázisból áll, az első lépésben az OCR hibákat detektáljuk egy BERT-alapú detekciós modellel, míg a második lépésben az így felismert OCR hibák javítása történik meg egy NMT alapú encoder-decoder neurális modell segítségével. A tanított/tesztelt modellek eredményei egy cikk alapjául szolgáltak.

A nagyszabású [ELE projekt](#)ben végzett adatgyűjtés során több mint 500 nyelvi erőforrást tartalmazó adatbázis készült el. Ez nemcsak azt mutatja be, mely kategóriákban van a legtöbb erőforrás, hanem azt is, hogy milyen erőforrások hiányoznak a magyar NLP fejlődéséhez, amely meghatározhatja a magyar NLP kutatás-fejlesztési irányát. Az elkészült adatbázis a következő linken kérdezhető le: <https://live.european-language-grid.eu/catalogue/search/Hungarian>.

A kutatócsoport belső munkaszervezését érintő fejlesztések a következők voltak: (1) a dokuwiki alapú open source dokumentációs célú wikipédia adoptálása egy fejlettebb és könnyen kezelhető, saját üzemeltetésű rendszerbe, melyet a kutatók és üzemeltetők egyaránt használhatnak, és amely weben keresztül elérhető: <https://wiki.nlp.nytud.hu/>. (2) Egy taiga alapú open source, saját kezelésű és teljes értékű projektkezelő rendszer is elkészült. Itt a munkatársak követhetik az egyes projektek és részfeladatok alakulását: <https://projects.nlp.nytud.hu/>. (3) Saját Docker konténerreink tárolását is megoldottuk egy saját kezelésű docker registry létrehozásával: <https://docker.nlp.nytud.hu/>.

További eredmény, hogy egy kutató PhD-fokozatot szerzett a Pázmány Péter Katolikus Egyetem Nyelvtudományi Doktori Iskolájában *The Right Edge of the Hungarian NP. A Computational Approach* című disszertációjával.

A kutatócsoport egy tagját a MANYE tiszteletbeli tagjává választotta.

Többszempvűségi és Edukációs Nyelvészeti Kutatócsoport

Tématerület: Emberi erőforrás: társadalom, bölcsészet, nyelv.

Az eredményeikből 2021-ben 30 magyar és 3 idegen nyelvű publikáció, köztük 1 tanulmánykötet és további 4 önálló kiadvány született. Kiemelendő eredmény a Kontra Miklós és Borbély Anna szerkesztésében megjelent *Tanulmányok a budapesti beszédről a Budapesti Szociolingvisztikai Interjú alapján* című kötet (Gondolat Kiadó: Budapest. 355 l.), melyben a kutatócsoport munkatársai 6 tanulmánnyal szerepelnek. Ugyancsak kiemelten fontos publikációs eredmény az Edulinguanet portál létrehozása az ott hozzáférhetővé tett kutatási eredményekkel s az elindított *EDULINGUA kiadványok* című kiadványsorozattal.

Lezárult az MTA és a Román Akadémia kétoldalú nemzetközi egyezményen alapuló közös kutatási projektje együttműködése keretében *a magyarországi román dialektológiai szótárát előkészítő munkálat*. Befejeződött a szótár módszertani szempontjainak kidolgozása; a meglévő nyelvi adatok rendezése és új adatok terepmunkával történő gyűjtése, az adatok korpuszá alakítása. Az

eredményekre épülő, 2021-ben benyújtott új pályázat a szótár szócikkeinek megírását, valamint az ehhez szükséges további adatgyűjtéseket célozzák.

Sikeresen zárult a 2016-ban az MTA Tantárgy-pedagógiai Kutatási Programja keretében megkezdett *NyelvEsély* című munkát, melynek számos korábbi és 2021-ben született új eredménye elérhetővé vált a projekt erre kialakított <http://edulingua.net/edupage/> felületén. A hazai és nemzetközi szakértői bírálatokra épülő záró értékelés a kutatócsoport munkáját a legmagasabb (6.) szintre, a tantárgy-pedagógiai kutatások nemzetközi kutatási élvonalába sorolta.

A jelnyelvi és romani modulok kutatási és fejlesztési eredményei 2021-ben jelentősen bővültek és véglegesedtek, hozzáférhetővé váltak az Edulinguanet portálon. A 2021-es évben kiemelendő új eredmények maga a portál, a kétnyelvű, multimodális Tanári-iskolai instrukciók tára, Online Oktatási fogalomtár, a Siketiskolai mérőeszköz-csomag, valamint az EDULINGUA kiadványok sorozat.

A projekt 2021-es évének egyik nemzetközi kitekintésben is új eredménye a kétnyelvű tanulástámogató [Tanári-iskolai instrukciók tára](#), amely a tanórákon, tankönyvekben és mindennapi oktatási-nevelési helyzetekben gyakran használt pedagógusi és szülői instrukciót (utasításokat, kifejezéseket) tartalmaz magyarul és jelnyelven. Az adatbázis a halló, jelnyelvet nem vagy kevésbé ismerő szülők és pedagógusok számára jelent hasznos segítséget.

Ugyancsak jelentős eredmény az [online Oktatási fogalomtár](#), amely öt kiemelten fontos iskolai tantárgy, tudásterület 1998 terminusát és 95 db informatikai definíciót tartalmaz jelnyelven.

Siketiskolai mérőeszköz-csomag

A bemeneti és kimeneti kompetenciák mérésére és az iskolai oktatás eredményességének feltérképezésére szolgál az újonnan kifejlesztett, [Siketiskolai mérőeszköz-csomag](#), amelynek létrejöttét egy, az oktatási-nevelési folyamat minden szereplőjére irányuló több évig tartó komplex kutatássorozat előzte meg. A mérőeszköz-csomagot megalapozó jelnyelvi tesztelés első fázisáról, a tesztelés módjáról [kisfilm](#) is készült. A mérőeszköz-csomagot leíró és a főbb eredményeket bemutató 2021-es rövidebb publikáció online formában is elérhető.



1. kép: Jelnyelvi tesztelés és mérés

A 2021-ben indított [EDULINGUA kiadványok sorozat](#) különálló, ugyanakkor tematikus egységekbe rendeződő füzeteket foglal magában: hazai és nemzetközi kutatási beszámolók, saját elemzések és fejlesztések, egyéb szakmai anyagok, bibliográfiák, policy dokumentumok, ajánlások, eredmények, módszertani jó gyakorlatok, konferencia-beszámolók, infografikák stb.

Az NYTK konzorciumvezetésével és a kutatócsoport szakmai irányításával 2021 júniusában új ERASMUS+ nemzetközi kutatási projektum indult magyarországi, romániai és szlovákiai partnerintézmények együttműködésével *Romani Digitális Tudástér (DigiDrom)* címmel. A megkezdett munkát fő célkitűzései között szerepel a romani (cigány) gyermekek, fiatalok oktatáshoz való hozzáféréseinek, esélyegyenlőségének támogatása. 2021-ben lezárultak a projekt előkészítő szakaszának munkálatai. Kutató- és gyűjtőmunka folyt az oktatási helyzetelemzéshez, hazai és külföldi két- és többnyelvű romani szótárak, mesék, mesekönyvek gyűjtése, elemzése, adatbázisba rendezése, digitalizálása, valamint a régiós összehasonlító vizsgálat megalapozása zajlott. Elkezdődtek a romani (lovar, cerhar, kárpáti változatokra épülő) szótár munkálatai. Rendszeres konzultációk folytak az NYTK Lexikológiai Intézete együttműködésbe bevont lexikográfus és nyelvtechnológus munkatársaival. Ugyancsak megkezdődtek a Digitális mesetár gyűjtőmunkálatai mindhárom partnerországban.

A *Longitudinális kutatás a román–magyar kétnyelvűségről (1990–2000–2010–2021)* című kutatás 4. terepmunkája keretében 2021-ben összesen 34 személlyel készült interjú, közülük 21 személlyel 1990-től immár negyedik alkalommal az elemzésből egy tanulmány született *Nemzetiségi utcanevék Magyarország nyelvi tájképében* címmel.

Elkészült a *Két- és többnyelvűség, jelnyelvek és siketoktatás. Elméletek, nemzetközi gyakorlatok és következtetések – tudománytörténeti elemzés*, amely a változatos nyelvi szocializációs háttérű siket gyermekek fejlődésére vonatkozó elméleti kutatások között hiánypótlónak számít.

Interakciós Rítus Kutatócsoport

Tématerület: Emberi erőforrás: nyelv.

A kutatócsoport sikeresen, „kiváló” minősítéssel zárta le 2016-ban indult MTA Lendület projektumát. 2021-re fő feladatuk a korábbi kultúraközi modell továbbfejlesztése volt az összehasonlító pragmatika területén. E vállalat sikerült: legfontosabb új eredményük egy olyan összehasonlító pragmatikai modell kidolgozása lett, amely segítségével a nyelvhasználat rendszerszerűen kutatható gyakorlatilag bármilyen adattípusban, illetve nyelvkultúrában. A modell a nyelvhasználat három fő egységén – a kifejezések, beszédaktusok és diskurzus – keresztül közelíti meg a vizsgált adatokat. A modell hiánypótló jellegű: noha az összehasonlító pragmatika a pragmatikai kutatások között az egyik legnagyobb tudományos hatásokkal bíró terület, az 1987-es Blum-Kulka, House és Kasper modell óta gyakorlatilag nem jött létre új, átfogó jellegű összehasonlító pragmatikai modell. A modell alkalmazásának fontossága abban rejlik, hogy empirikus megközelítésük új betekintést kínál a fent felsorolt területekbe, amelyeket sokszor rendkívül ideologikus módon kutatnak. Példának okáért, a System D1 folyóiratban megjelent nyelvoktatással foglalkozó cikkükben arra világítottak rá, hogy a köszönés látszólag egyszerű beszédaktusa milyen nehézségeket jelenthet haladó angoltanuló kínai diákoknak. E nehézségek vizsgálata során arra is rávilágítottak, hogy azok csak összehasonlító pragmatikai eszközökkel írhatók le replikábilis módon. Az új összehasonlító pragmatikai modellt több területre, így a nyelvoktatásra, a fordításelméletre, illetve a nyelv és politika vizsgálatára is alkalmazni kezdték, ezzel fektetve le 2022-ben induló TKP projektjük (ld. lentebb) alapjait. Kutatásaikat számos D1, illetve Q1 besorolású szakfolyóiratcikk formájában publikálták.

A kutatócsoport három panelt szervezett az International Pragmatics Association 17. Nemzetközi Konferenciáján (Winterthur, Svájc, 2021 július), a kutatócsoport vezetője 4 nemzetközi plenáris előadást tartott nemzetközi konferenciákon, valamint további 4 meghívott előadást tartott nemzetközi eseményeken. A kutatócsoport tagjai 11 nemzetközi előadást is tartottak. A csoport egyik munkatársa 2021 során elkészítette doktori értekezését a PPKE Nyelvtudományi Iskolájában, amelynek végleges formáját 2022 márciusában nyújtja be. Kutatása jelentősen hozzájárult a tudomány és társadalom közötti párbeszédhez: folytatta korábban megkezdett, a szlovákiai magyar szórványban végzett munkáját, melynek célja magyar anyanyelvű gyermekek kultúraközi kommunikációs kompetenciájának fejlesztése.

A kutatócsoport jelentősen tudta növelni kutatási támogatásait. 2021-ben sikerrel pályáztak a Tématerületi Kiválósági Program 2021, Nemzeti Kutatások Alprogramra, amelyet a Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal írt ki. Az elnyert támogatás összege 161.655.300 forint.

A kutatócsoport vezetőjét a 2021-es év során az *Academia Europaea* rendes tagjává választották (Oriental Studies Section).

Nyelvhasználati és Nyelvi Tanácsadási Kutatócsoport

Tématerület: Emberi erőforrás: nyelv.

A kutatócsoport munkáját a közcélú nyelvi szolgáltatások működtetése és az azokhoz kapcsolódó kutatások kettőssége jellemzi. A csoport kutatásai a nyelvi szolgáltatásokhoz kapcsolódnak, elsődlegesen az azokból nyert adatok feldolgozására, rendszerezésére épülnek.

Kutatási eredmények

A nyelvi közönségszolgálathoz érkező nyelvi megfigyelések, kérdések olyan kutatási anyagot szolgáltatnak, amely megmutatja a beszélők nyelvi problémáinak egy sajátos szegmensét. 2021-ben folytatódott a nyelvi tanácsadó szolgálat adatainak a feldolgozása, rendszerezése, a levéladatbázis címkéinek strukturálása. 2021 júliusában a nyelvi közönségszolgálat munkatársai kérdőíves vizsgálatba kezdtek annak feltérképezésére, hogy milyen társadalmi háttérű személyek fordulnak közönségszolgálathoz, milyen motivációkkal. Az adatgyűjtés folyamatosan zajlik, a rövid, öt kérdésből álló kérdőívet a kérdésre adott válaszlevélben küldik el a munkatársak az adatközlőkhöz.

2021-ben folytatódott a mai magyar levelezési, e-mailezési gyakorlatok nyelvi sajátosságainak, nyelvi problémáinak vizsgálatára irányuló kutatás. A felsőoktatási intézményekben 2020 márciusában elrendelt digitális távolléti oktatás a hallgató-oktató írásbeli kapcsolattartásra is hatással volt. 2021-ben megtörtént az átalakuló felsőoktatási e-mailezési gyakorlatnak a koronavírus-világjárvány kezdeti szakaszában történő vizsgálata. Az eredmények szerint az írásbeli kapcsolattartási gyakorlatokban a szolidaritás, a társas közeledés kifejezésének eszközei figyelhetők meg, főként az elektronikus levelezés jókívánásaiban, a levelek elkészítő részében, záróformuláiban.

- Domonkosi Ágnes – Ludányi Zsófia 2021. „Üdvözlettel, jó munkát és hozzá jó egészséget kívánva”. A felsőoktatás e-mailezési gyakorlatai a járványhelyzet idején. In: Istók Béla – Simon Szabolcs (szerk.) *Online oktatás – kontaktoktatás. Edukációs folyamatok és a Covid19*. Komárom, Szlovákia: Selye János Egyetem. 11–23. <http://real.mtak.hu/138836>
- Domonkosi Ágnes – Ludányi Zsófia 2021. Address forms and age roles in emailing practices between university students and their teachers. In: Uberman, Agnieszka – Trinder, Magdalena (eds.): *Text Sentence Word: Studies in English Linguistics IV*. Rzeszów: Wydawnictwo Uniwersytetu Rzeszowskiego. 26–38. <http://real.mtak.hu/138846/>

A kutatócsoport névtani tevékenysége igen összetett és szerteágazó: a jól körülhatárolható napi, szoros határidőkkel rendelkező közhasznú feladatok tudományosan hiteles ellátása mellett folyamatosan feldolgozza az összegyűlt adatokat. Folytatódott a meglévő névadattárak archiválása, szocioonomasztikai kutatása. A 2013-ban létrehozott Utónévkereső adatai frissítve lettek a 2021 decemberéig keletkezett adatokkal. Elkészült a névkérelmeket rögzítő adatbázis, amely a 2020–2021-es év adatait tartalmazza, a tervek között szerepel a korábbi évek adatainak feltöltése az adatbázisba.

Romani és beás interakciókutatás

Tématerületi megjelölés: 3. Emberi erőforrás: Társadalom, nyelv

Folytatódott a K 129378 sz. NKFIH-projekt keretében a romani interakciós adatok gyűjtése és feldolgozása: 36 online kapcsolati rítus dokumentálása fejeződött be. 8 fejezet készült el a gábor roma közösségek nyelvhasználatát vizsgáló monográfiából, amely a viszonyulás jelzés interakciós eszközeit és az azok kontextusául szolgáló kapcsolati rítusokat elemzi. Folytatódott a beás nyelvi interjúk feldolgozása, és a munkatárs a beás nyelv változatosságát elemző disszertációján dolgozott. Az eredmények egy részét publikációk és konferencia-előadások formájában tették közzé.

II. Tudományszervezési eredmények, események

1. Rövid általános ismertető az Intézet kutatóhely tudományos munkáját illetően

A Nyelvtchnológiai és Alkalmazott Nyelvészeti Intézet sajátos profilt képvisel a Nyelvtudományi Kutatóközpont intézetei között. Kutatási tématerületei az alkalmazott nyelvészet egész széles spektrumát képviselik a szociolingvisztika, pragmatika és kétnyelvűség vizsgálatán át a nemzetközi mércével mérve is egyedi jelnyelvi kutatásokig. A Nyelvtchnológiai Kutatócsoport kutatási területe túl is nyúlik a hagyományos bölcsész stúdiumokon, és elsősorban a digitalizáció fogalmával írható le.

Az elméleti kutatások mellett az Intézet valamennyi kutatócsoportja igen aktívan foglalkozik *kutatási infrastruktúra* építésével, amit számos nagy adatbázis építése és üzemeltetése jellemez, melyeket a szakmai és szélesebb közösség javára végez, a Top-50 infrastruktúrába kiválasztott Hun-CLARIN és más portálok. A kutatási tevékenység mellett az Intézet profiljában jelentős szerepet tölt be a *közcélú tevékenység*, melyet az Intézet kutatócsoportjainak több mint a fele végez (l. alább).

2. A tudomány és a társadalom kapcsolata, társadalmi hasznosság

A *Nyelvhasználati és Nyelvi Tanácsadási Kutatócsoport* munkatársai kutatómunkájuk mellett közcélú feladatokat is elláttak a nyelvi szolgáltatások végzésével. A nyelvi közönségszolgálatot, a folyamatosan működő levelezési és a heti két napon működő telefonos tanácsadást, valamint a nyelvi szakértői véleményezéseket 2021 második felében három új munkatárs segítette elő. A közönségszolgálat munkatársai a tanacs@nytud.hu címen folyamatosan fogadták a nyelvi, helyesírási kérdéseket, és folyamatosan válaszoltak rájuk. A kérdések jelentős része (81%) helyesírási jellegű volt, a többi pedig nyelvhasználati, szókészletteni, stilisztikai tárgyú nyelvi probléma. Az e-mailes megkeresések száma 2021-ben 923 volt, ez az előző évhez képest valamivel kevesebb (2020-ban ez a szám 1132 volt). A közönségszolgálat munkatársai 6 db szakvéleményt készítettek, és 35 darab emléktábla nyelvi ellenőrzését végezték el.

A kutatócsoport speciális, törvényi kötelezettségből adódó feladata az utónév-szakvéleményezés. A kutatócsoport névtani munkacsoportja a Kutatóközpontban erre a célra létrehozott Utónévbizottsággal együttműködve látta el a hatósági anyakönyvezésre beterjesztett új utónevek nyelvi szakvéleményezését és a bejegyzésre alkalmasnak tartott nevek listájának gondozását. A névkérelmek elbírálását, szükség szerint további szakértők bevonását, az illetékes hatóságokkal való kapcsolattartást és a névlista bővítését 2021 második felében egy új munkatárssal kiegészülve végezték. A munkatársak 2021-ben közel 580 keresztnévkérelem, 31 családnév és 4 névthonosítás estében készítettek szakvéleményt. Az utónévlistákat havonta folyamatosan frissítették.

A kutatócsoport munkatársai [két programmal](#) is részt vettek a Kutatók Éjszakája rendezvényein:

- „Csókolom. It lehet hejes írást kérdezni?” Nyelvi problémák és nyelvhasználati segédeszközök
- Névadási szokásainkról. Betekintés a Névtani munkacsoport munkájába

A csoport több tagja tartott ismeretterjesztő előadásokat:

- Parádfürdő, 2021. július 2. Számítógépes eszközök az írás támogatására, avagy különírás-egybeírás automatikusan. Gyöngyösi Berze Nagy János Gimnázium Természettudományos Önképzőköre.
- Paks, 2021. október 21. Névadási szokásainkról. Vak Bottyán Gimnázium. MTA Alumni Program.
- Eger, 2021. november 10. A könnyen érthető kommunikáció mozgalma mint nyelvmenedzselési folyamat. Eszterházy Károly Katolikus Egyetem.

Megjelent egy ismeretterjesztő cikk a hivatalos e-mailek megszólításaiával kapcsolatos nyelvi problémákról.

- Domonkosi Ágnes – Ludányi Zsófia 2021. Nyelvi problémák a hivatalos e-mailek megszólításaiiban. *Édes Anyanyelvünk* 43/5: 18.

Médiamegjelenések:

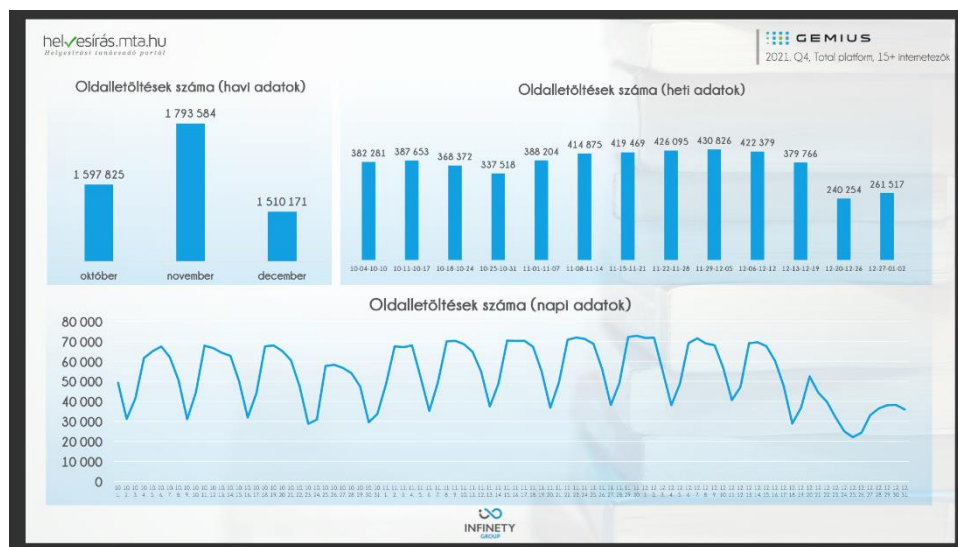
Nyelvhasználati kérdésekkel kapcsolatban:

- [Klubrádió Galaxis Kalauz](#)
- Kossuth Rádió, Édes anyanyelvünk – ([2021. július 14.](#), [2021. július 15.](#)).

Utónévügyekkel kapcsolatban:

- [Hír 24](#)
- [MTV5 Én vagyok itt!](#) (1.01 óra)
- [Klasszik reggeli 2021. 01. 25.](#)
- [Pandémia és az Éden keresztneveket is megpróbálták anyakönyveztetni 2020-ban](#)

A *Nyelvtechnológiai Kutatócsoport* üzemelteti a továbbra is nagy népszerűségnek örvendő *helyesiras.mta.hu* helyesírási tanácsadó portált. 2021-ben összesen 2,94 millió látogató (26,3%-uk visszatérő látogató) több mint 7,4 millió lekérdezést indított. A lekérdezések átlagos száma meghaladja a havi 300 000-et, nyaranta a látogatások száma – érthető módon – visszaesik kb. 250 000 felhasználóra, tavaszi időszakban azonban a 400 000-et is átlépi. A felhasználók 50%-a asztali gépről, 48%-a azonban már okostelefonról kereste fel az oldalt és átlagosan csak 3 percet töltenek az oldalon, köszönhetően a gyors működésnek. Elmondható, hogy a 2013 óta üzemeltetett egyik legsikeresebb szolgáltatásunkat továbbra is széles tömegek használják, a világ minden tájáról az Egyesült Államoktól Azerbajdzsánig. A Magyar Nemzeti Szövegtárban 2021-ben 144 125 lekérdezést futtattak.



2. kép: A helyesiras.mta.hu lap webes forgalmi adatai

A *Többnyelvűségi és Edukációs Nyelvészeti Kutatócsoport (TENyEK)* a korábbi és folyó kutatási projektjei során kialakított szakmai együttműködési hálózatának és a hazai felsőoktatásban (ELTE BTK, MNyFI; KRE TFK, BTK) végzett oktatási, program- és rendezvényszervezői tevékenységének, a jelnyelvi, romani, nemzetiségi szociolingvisztikai, alkalmazott nyelvészeti, tanárképzési és mentorálási vonatkozású kapcsolatai kiaknázásával igyekszik kiszélesíteni a társadalom különböző célcsoportjainak elérését, további szakértők, érintettek, érdeklődők bevonását. Részt vesz különböző programsorozatokon ismeretterjesztő rendezvények, programok és konferenciák szervezésével és ilyen eseményeken való részvétellel is, disszeminálva a kutatások és fejlesztések eredményeit.

2021-ben két önálló tudományos ismeretterjesztő rendezvény, valamint egy online tudományos konferenciát követő jelenléti ismeretterjesztő workshop valósult meg:

Kutatók Éjszakája – TENyEK Interaktív kitelepülés – 2021. szeptember 24.

A 2021. évi rendezvényen a kutatócsoport interaktív kitelepüléssel vonta be a kutatóhely, valamint a többnyelvűségi kutatási témák és fejlesztések iránt érdeklődőket. Az interaktív többnyelvűségi bazar keretében az érdeklődők megismerkedhettek a kutatócsoport Esély-projektjei során végzett fejlesztésekkel, valamint ismeretterjesztő betekintést nyerhettek a többnyelvűséget, a magyarországi nemzetiségeket, a siket és cigány közösségeket, a jelnyelvet és a romani nyelvet övező legfontosabb tények, hitek és tévhitek témájába is.

ELTE Kampusznyitó fesztivál – Esélyegyenlőségi sátor. Budapest, ELTE, 2021.szeptember 8–9.

A kutatócsoport előadásokkal és interaktív tevékenységeken keresztül népszerűsítette a TENyeK kutatásait, különösen a jelelő közösségeket, a jelnyelveket és a magyarországi nemzetiségeket érintőket, továbbá a TENyeK által támogatott és az ELTE által működtetett Magyar jelnyelv minorspecializációt.

A NyelvEsély projekt záró rendezvényéhez kapcsolódó digitális tanítási-tanulási workshop, Budapest, NYTK, 2021. augusztus 27.

A projektet záró online konferencia tudományos előadásait és fejlesztési eredményeinek interaktív bemutatását követően került megrendezésre tanárok, jelnyelvoktatók és más érdeklődők számára az a workshop, melynek keretében a résztvevők gyakorlatban sajátíthatták el a projekt tanulástámogató digitális anyagaihoz, feladataihoz is használt szoftverek, applikációk működését és tanórai felhasználását.

Médiamegjelenések

A Tanulmányok a budapesti beszédről című kötet (szerk. Kontra Miklós és Borbély Anna) *online könyvbemutatója*. 2021. október 5. 17 órai kezdettel a *Gondolat Kiadó Facebook oldalán* élőben bemutatták a szerkesztők a kötetet. Az adás megtekinthető [ezen a linken](#) (letöltés ideje: 2022.03.13.)

Rádióinterjú a *Jelenlét, Roma Kulturális Magazin* című műsor részeként, 2021.04.11-én, a 21:47 másodperctől a TENyeK kutatója beszélt munkásságáról és a Többnyelvűségi és Edukációs Nyelvészeti Kutatócsoporttal való együttműködéséről a *Kossuth Rádióban*.

Az adás elérhető [ezen a linken](#) (letöltés ideje: 2021.04.13.)

Rádióinterjú az *Ötös – Kultúrcsütörtök Warholik Zoltánnal* című műsor részeként, a *Klub Rádióban*, 2021. április 08. 10:00 órakor. A Többnyelvűségi és Edukációs Nyelvészeti Kutatócsoport tagja beszélt a hodászi helyzetképről, hagyományokról, a Tájházzal. Az adás elérhető [ezen a linken](#) (letöltés ideje: 2022.03.13.)

Rádióinterjú a *Három szólamra* című műsorban, az *MR4 Nemzetiségi Rádió* műsorának részeként, 2021.06.07. 12:03-tól. A Többnyelvűségi és Edukációs Nyelvészeti Kutatócsoport tagja beszélt munkásságáról. Az adás elérhető [ezen a linken](#) (letöltés ideje: 2022.03.13.)

Rádióinterjú a *Jelenlét, Roma Kulturális Magazin* című műsor részeként, 2021.03.21-én. Riport a kutatócsoport tagjával a tudományos és közösségszervező munkájáról. Az adás elérhető [ezen a linken](#) (letöltés ideje: 2022.03.13.)

A kutatócsoport egyik tagja rádióinterjúban (podcast) szerepelt, melyet a Károli Gáspár Református Egyetem Pedkaszt című műsorában hallhattak az érdeklődők. Az adás megtekinthető [ezen a linken](#) (letöltés ideje: 2022.03.13.).

3. A kutatóhely legfontosabb hazai és nemzetközi eredményes K+F+I együttműködései

Az intézet nemzetközi K+F+I együttműködéseivel kapcsolatban fontos hangsúlyozni, hogy a Nyelvtechnológiai Kutatócsoport, valamint a Többnyelvűségi és Edukációs Nyelvészeti Kutatócsoport európai ismertséggel bírnak, míg az Interakciós Rítus Kutatócsoport széleskörű nemzetközi együttműködést épített ki nemcsak az Egyesült Királyságban és Németországban, hanem Indiában és Kínában, valamint Japánban. A Nyelvtechnológiai Kutatócsoport immár több mint húsz éve aktív és elismert szereplő az európai nyelvtechnológiában, az utóbbi tíz évben regionális vezető szerepet játszik, és egymás után három projektben is egy héttagú konzorcium élén koordinátori szerepet töltött és tölt be. A projektek támogatási értéke gyakran nem tükrözi azt az értéket, amelyet az általuk biztosított ismertség és elismertség jelent. A Többnyelvűségi és Edukációs Nyelvészeti Kutatócsoport is koordinátori szerepet tölt be egy Erasmus+ pályázatban.

a. az állami költségvetési, a nonprofit és a kormányzati szektorral való kapcsolatok

Folyamatos nyelvi tanácsadói, nyelvi lektori munkakapcsolat kormányzati és egyéb állami szervekkel, minisztériumokkal és háttérintézményekkel: Köztársasági Elnöki Hivatal, Országgyűlés Jegyzői

Irodája, Nemzeti Adó- és Vámhivatal, Budapest Főváros Önkormányzata, Főpolgármesteri Hivatal, Nemzeti Emlékezet Bizottsága, Külgazdasági és Külügyminisztérium.

A névügyi szakvéleményezés tekintetében többféle állami intézménnyel folyik együttműködés. Folyamatos a kapcsolat és szakmai állásfoglalás adása a családnév-változtatásokban a Budapest Főváros Kormányhivatala Állampolgársági és Anyakönyvi Főosztály Névváltoztatási és Anyakönyvi Osztály munkatársainak, illetve szintén szakmai segítségnyújtás, állásfoglalás adása a névhonosításkor a Budapest Főváros Kormányhivatala Állampolgársági és Anyakönyvi Főosztály munkatársának. 2021-ben 35 szakvélemény készült ebben a témában.

A Magyarországon élő nemzetiségek jogainak védelmét ellátó biztoshelyettes és hivatala jogvédelmi munkájában a magyarországi nemzetiségeket érintő nyelvi, nyelvi jogi és oktatási kérdésekben rendszeresen támaszkodik a *Többszínű és Edukációs Nyelvészeti Kutatócsoport* kutatási eredményeire és tanácsadó tevékenységére. A kormányzati szektorral való együttműködés keretében a kutatócsoport vezetője szakértőként részt vett a biztoshelyettes által kiadott *A nemzetiségi nyelven való közösségi névhasználatról szóló 4/2021. számú elvi állásfoglalás* létrejöttében.

A Magyarországi Román Országos Önkormányzat számára elkészült és a MROÖ hivatalos honlapján elérhetővé vált az alapelveket, valamint a magyar anyakönyvbe bejegyezhető román férfi és női utónevek listáját, a nevek magyar átírását és megközelítő magyar változatát tartalmazó *Magyarországi román utónévjegyzék*.

Új együttműködési megállapodás született a *Károli Gáspár Református Egyetem Tanítóképző Főiskolai Karával*, melynek célja a kutatási és módszertani eredmények kölcsönös megosztása, valamint a kutatócsoport eredményeinek és fejlesztéseinek közvetlen bevonása a tanítóképzésbe és továbbképzésbe.

A TÁRKI által koordinált, „*A fogyatékos emberek kirekesztésének változó és új mechanizmusa*” (NKFIH) projektjében kollégánk a tanácsadó testületben szakértőként segítette a siket adatközlők kiválasztását, a releváns kérdések azonosítását, valamint ezek érthetőségének ellenőrzését a célcsoport számára.

Két munkatárs részt vett a SINOSZ jelnyelvi tolmácsolás képzésében. Az együttműködés célja a jelnyelvi és szociolingvisztikai tudományos eredmények közvetítése a jelnyelvi tolmácsolás számára, valamint tolmácsolás nyelvészeti támogatása.

b. felsőoktatási intézményekkel való együttműködések

A *Nyelvtechnológiai Kutatócsoport* munkatársai számos magyar egyetemmel állnak kapcsolatban. Az egyik legfontosabb ezek közül a *Pécsi Tudományegyetem PTE Alkalmazott Adattudományi és Mesterséges Intelligencia Kutatócsoportjával* való együttműködés. A két intézmény által létrehozott HILANCO konzorcium intenzív kutatási irányba a nyelvi modellek fejlesztése és tesztelése.

Igen szoros kapcsolatot ápolnak a *Pázmány Péter Katolikus Egyetemmel* is: annak Bölcsész- és Társadalomtudományi és Informatikai Karával egyaránt. A kutatócsoport ott oktató tagjai 4 szakdolgozat elkészülésében segítettek témavezetőként. Ezek közül eddig egy szakdolgozat sikeres védelme történt meg. A kutatócsoport három kollégája dolgozott tudományos munkatársként az MTA-PPKE Magyar Nyelvtechnológiai Kutatócsoportjában 2021-ben, ahol szintén nyelvtechnológiai feladatok megvalósításában vettek részt (pl. gépi fordítás, szentimentelemzés, szövegösszefoglalás, nyelvi modellek, hibás szövegek automatikus javítása). Közülük ketten a PPKE ITK és az Idomsoft cég közötti K+F nemzeti laboratórium projektben is dolgoznak, ahol az állami rendszerek és ügyintézési folyamatokba integrálnak nyelvtechnológiai eszközöket.

Az együttműködés számos más egyetemmel is hangsúlyos: a *Debreceni Egyetem Általános Orvostudományi Karának Magatartástudományi Intézetével* folyó közös kutatás célja egy olyan korpusz-annotációs séma kidolgozása, amellyel a különböző, betegellátással kapcsolatos blogbejegyzések, személyes írások annotálhatók, azokból egy korpusz állítható elő, amely alapján lehetővé válik magyar páciensek narratív tapasztalatainak elemzése.

Az *Eötvös Loránd Tudományegyetem Társadalomtudományi Kar, Szociológia Tanszék és a Clementine Kft.* közötti együttműködés keretében egy munkatársunk vizsgálja a nemek közötti

nyelvhasználati különbségeket a Facebook bejegyzésekben. Fentiekén túl az *Eszterházy Károly Egyetem Informatikai Karával* is folynak közös kutatások.

A Többszínű és Edukációs Nyelvészeti Kutatócsoport egyik tagja a magyar nyelvű társalgások vizsgálatával közreműködik a *Károli Gáspár Református Egyetem* kutatójával együttműködve a „Beszédjegyzetek fonetikai jellemzőinek összefüggései, társalgáselemzés” (NKFI-128810) c. projektben, valamint az ELTE BTK Diagram Funkcionális Nyelvészeti Központ Kognitív Stilisztika Műhelyében részt vett a magyar nyelv stilisztikai jelenségeinek funkcionális keretű kutatásában.

4. A kutatóhely hazai vonatkozású (így különösen NKFI Alap, OP) pályázati adatai (l. a 3. Melléklet excel-táblázatában)

5. A legfontosabb tudományos publikációk

Monográfia:

1. 2021 Juliane House és Dániel Z. Kádár: *Cross-Cultural Pragmatics*. CAMBRIDGE UNIVERSITY PRESS, CAMBRIDGE ISBN: 9781108949545, 286+ xv pp.

Szerkesztett kötetek:

2. 2021 Juliane House és Dániel Z. Kádár: *Political Language in Contrast*, special issue of JOURNAL OF PRAGMATICS <https://www.sciencedirect.com/journal/journal-of-pragmatics/special-issue/10TFL6347T0> D1
3. 2021 Michael Haugh, Dániel Z. Kádár és Marina Terkourafi: *The Cambridge Handbook of Sociopragmatics*. CAMBRIDGE UNIVERSITY PRESS, CAMBRIDGE ISBN: 9781108844963, 783+xii pp.
4. 2021 Dániel Z. Kádár, Vahid Parvaresh és Rosina Marquez Reiter: *Alternative Approaches to Politeness*, special issue of JOURNAL OF POLITENESS RESEARCH 17(1), MOUTON DE GRUYTER, ISSN: 1613-4877 Q1
5. Kontra Miklós – Borbély Anna (szerk.) 2021. *Tanulmányok a budapesti beszédéről a Budapesti Szociolingvisztikai Interjú alapján*. Gondolat Kiadó: Budapest. 355 l.

Szakfolyóirat-cikkek:

6. 2021 Juliane House, Dániel Z. Kádár, Fengguang Liu és Zhuo Bi: 'Altered speech act indication: A problem for foreign language learners?', SYSTEM. <https://doi.org/10.1016/j.system.2021.102554> SSCI D1
7. 2021 Dániel Z. Kádár és Juliane House: 'Mimesis and ritual frames: A case study of military training in Chinese universities', LANGUAGE & COMMUNICATION 80: 21–30. <https://doi.org/10.1016/j.langcom.2021.05.006> SSCI D1
8. 2021 Dániel Z. Kádár és Juliane House, 'Interaction ritual and (im)politeness', JOURNAL OF PRAGMATICS 179: 54–60, SPECIAL ISSUE IN HONOUR OF MICHAEL HAUGH. <https://doi.org/10.1016/j.pragma.2021.04.021> SSCI D1
9. 2021 Juliane House, Dániel Z. Kádár, Fengguang Liu, Shiyu Liu, Wenrui Shi, Zongfeng Xia and Lin Jiao, 'Interaction, speech acts and ritual: An integrative model', LINGUA 257. <https://doi.org/10.1016/j.lingua.2021.103082> SSCI Q1
10. 2021 Dimitra Vladimirova, Juliane House és Dániel Z. Kádár: 'Complaining on social media: The case of #MuckyMerton', JOURNAL OF PRAGMATICS 177: 51-64. <https://doi.org/10.1016/j.pragma.2021.01.017> SSCI D1

11. 2021 Juliane House és Dániel Z. Kádár: 'War crime apologies: A contrastive pragmatic case study', JOURNAL OF PRAGMATICS 177: 109-121.
<https://doi.org/10.1016/j.pragma.2021.02.001> SSCI D1
12. 2021 Fengguang Liu, Dan Han, Dániel Z. Kádár és Juliane House: '(M)minzu-zhuyi' and 'Nationalism' in the Chinese and U.S. American Linguacultures: A Contrastive Pragmatic Analysis', JOURNAL OF PRAGMATICS 174: 168-178.
<https://doi.org/10.1016/j.pragma.2020.12.013> SSCI D1
13. 2021 Dániel Z. Kádár és Juliane House: 'Politeness markers' revisited: A contrastive pragmatic perspective', JOURNAL OF POLITENESS RESEARCH 17(1): 79-110.
<https://doi.org/10.1515/pr-2020-0028> SSCI Q1
14. 2021 Dániel Z. Kádár, Vahid Parvaresh és Rosina Marquez Reiter: 'Alternative approaches to politeness and impoliteness: An introduction', JOURNAL OF POLITENESS RESEARCH 17(1): 1-8.
<https://doi.org/10.1515/pr-2020-0028> SSCI Q1
15. 2021 Dániel Z. Kádár, Juliane House, Fengguang Liu and Yulong Song: 'Admonishing – A Paradoxical Pragmatic Behaviour in Ancient China'. PRAGMATICS 31(2): 173–197.
<https://doi.org/10.1075/prag.20022.kad> SSCI D1

KUTATÓKÖZPONT NEVE:
Nyelvtudományi Kutatóközpont
KUTATÓINTÉZET NEVE:
Történeti Nyelvészeti és Uralisztikai Intézet
Cím:
1068 Budapest, Benczúr u. 33.
Postafiók címe:
1394 Budapest, Pf. 360
telefonszám:
(1) 342-9372
honlap:
www.nytud.hu
e-mail:
linginst@nytud.hu

I. A Történeti Nyelvészeti és Uralisztikai Intézet kutatási eredményei a 2021. évre vonatkozóan

Tématerület: emberi erőforrás

A **Történeti és összehasonlító mondatlan kutatócsoport** jelentős eredményeket ért el az uráli nyelvek szintaxisának kutatásában. Számos aspektusból elemzés tárgyai voltak az uráli nyelvekben zajló SOV > SVO szórendi változás következményei, és ez segítette az ősmagyarban-ómagyarban feltételezett hasonló mondatlan változások rekonstruálását is. Vizsgálat tárgya volt a vonatkozó névmással bevezetett poszt-nominális vonatkozó mellékmondatok kialakulása a hantiban és az ómagyarban; hasonlóképpen az ugor korrelatív szerkezet véges tárgyi mellékmondatá alakulása a magyarban. Sikerült megoldani számos, a mari főnévi igeneves kifejezések dativusi alanya által felvetett szintaxiselméleti problémát. Kutatás tárgya volt továbbá a szünetikus mellérendelés és az ellipsis megjelenése és a két jelenség korrelációja a hantiban; ugyanígy a határozatlan névmások egzisztenciális, univerzális, szabad választást kifejező és negatív névmási kvantorokká való specifikálódása az uráli nyelvekben és az ómagyarban. Górcső alá került az információs szerkezet és a prozódia kölcsönhatása a szórendi változásokban. Sikeres kutatás tárgya volt az is, hogy hogyan oldható fel a mariban és az udmurtban (és számos más SOV szórendű nyelvben) a tagadó elem mondatlan helye és hatóköre között tapasztalható ellentmondás. A tagadással kapcsolatban a magyarban is megoldást nyert egy hatóköri probléma (a *Mit nem mondasz!* típusú felszólító mondatokban megfigyelhető ún. expletív tagadás kérdése). Egy mai magyar beszélt nyelvi szerkezet elemzésével sikerült alátámasztani az állítást, hogy a magyar VP még őrzi az uráli SOV szórendet. Kutatás tárgyai voltak még az uráli nyelvek határozói esetragos és névutós szerkezetei. Az abesszív PP-ben tapasztalható határozatlansági effektus lehetőséget adott a versengő határozatlansági effektus elméletek tesztelésére és a köztük való választásra. Sikerült magyarázatot találni az ugor és szamojéd nyelvek egyik meglepő sajátosságára: az 1. és 2. személyű (egyres nyelvekben valamennyi) tárgyi személyes névmáson megjelenő birtokos személyragra (pl. *eng-em, tég-ed*). Vizsgálat tárgya volt az is, hogyan alakult át az ugor kori tárgyjelölés és a tárgyi egyeztetés a három ugor nyelvben.

Jelentős kutatási eredmények születtek a tundrai nyenyec mondatlan és prozódia területén is. Az első a kiegészítendő kérdés mondatlan és prozódiai szerkezetét érinti. A kérdő kifejezés három mondatlan pozícióban állhat a kérdésben: lehet mondatkezdő helyzetű (az alany előtt), állhat a mondat közepén (az alanyt követő és a tárgyat megelőző helyen), megjelenhet közvetlen ige előtti helyzetben (a direkt tárgy után). A korábban megfigyelt *[op...wh] szintaktikai megszorítás (ti. kérdő kifejezést nem előzhet meg operátor) mindhárom pozícióban érvényesül. Az elsődleges prozódiai megfigyelések alapján a kérdő kifejezés mindhárom szintaktikai pozícióban hangsúlyos. A hangsúlyozási mintázatban aszimmetriát lehet megfigyelni a három pozícióban: azokban a kérdésekben, ahol a kérdőszó nem mondatkezdő helyzetű, a többi összetevő hangsúlytalanodik, míg a mondatkezdő kérdő kifejezés esetén minden összetevőre hangsúly esik. A jelenség magyarázatához további tesztek elvégzése szükséges. A kérdő kifejezések többségére jellemző, hogy

a szón megfigyelt dallamkontúr csúcса a második szótagon éri el alapfrekvencia-maximumát, ami arra utal, hogy ez a szótag hordozza a szóhangsúlyt. Ez azért meglepő, mert a szakirodalom alapján a tundrai nyenyecben a szóhangsúly mindig a szó első szótagjára esik. Első szótagi hangsúlyra utaló intonáció viszont csak a temporális kérdőszó esetében figyelhető meg.

A második a többszörös kérdés mondattani szerkezetével kapcsolatos. A tundrai nyenyecben megjelenhet egyszerre két vagy több kérdő kifejezés. A többszörös kérdésben a kérdő kifejezések relatív sorrendjében érvényesül az ún. specifikussági megszorítás, vagyis a nagyobb hatókörű kérdő kifejezés kötelezően specifikus. Ennek egyik bizonyítéka a tárgyi egyezés megjelenése az igén elől álló kérdőszói tárgy esetén, valamint bizonyos kérdő kifejezések sorrendjének agrammatikussága.

A harmadik a választó kérdés mondattani szerkezetét érinti. A választó kérdésre fókuszáló teszteléses vizsgálatok három szerkezetet azonosítottak a nyelvben, amelyek a választó kérdés funkcióját tölthetik be: (i) választó értelmű (kimerítő) kiegészítendő kérdés, amely a lehetséges alternatívákat egy zárt halmaz elemeiként felsorolva megadja; (ii) eldöntendő kérdés, amely az alternatívákat 'p ṽp?' formában tartalmazza (tehát egy tagadó elemet is tartalmaz a kérdés); (iii) mellérendelő eldöntendő kérdések formájában, amelyek nem tartalmaznak kötőszót. Az eddigi megfigyelések egyelőre leíró nyelvészeti megállapításokra alkalmasak, további vizsgálatokra van szükség a témában.

A projekt keretében ezeken felül egy tundrai nyenyec (egynyelvű) korpusz is építés/fejlesztés alatt áll. A korpusz jelenleg több mint 460 000 szövegszót tartalmaz. 2021-ben az alábbi fejlesztések lettek végrehajtva. Online elérhetővé vált a tundrai nyenyec korpusz. A korpuszkezelő felhasználói felülete módosítva lett, olyan funkciók lettek beállítva, amelyek a szövegek jelenlegi feldolgozottsági szintjén használhatóak. A korpuszban az egyszerű keresés mellett jelenleg frázisokra, szavakra, karakterekre/karakterláncokra lehet keresni. Emellett reguláris kifejezések segítségével lehet komplex keresési mintákat alkotni. A keresés egyszerűbbé tételének érdekében létre lett hozva egy cirill billentyűzet a felületen. A feldolgozott és a korpuszba feltöltött tundrai nyenyec szövegek részleges annotálása elkezdődött. A szöveg szintjén jelölve vannak a kérdések, a rövidítések és a lábjegyzetek. Ezek mindegyikéhez szükség volt manuális feldolgozásra is. A szavak szintjén a személyes névmások lettek ellátva grammatikai információval (szófaji és morfológiai címkékkel).

Az **Uralisztikai Kutatócsoport** kutatásai szintén több területen jártak jelentős eredményekkel. Az első ezek közül az északi–sarkvidéki városokban élő őshonos kisebbségek identitáskonstrukciói. A számi, a manysi, valamint a tajmir-félszigeti őshonos kisebbségek (elsősorban nganaszan, továbbá pl. dolgán) identitáskonstrukcióinak összehasonlító vizsgálata lezárult. Fel lett tárva, hogy a nyelv központi helyet foglal el a vizsgált kisebbségek identitásában, annak ellenére, hogy a nyelvhasználati szokások ezzel nincsenek összhangban. A vizsgált kisebbségekre ez egyaránt jellemző, de attitűdjeikben eltérő jegyek is vannak. Az őslakos identitást övező pozitív attitűd általános a vizsgált népek körében. A várakozásoknak megfelelően, ha a kisebbségi nyelvet nem beszélik is, a kulturális kötődések egyértelműen jelen vannak.

A második ide tartozó kutatási terület a térbeli viszonyok kifejezőeszközeit foglalja magában. Elkészült két [nganaszan](#), egy [hanti](#) és egy manysi testtartásige grammatikalizációjának leírása. Az 'ül' jelentésű igék kvázikopolává válásának kulturális meghatározottságát kognitív nyelvészeti irányból igazolták. Kérdőíves felmérést végeztek a magyar nyelv lokatív mondataiban az igei elemek felgyűjtéséhez, ehhez adaptálták a nemzetközi kutatásokban használatos képsort. Feltárták, hogy a magyar lokatív mondatokban más igei elemek használata mellett a rezultatív szerkezetek szerepe is jelentős. Megkezdődött a fiktív mozgást tartalmazó lokatív mondatok leírása: az előforduló igék csoportosítása, illetve az alany–állítmány megkötések vizsgálata. Elkészült továbbá a magyar lokatív mondatok lehetséges igei elemeinek áttekintése, csoportokba rendezése, valamint szótárszerű feldolgozása.



1. kép: Terepmunka nganaszan adatközlővel

A harmadik terület a nyelvi őstörténet. Egy háromszerzős könyvfejezetben áttekintették a magyarországi uralisztika állásfoglalásait a többnyelvűség-kutatás relevanciájáról a nyelvrokonság-kutatásban: összefoglalták a többnyelvűség-modellezés lehetőségeit korábbi nyelvtörténeti időszakokra nézve. Felsorakoztatták a társtudományok érvelését a nyelvi kapcsolatok meglétéről a preuráli és alapnyelvi korokban. Esettanulmányokat ismertettek a törzsi nyelvek érintkezésének változatos szociolingvisztikai következményeiről. A nemzetközi uralisztika eredményeiből bemutatták azokat, amelyek a többnyelvűség meglétét és sikeres kutathatóságát bizonyítják a történelem előtti, illetve nyelvemlékek nélküli korokban.

A negyedik terület a kontaktusnyelvészet. A kutatás tárgyát a szláv hatásra finnugor nyelvekben fellépő – hagyományosan kódváltásnak nevezett – jelenségek képezik. Megvizsgálták, hogy a magyar–horvát anyag alapján felállított kategorizáció érvényesíthető-e az udmurt–orosz korpusz adataira is. Megállapították, hogy mindkét nyelvpárhoz tartozó adatok beilleszthetők a rendszerezésbe, s az adateloszlásban a megfeleltethető kategóriák között nagymértékű hasonlóság van. Az adatbázis jelenlegi feltöltöttségében a szerkezetkölcsonzésre jóval kevesebb udmurt példa található, mint magyar. A következő szakaszban rögzítendő további udmurt–orosz adatok ezt a képet módosíthatják. Az írásbeliség nélküli nyelvekben nehéz hozzáférni a masszív orosz hatás előtti nyelvállapot jellemzőihez. Az egyik jelenség, amelyről nehéz eldönteni, hogy orosz hatás következtében jött-e létre, az az állandó/ideiglenes tulajdonságokat leíró mondatoknak a mai hantiban megfigyelhető morfológiai megkülönböztetése. Az eredmények szerint áttételes orosz hatással lehet számolni: a megkülönböztetés csírái, valamint saját eszközei az orosz hatás előtt is megvoltak a hantiban.

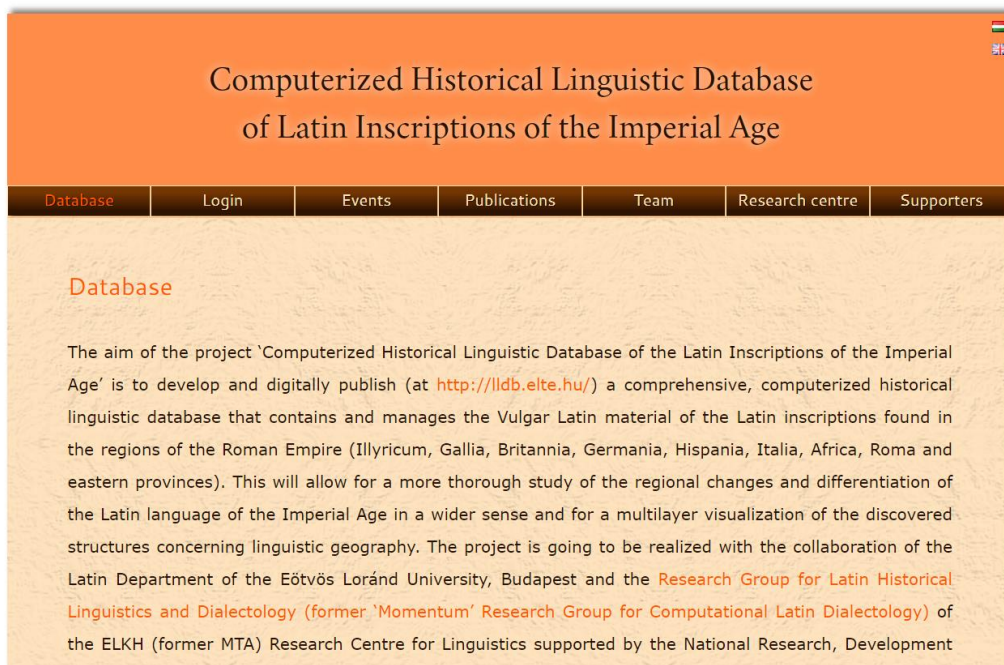
A **Magyar Nyelvtörténeti Kutatócsoport** tagjai a *Regiszterfüggő változatok a középmagyarban* projektum (NKFI FK 135186) keretében kétféle feladaton dolgoztak. A 2021-es év volt az első teljes éve ennek a munkálatnak, s ezen első év fő célkitűzése a projekt módszertani megalapozása volt. Ennek keretében egyrészt megtervezték és építeni kezdték a középmagyar emlékirat- és drámakorpuszt. Tíz emlékiratból jelölték ki egyenként 70 000 karakternyi terjedelmű szakaszokat, ezeket szükség esetén digitalizálták és OCR-ezték. Az év során elkészült hét szövegegység normalizált változata 517 000 karakter terjedelemben, valamint a lekérdezőfelület belső tesztverziója. Szintén a módszertani megalapozást szolgálta a statisztikai ismeretek közös bővítése, az év utolsó szakaszában ez is megkezdődött.

A projekt keretében folytatott vizsgálatok egyrészt különböző morfoszintaktikai kategóriákra irányultak: a) A múlt idők témakörében részletes vizsgálat készült a visszaszorulóban lévő *-t+vala/volt* összetett múlt időről a kolozsvári és a magyarországi boszorkányperekben, illetve a múlt idők közötti választás

és váltakozás jelenségeiről ugyancsak a kolozsvári boszorkányperekben. b) 16–18. századi levelezések és boszorkányperek szolgáltak alapjául a kötőmód tanulmányozásának (formai és szemantikai jellemzők feltárása). c) A vonatkozó mellékmondatok grammatikai jegyei a boszorkányperes és az úriszéki peres iratok tükrében: kirajzol-e a vizsgált grammatikai kategória regiszterkülönbségre visszavezethető eltéréseket a két szövegtípus között. d) A tagadás témakörében készült egyrészt egy átfogó jellegű leíró elemzés a pleonasztikus tagadásról (PT), másfelől pedig [egy összehasonlítás egy formális regisztert képviselő szövegkorpuszal](#) (Pázmány Péter műveivel). Beigazolódott, hogy a PT tág teret enged az egyéni, illetve a regiszteralapú variációnak.

Másrészt a projekt keretében pragmatikai változók vizsgálatára is sor került: a) A diskurzusjelölők kutatásán belül a *mondván* esetében jelentős következményekkel járó változás, hogy a középmagyarban kiterjed a használata, és a függő idézésekben is szokásossá válik; a perek tanúvallomásait ugyanakkor elsősorban az egyenes idézetek melletti használat jellemzi. b) A pragmatikai markerek 18. századi történetével három esettanulmány foglalkozott. Ezeknek fő eredményei: i) Egy szerző írásain belül [regiszter szerint oszlik meg a markerek használata](#) a fő funkciókat tekintve. ii) A közel szinonim jelentésű pragmatikai markerek a korszakban eltéréseket mutatnak a módosított elemek (alaptagok) kategóriái szerint, a negatív polaritású nyomatékosítók pedig a semlegessé válás irányába mozdulhattak el.

A **Latin Nyelvtörténeti és Dialektológiai Kutatócsoport** tevékenysége elsősorban a [Császárkori latin feliratok számítógépes nyelvtörténeti adatbázisa](#) alapján folyt, amely több mint 14 000 új adatlappal és informatikai fejlesztéssel gyarapodott az év folyamán. A kutatások részint az itáliai latin nyelv hangrendszerének dialektológiai jellegzetességeinek leírására irányultak a késő császárkori és kora középkori feliratok alapján. Sikerült kimutatni, hogy a kiválasztott jelenségek tekintetében az adatok latin területi megoszlása előre jelzi, megelőlegezi az olasz dialektusokon belül, ill. az olasz és szárd dialektusok között megfigyelhető jellegzetes területi különbségeket. Szintén sikerült egy Aquincumban, leletmentő ásatásról előkerült mágikus lemez szövegmegállapítását és nyelvi feldolgozását elvégezni. A verses feliratokról rögzített metrikai alapú nyelvi adatok ellenőrzésével, gondozásával és kiegészítésével is foglalkozott a kutatócsoport. Mindezek mellett folyamatosan zajlott a fent említett Adatbázisban az adatlapok felvitele, ellenőrzése és gondozása.



2. kép: Császárkori latin feliratok számítógépes nyelvtörténeti adatbázisa - nyitólap

Születtek kutatási eredmények a latin igeragozási rendszer sajátos fonológiai szerkezete és a rendszeren belül fennálló implikációs viszonyok területén is, amelyek a latin nyelvészetet a korszerű fonológia- és morfológiaelméleti kutatásokhoz tudják kapcsolni.

A **Magyar Dialektológiai Kutatócsoport** további anyagokkal bővítette a 2020-ban elindított, [moldvai hanganyagok hanggal összekapcsolt fonetikus lejegyzéseiből](#) álló adattárat. A vállalkozás egyedülálló a magyar dialektológia történetében olyan tekintetben, hogy korábban nem volt példa ilyen hosszú felvételek fonetikus lejegyzéseiből készült adattár építésére. 2021 szeptemberében elindult egy új projekt is, amelynek legfontosabb célja a magyar magánhangzók akusztikai minőségeinek feltérképezése, a magánhangzórendszerek összevetése és magyar nyelvjárások dialektometriai elemzése.

II. Tudományszervezési eredmények, események

1. Rövid általános ismertető a kutatóhely tudományos munkáját illetően.

A Történeti Nyelvészeti és Uralisztikai Intézetben folyó tudományos munka kiterjed a magyar nyelv szoros értelemben vett történetére, azaz a magyar nyelvű írásbeliség megjelenése óta lezajlott változások vizsgálatára (ómagyar, középmagyar korszak), másfelől a magyar nyelv őstörténetére, azaz az uráli/finnugor nyelvészetre, a magyar nyelv rokonsági körébe tartozó nyelvek összehasonlító vizsgálatára; a rokon nyelvek tekintetében a nyelvi veszélyeztetettséggel és a nyelvi kontaktusok következményeinek feltárásával is foglalkoznak az Intézet kutatói. Az itt végzett munka kiterjed a történeti nyelvészet módszertani és elméleti vonatkozásaira is, valamint olyan nyelvek vizsgálatára, amelyekkel a magyar nyelv a története során érintkezésbe lépett, és amelyek komoly befolyást gyakoroltak rá (pl. szláv, latin). Etnolingvisztikai, illetve a kulturális nyelvészet módszertanához tartozó kérdéseket is vizsgál, mivel számos nyelvi jelenség összefüggésbe hozható társadalmi és kulturális tényezőkkel. A magyar nyelvjárások vizsgálata szintén az intézet munkájához tartozik; a dialektometria eszköztárával folyik a nyelvjárások közti hasonlóság, a nyelvjárások csoportosításának lehetséges irányai és a nyelvi kontinuum kutatása.

Az intézetnél folyó munka változó intenzitással bár, de kiterjed a központi nyelvi szintekre: szintaxis, morfológia, fonológia, szókészlet. Ezek kutatásának kereteit a legkorszerűbb elméleti megközelítések adják. A rokon nyelvek kutatása szintén a szociolingvisztika legkorszerűbb megközelítésein alapszik; ezeknek a kutatásoknak a fontosságát inherens tudományos értékükön felül az adja, hogy ezeknek a nyelveknek jó része veszélyeztetett nyelv, élő formában való kutatásuk előreláthatólag mindössze néhány generáción át lesz lehetséges.

A magyar nyelvtörténeti vizsgálatok középpontjában a fentiek mellett a nyelvhasználati változatok és maga a nyelvi változás áll. Amikor egy új nyelvi jelenség felbukkan, egyúttal a nyelvi variációnak is új színtere nyílik, az új jelenség pedig – a nyelv rendszerébe bekerülve – egyben változattá is válik, amely a potenciális egyéb változatokkal (beleértve a zéró változatot is, azaz a jelenség használatának mellőzését) vagy stabil, vagy instabil viszonyban áll. Az egyes szinkrón nyelvallapotok leírásához nélkülözhetetlen a változatok viszonyának feltárása, a diakrón változásokat pedig a változatok története rajzolja ki. Azt, hogy a beszélők hogyan választanak a változatok között, számtalan nyelven belüli és nyelven kívüli tényező befolyásolhatja, és ezeknek a tényezőknek a szerepe is változhat az idők során. Az utóbbi években a kutatócsoport fókuszában elsősorban a középmagyar kor áll, de egyes vizsgálatok más korszakokat is érintenek, illetve akár átfogó, a teljes adatolt nyelvtörténeti időszakot lefedő képet alkotnak egy-egy jelenségről.

Egyes adatbázisok építése és folyamatos gondozása is az intézet feladatai közé tartozik. A magyar nyelv uráli eredetű szókincsének vizsgálatát segíti elő az *Uralonet* (uráli/finnugor etimológiai adatbázis) és az *Új magyar etimológiai szótár* összekapcsolása. Megtörtént a két állomány releváns lexikai egységeinek kijelölése, elkezdődött a keresőstruktúra megtervezése. Tovább folyik a történeti magánéleti korpusz gondozása, jelenleg épül a Középmagyar emlékirat- és drámakorpusz, valamint a moldvai nyelvjárási szövegtár és hangzó adatbázis.

Az intézethez kapcsolódnak azok a kutatások is, amelyek középpontjában a magyar nyelv térbeli változatossága áll, elsősorban a Kárpát-medencei magyar nyelvváltozatoké. Az intézetben a dialektológiai kutatások közel százéves múltra tekinthetnek vissza, a kutatócsoportban folyó projektek szervesen építenek az elődök munkájára, az elemzésekben kiaknázva a legújabb nyelvészeti technológiák adta módszertani lehetőségeket. Ezek a kutatások többnyire a kvantitatív nyelvfeldrajz körébe sorolhatók. A kutatók a dialektometria eszköztárával vizsgálják a nyelvjárások közti hasonlóságot, a nyelvjárások csoportosításának lehetséges irányait és a nyelvi kontinuumot. Hangfelvételek akusztikai elemzése révén kutatják a magyar nyelvjárások magánhangzórendszereit, céljuk a rendszereken belüli és a rendszerek közötti összefüggések feltárása.

2. A tudomány és a társadalom kapcsolata, társadalmi hasznosság.

Azon túl, hogy a magyar nyelv történetének minél alaposabb megismerése önmagában is fontos kutatási célnak tekinthető, társadalmi szempontból két feladatot tekintünk különösen lényegesnek. Az egyik ezek közül a különböző nyelvtörténeti szövegek hozzáférhetővé tétele minél szélesebb körben. Ezt szolgálja a Magyar Nyelvtörténeti Kutatócsoport által már létrehozott, a szakirodalmi visszhang alapján is kifejezetten felhasználóbarátnak tekinthető *Történeti magánéleti korpusz*, ezt célozza a jelenleg épülő *Középmagyar emlékirat- és drámakorpusz*, és ugyanez a céljuk a kutatócsoport tagjai által korábban megjelentetett, filológiai is megbízható szövegkiadásoknak. Ugyanígy nyilvánosan elérhető az Intézet munkatársai által épített moldvai nyelvjárási hangfelvételek lejegyzéséből álló szövegtár, már a munkafolyamat kezdetétől. A szövegeket a munka során a kutatók folyamatosan javítják, frissítik (<https://nlp.nyud.hu/csango/index.html>)

A másik, hasonlóan fontos feladat annak tudatosítása, hogy a nyelv elidegeníthetetlen tulajdonsága a változás, s ennek érdekében az intézet kutatói rendszeresen tartanak ismeretterjesztő előadásokat (részint az MTA Alumni Programjában), illetve írnak ismeretterjesztő cikkeket különböző jelenségek változásairól, avagy tágabb értelemben vett történeti nyelvészeti és dialektológiai kérdésekről.

Az intézet kutatói egy részének van formális együttműködése hazai felsőoktatási intézménnyel (ld. 3. pont lentebb), de informális módon, valamint eseti jelleggel a kutatók többsége részt vesz különböző, a felsőoktatáshoz kapcsolódó feladatokban (pl. opponensi vélemények írása, diákműhelyben való mentori részvétel, absztraktbírálat hallgatói, illetve doktoranduszkonferenciára).

A zirci Reguly Antal Múzeum és Népművészeti Alkotóház részére a *Regulyversum – mindent Regulyról* című kiállításához az uráli nyelvcsaládról, a történeti összehasonlító nyelvészet módszertanáról digitális felhasználásra alkalmas tartalmak létrehozása is az intézet kutatóinak bevonásával történt.

A Nyelvtudományi Kutatóközpont (benne intézetünk munkatársai) kezdetektől fogva központi szerepet játszik a Magyar Nyelvészeti Diákolimpia magyarországi válogatóversenyeinek lebonyolításában és a csapatok nemzetközi versenyre való felkészítésében. A feladat részét pályázati tevékenység is képezi. A beszámolási évben sikeresen pályáztak a 2022-es nemzetközi versenyen való részvételre. A vírushelyzet miatt a nemzetközi versenyt online rendezték. A versenyhéten a csapatok Budapesten tartózkodtak, a verseny helyszíne a Nyelvtudományi Kutatóközpontban volt. Egy ezüst- és egy bronzérmét, valamint két dicséretet szereztek a diákok.

Az intézet munkatársai számos honi és nemzetközi konferencián tartottak előadásokat; ezek javarészt még online formában zajlottak a COVID-járvány miatt. De az intézet maga is több konferenciát szervezett 2021 folyamán: a Szegedi Tudományegyetem Finnugor Tanszékével közösen meghívásos alapú, nyílt lebonyolítású konferencia: *A nyelvészeti antropológiától az etnoszintaxisig „Régi-új” kutatási területek az uralisztikában és azon túl* (Nyelvtudományi Kutatóközpont, 2021. november 5.); egy nemzetközi akadémiai emlékülés: *Sajnovics János Demonstráció 250* címmel. (Nyelvtudományi Kutatóközpont, 2021. október 21., az MTA Uráli Nyelvészeti Munkabizottsági tagság keretében); az ELTE Finnugor Tanszékével közösen: *Jó gyakorlatok a finnugor nyelvek oktatásában módszertani konferencia* (nemzetközi, 2021. december 1.).

3. A kutatóhely legfontosabb hazai és nemzetközi eredményes K+F+I együttműködései az elmúlt évben.

Az intézet munkatársai közül többen is részt vesznek a felsőoktatásban, tehetséggondozásban, utánpótlás-nevelésben, beleértve a doktori iskolákban oktatóként és törzstagként való részvétel, egyetemi kurzusok tartását, szakdolgozati és OTDK-témavezetést, doktori témavezetést, szuperkurzusok, vendégelőadások és kutatószemináriumok tartását nemcsak Magyarországon, de neves külföldi egyetemeken is. Ezek az együttműködések jelentősen hozzájárulnak mind az intézet képviseléséhez a kutatóközponton kívül, mind pedig az utánpótlás-neveléshez.

Az intézet kutatói ezeken felül gyakran végeztek szaklektori munkát folyóiratoknak, kiadóknak, pályázati szerveknek. Megjegyzendő még, hogy az intézet kutatói közül kerül ki mind a *Nyelvtudományi Közlemények*, mind pedig az *Acta Linguistica Academica* c. folyóirat főszerkesztője.

4. A kutatóhely hazai vonatkozású (így különösen NKFI Alap, OP) pályázati adatai. (részletesebben lásd 3. sz. melléklet)

A Magyar Dialektológiai Kutatócsoport új pályázatot nyert el a most indult stratégiai tervezési időszakra, „Magyar nyelvjárások akusztikai és dialektometriai vizsgálata” címmel (NKFIH FK-138396, vezető kutató: Vargha Fruzsina Sára, 2021–2025).

A Latin Nyelvtörténeti és Dialektológiai Kutatócsoportban, a Császárkori latin feliratok számítógépes nyelvtörténeti adatbázisa c. projekthez kapcsolódóan két OTKA K pályázat is futott Adamik Béla témavezetésével 2021-ben, a K 124170 sz. projekt utolsó éve és a K 135359 projekt első éve egyaránt november végén zárult. Továbbá két NKFIH MEC_R (141374 és 141509 sz., Adamik Béla, Cser András) és egy MTA INKP 2022 (INKP 2022_22 sz., Barta Andrea) nemzetközi konferencia-részvételi pályázat is támogatást nyert az év végén.

A Magyar Nyelvtörténeti Kutatócsoport egy NKFIH alapkutatási pályázati projektum (FK 135186, „Regiszterfüggő változatok a középmagyarban”), keretében dolgoznak, ennek a pályázatnak négy év a futamideje (2020. november 1. – 2024. október 31.), a teljes költségvetése pedig 39 990 000 Ft.

Az Uralisztikai Kutatócsoport 2021-ben az alábbi pályázati projektumokon dolgozott:

- Способы выражения пространственных отношений на материале русского и угорских языков. [Téri relációk kifejezőmódjainak kérdései orosz és ugor nyelvi anyagon] RFFI (2021. márc. – 2022. dec.)
- Прагматические функции кодовых переключений (на материале славянских и финно-угорских языков) [A kódváltás pragmatikai funkciói (szláv és finnugor nyelvekben)] (2021. március – 2022. december)
- A Schmidt-hagyaték feldolgozásának előkészítése (ELKH Kiemelt Téma). (2021. nov. – 2022. nov.)

5. A 2021-es év legfontosabb publikációi:

Bácsi Enikő 2021. Mihez kötődik a kötőmód? Kötőmódhasználat a Nadasdy-levelezésben. *Nyelvtudományi Közlemények* 117: 281–300.

<http://real.mtak.hu/138337/1/B%C3%A1csi%20Enik%C5%91.pdf>

Dömötör Adrienne (szerk.), Gugán Katalin – Varga Mónika (társszerk.) 2021. Versengő szerkezetek a középmagyar kor nyelvében. *Nyelvtudományi Értekezések* 169. Akadémiai Kiadó, Budapest.

https://mersz.hu/keres/hal%C3%A1sz/hivatkozas/m883vsakke_7_p8/

[A kötetből öt tanulmány és a bevezető tanulmány az Intézet kutatóinak munkája.]

Gugán Katalin 2021. Tagadás és pleonasztikus tagadás Pázmány Péter műveiben. In: Balogné Bérces Katalin – Hegedűs Attila – Surányi Balázs (szerk.), *Nyelvelmélet és dialektológia* 5. PPKE BTK Elméleti Nyelvészeti Tanszék – Magyar Nyelvészeti Tanszék. Budapest. 13–37.

<http://btk.ppke.hu/uploads/articles/2616708/file/3-Gug%C3%A1n-%C3%A9s-2021-final.pdf>

Varga Mónika 2021. Történeti pragmatikai variációról a regiszter tükrében. A feltételt kifejező határozói igenevek esetei. *Jelentés és Nyelvhasználat* 8/1: 81–104.

<http://www.jeny.szte.hu/JENY-2021-VargaM>

Adamik, Béla 2021. Romanization and Latinization of the Roman Empire in the light of data in the Computerized Historical Linguistic Database of Latin Inscriptions of the Imperial Age. *Journal of Latin Linguistics* 20(1): 1–19. <https://www.degruyter.com/document/doi/10.1515/joll-2021-2016/pdf>

Salánki Zsuzsa – Oszkó Beatrix – Sipos Mária 2021. A kétnyelvűség helye és szerepe az alapnyelvi korban. In: Klima László – Türk Attila (szerk.): *Párhuzamos történetek. Interdiszciplináris őstörténeti konferencia a PPKE Régészettudományi Intézetének szervezésében*. 2020. nov. 11–13. PPKE BTK Régészettudományi Intézet – Martin Opitz Kiadó – BTK Magyar Őstörténeti Kutatócsoport, Budapest. 11–38.

Sipos Mária 2021. A 'fekszik' ige grammatikalizációja a színjai hantiban. *Folia Uralica Debreceniensia* 28: 319–332. <https://finnugor.arts.unideb.hu/fud/fud28/fud28.pdf>

Várnai Zsuzsa 2021. Az 'áll' ige grammatikalizációja a nganaszban. *Nyelvtudományi Közlemények* 117: 101–122. <http://www.nytud.hu/nyk/nyk117.pdf>

Dékány Éva. 2021. *The Hungarian nominal functional sequence*. Cham: Springer.

Dékány Éva, Veronika Hegedűs. Postpositions: formal and semantic classification. 2021. In: Katalin É. Kiss and Veronika Hegedűs (eds.), *Syntax of Hungarian: Postpositions and postpositional phrases*. Amsterdam: Amsterdam University Press. 11–191.

<https://library.oapen.org/handle/20.500.12657/47855>

É. Kiss Katalin 2021. Definiteness effect in the PP. *Linguistic Inquiry* 2021 online early [Q1] https://doi.org/10.1162/ling_a_00453

Halm Tamás 2021. Radically Truncated Clauses in Hungarian & Beyond: Evidence for the Fine Structure of the Minimal VP. *Syntax*. [D1] <https://doi.org/10.1111/synt.12214>

Borise, Lena 2021. Focus projection need not be based on pitch accents: evidence from Georgian. *Glossa: a journal of general linguistics* 6(1): 128, 1-46. <https://www.glossa-journal.org/article/id/5733/>

Burukina, Irina 2021. Profile of reflexives in Hill Mari. *Folia Linguistica: Acta Societatis Linguisticæ Europææ* 55(1): 127--162. <https://www.degruyter.com/document/doi/10.1515/flin-2020-2072/html>

Csepregi Márta 2021. *K. F. Karjalainen keleti hanti szöveggyűjtése 1899-1901*. MTA doktori disszertáció. (A 120 éves kézirat anyag megfejtése, átírása, lefordítása, nyelvtani és lexikai elemzése és kommentálása)