

Főnévi igeneves szerkezetek a magyarban

Egy kutatási téma helye a (számítógépes) nyelvészetben

A nyelvtudomány nemzetközi fejlődését a 20. század közepétől két tendencia jellemzi: a modern nyelvészet¹ központi területein belüli differenciálódás, valamint a nyelvészettel rokon diszciplínák felé való nyitás, és – ennek eredményeképpen – új nyelvészeti tudományterületek megjelenése és művelése. A nyelvtudomány társdiszciplínáinak egy része társadalomtudomány, de találhatóak közöttük kognitív, matematikai és műszaki tudományok is.

Jelen dolgozat két, önmagában is megálló részből áll. Az első rész az MTA 56 éves Nyelvtudományi Intézetében folyó interdiszciplináris kutatásokról ad rövid áttekintést,² a második részben saját interdiszciplináris kutatásaimról számolok be. A dolgozat tehát egyrészt „elhelyezi” kutatómunkámat a nyelvtudomány tágabb kontextusában, másrészt a főnévi igeneves szerkezetekkel kapcsolatos kutatási célok, az alkalmazott eszköztár és az eddigi eredmények ismertetésével arra is rá fog mutatni, hogyan juthatunk egy számítástechnikával támogatott újszerű kutatómódszertannal az elméleti nyelvészek számára is értékelhető eredményekhez.

Interdiszciplináris kutatások a Nyelvtudományi Intézetben

A bevezetésben említett két nemzetközi tendencia – a szakosodás és a tematikai gazdagodás – az elsősorban a magyar nyelvvel kapcsolatos kutatásokat folytató Nyelvtudományi Intézet munkáját is

BOTTYÁN GERGELY
2001-ben végzett
– Eötvös-kollégistaként – az ELTE BTK angol nyelv és irodalom szakán, majd 2006 júliusában közgazdasági diplomát is szerzett a győri Széchenyi István Egyetemen. Munkahelyei: 2004-ig az ELTE BTK Nyelvtudományi Doktori Iskolájának ösztöndíjas PhD-hallgatója, azóta az MTA Nyelvtudományi Intézetének fiatal kutatója. 2005 februárja és 2006 júliusa között az Eötvös József Collegium Angol–Amerikai Műhelyének vezetője. Kutatási területei: a nyelvi jelentés elméletei, illetve a korpuszok nyelvészeti felhasználásának lehetőségei.

jellemzi. Az intézet tíz osztálya közül jelenleg öt osztályon folynak állandó jelleggel interdiszciplináris jellegű kutatások: (ábécésorrendben) az Elméleti Nyelvészeti Osztályon, az Élőnyelvi Osztályon, a Fonetikai Osztályon, a Kísérletes Nyelvészeti (2006-ig: Neurolingvisztikai és Pszicholingvisztikai) Osztályon és a Nyelvtechnológiai (2006-ig: Korpusznyelvészeti) Osztályon. A II. F fiatal Kutatók Fóruma megrendezésének idején mind az öt osztály rendelkezett fiatal kutatóval.

Az Elméleti Nyelvészeti Osztály elsősorban a mai magyar nyelv rendszerének feltérképezésével, magyar hangtannal, alaktannal, mondattannal és jelentéstannal, valamint nyelvemléttel foglalkozik. A magyar nyelvtan egyes részterületeivel kapcsolatos kutatások az általános nyelvészet eredményeit alkalmazzák, és az így létrejövő új eredmények maguk is beépülnek a nemzetközi nyelvtudományba.³ A leíró apparátus sokat merít a matematikából és – a jelentéstani vizsgálatok esetében – a logikából is. Az osztályon folyó interdiszciplináris kutatások a modern magyar nyelvtudomány eredményeit alkalmazzák a korszerű anyanyelvi nevelésben és a számítógépes nyelvészetben.⁴

Az Élőnyelvi Osztályon elsősorban szociolingvisztikai (társasnyelvészeti) kutatások folynak, amelyek a magyarországi magyar nyelv variabilitásával és változásaival, a környező országokban és a tengeren túl beszélt magyar nyelvváltozatokkal, a nyelvi jogokkal, a nyelvcserével, a normatív nyelvhasználattal kapcsolatos ideológiákkal, a jelnyelvekkel és a hazánkban élő kisebbségek nyelvhasználatával foglalkoznak.

A Fonetikai Osztály kutatásainak tárgya a (magyar nyelvű) beszéd. Ez egy rendkívül összetett jelenségkör, amelynek tanulmányozásához szükség van a fonetikával érintkező tudományterületek művelőivel való aktív együttműködésre. Néhány fontos kutatási irány: a (magyar) beszédhangok és egymásra hatásuk elemzése, a (magyar) beszédhang és a mondatfajta összefüggéseinek meghatározása, magyar írott szöveget beszédalakító (text-to-speech) és a beszélő személyt beszédének elemzése alapján azonosító számítógépes rendszerek fejlesztése, a beszéd megértésének modellezése. Kapcsolódó tudományterületek: akusztika, informatika, kognitív pszichológia.

A Kísérletes Nyelvészeti (2006-ig: Neurolingvisztikai és Pszicholingvisztikai) Osztályon kutatásai igénylik talán a legtöbb diszciplína együttműködését. A neurolingvisztika a nyelvészet, a neurológia és a kognitív pszichológia eredményeit kísérli meg szintetizálni. Művelőit elsősorban az foglalkoztatja, hogy a nyelvi kompetencia különféle összetevőit milyen agyi mechanizmusok teszik lehetővé. Erre a kérdésre a pszicholingvisztika is keresi a választ, emellett modellezi például az anyanyelv- és az idegen nyelv-elsajátítás, valamint a beszédértés folyamatait – a nyelvészetten kívül elsősorban a kognitív pszichológia eredményeire támaszkodva. A Kísérletes Nyelvészeti Osztály a fenti tudományközi területeken kívül cigánynyelvi kutatásokat is folytat.⁵ Ez utóbbi irány elsősorban a kisebbségkutatáshoz és a kulturális antropológiához kapcsolódik.

A Nyelvtechnológiai (2006 elejéig: Korpusznyelvészeti) Osztályon számítógépes nyelvészeti kutató-fejlesztő tevékenység folyik. Ennek két fő területe: 1. a nyelvészeti

munka támogatása a számítástechnika eszközeivel („számítógép a nyelvtudományban”), és 2. a nyelvészet eredményeinek alkalmazása a számítástechnikában („nyelvtudomány a számítógépben”). Szűk értelemben az 1. területet nevezik korpusznyelvészetnek, a 2. területet pedig nyelvtechnológiának.⁶ Az 1. terület fő kutatási irányai: (a) számítógépes szöveggyűjtemények, korpuszok összeállítása, (b) nyelvészeti hipotézisek tesztelése korpuszok segítségével, (c) a helyesnek bizonyult nyelvi elemzések bejelölése a korpuszokban, (d) nyelvészeti hipotézisek generálása korpuszok segítségével. A 2. területre néhány példa: helyesírás- és nyelvhelyesség-ellenőrző programok, nyelvi elemzéssel támogatott keresőmotorok, lexikai adatbázisok, automatikus fordítórendszerek létrehozása. Az osztály mindkét fő területen aktív. Tevékenységében az 1. terület minden említett kutatási iránya megjelenik. A 2. területen belül elsősorban gépi és emberi használatra készülő lexikai adatbázisok, valamint egy magyar–angol automatikus fordítórendszer létrehozásában vettek, illetve vesznek részt az osztály munkatársai.⁷

A magyar főnévi igeneves szerkezetekkel kapcsolatos kutatás

A vizsgálat tárgya és a korábbi kutatások

A szláv nyelvekben, valamint a legtöbb germán és újlatin nyelvben csak személyrag nélküli főnévi igenevek fordulnak elő. A magyarban azonban – néhány egymással különben nem rokon nyelvhez, például a walesihez és a portugálhoz hasonlóan – található személyrag nélküli és személyraggal ellátott főnévi igeneveket is.⁸

A magyar főnévi igenév két típusát mutatja be az 1. táblázat.

1. táblázat

Személyragos és személyrag nélküli főnévi igenevek a magyarban

| | |
|-----|----------------------------|
| (a) | Sikerült időben beérnie. |
| (b) | A levelet el kell küldeni. |

A főnévi igenév személyragos változatára példa az 1. táblázat (a) mondatának „beérnie” szóalakja: „beér” – ige (szótó), „-ni” – főnéviigenév-képző, „-e” – egyes szám harmadik személyű rag. A (b) mondatban szereplő főnévi igenév, a „küldeni” rag nélküli: ebben az alakban a „küld” igei szótó csak az „-(e)ni” főnéviigenév-képzővel kapcsolódik össze.

A magyar főnévi igeneves szerkezetek, konstrukciók az utóbbi időben a mondat-szerkezet ábrázolásával foglalkozó elméleti nyelvészek érdeklődésének homlokterébe kerültek. A magyar személyragos főnévi igenevekről jelentős tanulmányok születtek itthon és külföldön is.⁹ A kutatások számos olyan kérdést felvetettek, amelyeknek megfogalmazására korábban nem került sor. A legfontosabb kutatási kérdések a 2. táblázatban követhetők nyomon.

2. táblázat

A személyragos főnévi igenevekre vonatkozó legfontosabb kérdések

| | |
|--|--|
| A mai magyar nyelv anyanyelvi beszélői milyen mondatkörnyezetekben használják... | |
| (a) | <i>kizárólag a főnévi igenév személyragos változatát?</i> |
| (b) | <i>kizárólag a főnévi igenév személyrag nélküli változatát?</i> |
| (c) | <i>a főnévi igenév személyragos és személyrag nélküli változatát egyaránt?</i> |
| (d) | <i>Az utóbbi esetben mitől függ az alakok közötti választás?¹⁰</i> |

A 2. táblázat (a)–(d) kérdéseinek megválaszolásához természetesen szükség van a „mondatkörnyezet” fogalmának tisztázására. Ezt nehezíti, hogy nincsen szakmai konszenzus a főnévi igeneveket tartalmazó kifejezések szerkezetének ábrázolásáról.¹¹ Ugyanakkor a téma kutatói egységesen a főnévi igeneves bővítményt megengedő predikátumot tartják a mondatkörnyezet legmeghatározóbb részének. (Ez az 1. táblázat (a) mondatában a „sikerült” szóalak, a táblázat (b) mondatában pedig a „kell”). A kérdések megválaszolásához tehát mindenképpen meg kell határozni azoknak a predikátumoknak a körét, amelyek személyragos, illetve személyrag nélküli főnévi igeneves bővítményeket megengednek.

A legfontosabb lépést e felé Tóth Ildikó Hollandiában készült elméleti nyelvészeti doktori értekezése¹² tette meg, amely vállalkozott arra, hogy megadja azoknak a predikátumoknak a tipológiáját, amelyeknek személyragos főnévi igeneves bővítményei lehetnek. A tipológia céljai szerint teljes, azaz a személyragos főnévi igeneves bővítményeket megengedő predikátumok mindegyikét felsorolja és csoportokba rendezi.¹³

A vizsgálat módszere, célja és eredményei

A mondattannal foglalkozó elméleti nyelvészek nyelvi adataikhoz túlnyomórészt azoknak az ítéleteknek a segítségével jutnak, amelyekkel az általuk szerkesztett mondatokat a vizsgált nyelv anyanyelvi beszélői grammatikusnak (nyelvtanilag helyesnek) vagy agrammatikusnak (nyelvtanilag helytelennek) minősítik. Amennyiben egy elméleti nyelvész anyanyelvével (vagy anyanyelvi nyelvváltozatával) kapcsolatos vizsgálatot végez, gyakran saját nyelvi intuícióira hagyatkozik.

A korpusznyelvészek körében azonban általános az a nézet, hogy a nyelvtudományi vizsgálatoknak egy kifejezetten erre a célra szerkesztett szöveggyűjteményen, korpuszon kell alapulnia. Vagy – ahogyan a szakma vezető folyóiratának főszerkesztője fogalmaz – „a kutató nyelvi intuíción alapuló feltételezéseket korpuszadatokkal kell hitelesíteni”.¹⁴ Emellett szól, hogy a megvalósítandó feladat befolyásolhatja, vagyis elbizonytalaníthatja a vizsgálatot végző személy intuícióját, és hogy az is megeshet, hogy az alapul vett nyelvi adatok kizárólag egy nyelvváltozatra (szélsőséges esetben: a kutató idiolektusára, azaz egyéni nyelvhasználatára) jellemzőek. Ugyanakkor természetesen előfordulhat, hogy a nyelvész a rendelkezésére álló korpuszban nem talál releváns szerkezeteket, mondatokat. Ilyenkor a saját intuíciója segítségével nyert nyelvi adatokat anyanyelvi adatközlőkkel ellenőriztetheti.¹⁵

A Nyelvtudományi Intézetben informatikus kollégámmal, Sass Bálinttal azt vizsgáltuk meg, hogy egy gondosan összeállított nagyméretű korpusz és korszerű számítógépes nyelvészeti eszközök segítségével juthatunk-e olyan nyelvi adatokhoz, amelyek az említett – elsősorban a szerző saját anyanyelvi intuíción alapuló – tipológiában nem szerepelnek. Az adatgyűjtésre azért vállalkoztunk, mert feltételeztük, hogy kutatásunkkal új adatokat is felszínre hozhatunk, tehát hozzájárulhatunk a 2. táblázatban felsorolt kérdések megválaszolásához.

Vizsgálatunk alapjául a Magyar Nemzeti Szövegtár nevű korpusz *első változata* szolgált, amely 1998 és 2002 között készült az MTA Nyelvtudományi Intézetének Korpusznyelvészeti Osztályán. A gyűjtemény főbb jellemzőit a 3. táblázat foglalja össze.

3. táblázat

A Magyar Nemzeti Szövegtár jellemzői

| | |
|-----|--|
| (a) | a mai magyar írott nyelv reprezentatív korpusza; |
| (b) | a sajtónyelvet, a szépirodalmi nyelvet, a tudományos értekező prózát, a hivatali nyelvet és a személyes közlést (öt markánsan elkülönülő nyelvhasználatot) is bemutatja; |
| (c) | a szavakhoz a megfelelő szótót, szófaji meghatározást és alaktani elemzést is hozzárendeli; |
| (d) | eleve elektronikus formában lévő szövegekből készült; |
| (e) | 153,7 millió szövegszót tartalmaz; |
| (f) | internetes keresőfelülete 2002 óta regisztráció után ingyenesen elérhető a http://corpus.nytud.hu/mnsz címen. ¹⁶ |

A konkrét feladatokat, célokat a 4. táblázat mutatja be.

4. táblázat

A személyragos főnévi igenevekkel kapcsolatos vizsgálatunk céljai

| | |
|--|---|
| Állítsuk elő a Magyar Nemzeti Szövegtár anyagából... | |
| (a) | a személyragos főnévi igeneves bővítményeket megengedő predikátumok (a továbbiakban: FIP-ek) listáját, és |
| (b) | az egyes FIP-ekkel előforduló főnévi igenevek listáját. |

A számítógépes nyelvészetben általános módon a feladatokat félig automatikusan, félig manuálisan oldottuk meg, a következő módszerrel.

Először a szavak alaktani elemzése alapján automatikusan kigyűjtöttük a Magyar Nemzeti Szövegtár személyragos főnévi igenevet tartalmazó mondatait. Munkahipotézisünk az volt, hogy a FIP-ek a személyragos főnévi igenévvél azonos tagmondatban fordulnak elő. Mivel a korpuszban a tagmondathatárok nincsenek bejelölve, a tagmondathatárokat közelítő módszerrel próbáltuk megfogni. A közelítő módszerhez a tagmondathatároló írásjelek és tagmondathatározó kötőszavak listáját használtuk fel.¹⁷

A kapott tagmondathatározókból automatikusan kiszűrtük azokat, amelyek a legkorábbi magyar leíró nyelvtenban¹⁸ azonosított FIP-et tartalmaztak, az előforduló személyragos főnévi igenevet pedig feljegyeztük a megfelelő FIP-hez. Ezután következett az eljárás manuális része: a megmaradt tagmondathatározókban új FIP-eket kerestünk, majd az új, kiegészített FIP-listával újra alkalmaztuk az eljárást előlről.

Három iteráció után a következő számszerű eredményekre jutottunk. Szótövesített FIP-listánk 197 tagú. Az ezeket tartalmazó tagmondatok a Magyar Nemzeti Szövegtárban szereplő 228 367 személyragos főnéviigenév-előfordulásból 223 140-et (98 százalék) fednek le. Szótövesítve 17 874 FIP – személyragos főnéviigenév-párt kaptunk.¹⁹ Vizsgálatunk igazolta feltételezésünket: a korpuszból a fenti módszerrel sikerült olyan FIP-eket is kigyűjtenünk, amelyek az eddigi tipológiában nem szerepeltek. Ezek közül mutat be néhányat az 5. táblázat.

5. táblázat

Néhány először általunk azonosított FIP

| | |
|---------------------|---|
| (a) egytagú elemek | adódik, akaródik, megéri |
| (b) többtagú elemek | van hol/honnan/hová/miért/kit/mit/kinek/kiért/miért/mire |

Álljon itt egy-egy olyan mondat is a Magyar Nemzeti Szövegtárból, amely az 5. táblázatban kiemelt FIP-eket tartalmazza (6. táblázat).

Újjonnan talált FIP-eket tartalmazó korpuszmondatok

| | |
|-----|---|
| (a) | Volt, aki különbuszon utazott, de nem volt hol aludnia . |
| (b) | Az edző rá is állt volna a békülésre, csakhogy nem volt kinek edzést tartania... |
| (c) | Volt mire támaszkodniuk a kollégáknak. |

Az újjonnan talált nyelvi adatokat természetesen még rendszerezni, csoportosítani kell. Ez a rendszerezés, csoportosítás reményeim szerint a 2. táblázatban felsorolt kérdések megválaszolását is megkönnyíti majd.

Irodalom

- Bottyán Gergely: A korpuszhasználatban rejlő kiaknázatlan lehetőségekről. In: Klaudy Kinga–Dobos Csilla (szerk.): *A XV. Magyar Alkalmazott Nyelvészeti (MANYE) Kongresszus előadásai*. Miskolci Egyetem, Miskolc, 2005. I. kötet. 271–275. [http://www.nytud.hu/oszt/korpusz/resources/bg_manye2005.pdf].
- Bottyán Gergely–Sass Bálint: Conjugated infinitives in the Hungarian National Corpus. In: Garabik, Radovan (szerk.): *Computer Treatment of Slavic and East European Languages*. Szlovák Tudományos Akadémia, Pozsony, 2005. 27–30. [http://www.nytud.hu/oszt/korpusz/resources/bg_sb_slovko2005.pdf].
- É. Kiss Katalin–Kiefer Ferenc–Siptár Péter: *Új magyar nyelvtan*. 3. kiadás. Osiris Kiadó, Budapest, 2003.
- É. Kiss Katalin: A személyragos alaptagú főnévi igenes kifejezés. In: É. Kiss Katalin–Kiefer Ferenc–Siptár Péter: *Új magyar nyelvtan*. 3. kiadás. Osiris, Budapest, 2003. 120–123.
- É. Kiss Katalin: A főnévi igenév személyragozásának kérdéseiről. In: Bakró-Nagy Marianne–Bánréti Zoltán–É. Kiss Katalin (szerk.): *Újabb tanulmányok a strukturális magyar nyelvtan és a nyelvtörténet köréből. Kiefer Ferenc tiszteletére barátai és tanítványai*. Osiris Kiadó, Budapest, 2001. 42–58.
- É. Kiss Katalin: Mi legyen a magyar nyelvészet Széchenyi-programja? In: *Magyar Tudomány*, 2002. 2. 198–202. [<http://www.matud.iif.hu/02feb/ekiss.html>].
- É. Kiss Katalin: Agreeing infinitives with a case-marked subject. In: *The syntax of Hungarian*. Cambridge University Press, Cambridge, 2002. 210–221.
- Fábrí György: A nyelv kérdése mindenütt felbukkan. Beszélgetés Kenesei Istvánnal, az MTA Nyelvtudományi Intézet igazgatójával. *Világosság*, 2005. 4. 5–12. [<http://www.vilagosság.hu/pdf/20050830112138.pdf>].
- Gábor Kata–Héja Enikő–Mészáros Ágnes: Kötőszók korpuszalapú vizsgálata. In: Alexin Zoltán–Csendes Dóra (szerk.): *MSzNy 2003 – I. Magyar Számítógépes Nyelvészeti Konferencia*. Szegedi Tudományegyetem, Szeged, 2003. 305–306.
- Gervain Judit–Zemplén Gábor: Az elefántcsonttorony átépítése. *Világosság*, 2003. 5–6. 219–224. [<http://www.vilagosság.hu/pdf/20030702151506.pdf>].
- Kenesei István: A nyelvmélettől a nyelvművelésig. A nyelvtudomány egységéről és változatosságáról. *Magyar Nyelvőr*, 2005. 1. 1–12. [<http://www.c3.hu/~nyelvor/period/1291/129101.pdf>].
- Kiefer Ferenc: Paradigmaváltás a nyelvtudományban is? *Magyar Tudomány*, 1990. 9. 1031–1040.
- Kiefer Ferenc: *A magyar nyelv kutatásának általános nyelvészeti vonatkozásai*. Akadémiai székfoglaló, 1988. február 15. Akadémiai Kiadó, Budapest, 2001.
- Kiefer Ferenc: A nyelvtudomány néhány tudományelméleti-módszertani kérdése. In: Ruzsa Imre (szerk.): *Tertium non datur*. Osiris, Budapest, 2000. 123–171.
- Kiefer Ferenc: *A diszciplínák művelése. Nyelvtudomány*. Magyar Tudományos Akadémia, Budapest, 2001.

- Kis Balázs: A nyelv és a számítógép. In: Kenesei István (szerk.): *A nyelv és a nyelvek*. Ötödik, javított, bővített kiadás. Akadémiai Kiadó, Budapest, 2004. 171–189. [http://www.nytud.hu/nyelv_es_nyelvek].
- Miller, D. Gary: Where do conjugated infinitives come from? *Diachronica*, 2003. 1. 45–81.
- Prószéky Gábor: *A nyelvtudomány (és) alkalmazásai*. Aranykönyv Kiadó, Budapest, 2005.
- Prószéky Gábor–Olaszy Gábor–Várad Tamás: Nyelvtudomány. In: Kiefer Ferenc (szerk.): *Magyar nyelv*. Akadémiai Kiadó, Budapest, 2006. 1038–1072.
- Szirmai Mónika: *Bevezetés a korpusznyelvészetbe*. Tinta Könyvkiadó, Budapest, 2005.
- Teubert, Wolfgang: My version of corpus linguistics. *International Journal of Corpus Linguistics*, 2005. 1. 1–13. [<http://www.benjamins.com/jbp/series/IJCL/10-1/art/0001a.pdf>].
- Tóth Ildikó: Can the Hungarian infinitive be possessed? In: Kenesei István–Siptár Péter (szerk.): *Approaches to Hungarian 8: Papers from the Budapest Conference*. Akadémiai Kiadó, Budapest, 2002. 135–160. o.
- Tóth Ildikó: *Inflected infinitives in Hungarian*. PhD-értekezés. University of Tilburg, Tilburg, 2000.
- Várad Tamás: Szótár, korpusz – Magyar Nemzeti Szövegtár. In: Gecső Tamás (szerk.): *Lexikális jelentés, aktuális jelentés*. Tinta Könyvkiadó, Budapest, 2000. 263–270.
- Várad Tamás: Modern nyelvi technológiák a magyar nyelvért. In: Kiefer Ferenc–Gósy Mária (szerk.): *Helyzetkép a magyar nyelvtudományról*. MTA Nyelvtudományi Intézet, Budapest, 2000. 146–156.
- Várad Tamás: A nyelvhasználat empirikus vizsgálatáról. In: Andor József–Szűcs Tibor–Terts István (szerk.): *Színes eszmék nem alszanak... Szépe György 70. születésnapjára*. Lingua Franca Csoport, Pécs, 2001. 1285–1291.
- Várad Tamás és munkatársai: *Magyar Nemzeti Szövegtár*. MTA Nyelvtudományi Intézet, Budapest, 2002. [<http://corpus.nytud.hu/mnsz/>].

Jegyzetek

- 1 A „nyelvtudomány” és a „nyelvészet” szavakat – a bevett gyakorlatnak megfelelően – egymás szinonimáiként használom.
- 2 Az áttekintés szükségképpen elnagyolt lesz. A <http://www.nytud.hu> honlapon az intézet jelen dolgozatban nem tárgyalt osztályainak tudományos tevékenységéről is pontosan tájékozódhat az olvasó. Átfogó képet nyújt a Nyelvtudományi Intézetéről a *Világosság* című folyóirat Akadémiai Műhely-sorozatában megjelent interjú is: A nyelv kérdése mindenütt felbukkan. Kenesei Istvánnal, az MTA Nyelvtudományi Intézet igazgatójával beszélget Fábri György. *Világosság*, 2005. 4. 5–12. [<http://www.vilagosság.hu/pdf/20050830112138.pdf>].
- 3 Vö. Kiefer Ferenc: *A magyar nyelv kutatásának általános nyelvészeti vonatkozásai*. Akadémiai székfoglaló, 1988. február 15. Akadémiai Kiadó, Budapest, 2001. 5–7.; É. Kiss Katalin–Kiefer Ferenc–Siptár Péter: *Új magyar nyelvtan*. 3. kiadás. Osiris Kiadó, Budapest, 2003. 13–14.
- 4 Az elmúlt másfél évtizedben számos olyan tanulmány született az Elméleti Nyelvészeti Osztályon, amely a modern magyar nyelvtudomány célkitűzéseit, módszertani választásait, kihívásait és dilemmáit fogalmazza meg a szélesebb tudományos közönség számára. Ezek közül különösen a következők lehetnek érdekesebb a társadalomtudományok művelői szempontjából: Kiefer Ferenc: Paradigmaváltás a nyelvtudományban is? *Magyar Tudomány*, 1990. 9. 1031–1040.; Kiefer Ferenc: *A diszciplínák művelése. Nyelvtudomány*. Magyar Tudományos Akadémia, Budapest, 2001; É. Kiss Katalin: Mi legyen a magyar nyelvészet Széchenyi-programja? *Magyar Tudomány*, 2002. 2. 198–202. [<http://www.matud.iif.hu/02feb/ekiss.html>]; Kenesei István: A nyelvelméletől a nyelvművelésig: A nyelvtudomány egységéről és változatosságáról. *Magyar Nyelvőr*, 2005. 1. 1–12. [<http://www.c3.hu/~nyelvor/period/1291/129101.pdf>]. A számítógépes nyelvészet területeiről a Nyelvtudományi Osztály profiljának ismertetésénél lesz szó.
- 5 Ez a tevékenység történeti okokból folyik a Kísérletes Nyelvészeti Osztályon: a cigánynyelvi kutatások még az osztály jogelődjén, a 2001-ig létező Alkalmazott Nyelvészeti Osztályon kezdődtek.
- 6 Vö. Prószéky Gábor: *A nyelvtudomány (és) alkalmazásai*. Aranykönyv Kiadó, Budapest, 2005. 11.

- 7 A számítógépes nyelvészetről és a magyar nyelvtechnológia eredményeiről – Prószéky Gábor már idézett művén kívül – többek között a következő források adnak a fentieknél bővebb, de tömör áttekintést: Váradi Tamás: Modern nyelvi technológiák a magyar nyelvért. In: Kiefer Ferenc–Gósy Mária (szerk.): *Helyzetkép a magyar nyelvtudományról*. MTA Nyelvtudományi Intézet, Budapest, 2000. 146–156.; Kis Balázs: A nyelv és a számítógép. In: Kenesei István (szerk.): *A nyelv és a nyelvek*. Ötödik, javított, bővített kiadás. Akadémiai Kiadó, Budapest, 2004. 171–189. [http://www.nytud.hu/nyelv_es_nyelvek/12fej.pdf]; Prószéky Gábor–Olasz Gábor–Váradi Tamás: Nyelvtechnológia. In: Kiefer Ferenc (szerk.): *Magyar nyelv*. Akadémiai Kiadó, Budapest, 2006. 1038–1072.
- 8 Vö. Miller, D. Gary: Where do conjugated infinitives come from? *Diachronica*, 2003. 1. 45–81.
- 9 A következő könyvfejezetben és dolgozatban a tárgyban végzett összes eddigi elméleti nyelvészeti kutatás eredményei megtalálhatók: É. Kiss Katalin: Agreeing infinitives with a case-marked subject. In: *The syntax of Hungarian*. Cambridge University Press, Cambridge, 2002. 210–221.; Tóth Ildikó: Can the Hungarian infinitive be possessed? In: Kenesei István–Siptár Péter (szerk.): *Approaches to Hungarian 8: Papers from the Budapest Conference*. Akadémiai Kiadó, Budapest, 2002. 135–160.
- 10 A 2. táblázatban felsorolt kérdések – továbbiakkal együtt – magyarul itt jelentek meg először: É. Kiss Katalin: A főnévi igenév személyragozásának kérdéseiről. In: Bakró-Nagy Marianne–Bánréti Zoltán–É. Kiss Katalin (szerk.): *Újabb tanulmányok a strukturális magyar nyelvtan és a nyelvtörténet köréből*. Kiefer Ferenc tiszteletére barátai és tanítványai. Osiris Kiadó, Budapest, 2001. 42.
- 11 Vö. Tóth Ildikó: i. m. (2002). 135.
- 12 Tóth Ildikó: *Inflected infinitives in Hungarian*. PhD-értekezés. University of Tilburg, Tilburg, 2000.
- 13 A tipológiát a szerző következő dolgozata foglalja röviden össze: Tóth Ildikó: Can the Hungarian infinitive be possessed? In: Kenesei István–Siptár Péter (szerk.): *Approaches to Hungarian 8: Papers from the Budapest Conference*. Akadémiai Kiadó, Budapest, 2002. 135–139. A tipológia magyar nyelvű összefoglalása: É. Kiss Katalin: A főnévi igenév személyragozásának kérdéseiről. In: Bakró-Nagy Marianne–Bánréti Zoltán–É. Kiss Katalin (szerk.): i. m. 45–47.
- 14 Eredetiben: „No introspection can claim credence without verification through real language data.” Teubert, Wolfgang: My version of corpus linguistics. *International Journal of Corpus Linguistics*, 2005. 1. 1. [<http://www.benjamins.com/jbp/series/IJCL/10-1/art/0001a.pdf>].
- 15 Vö. Kiefer Ferenc: A nyelvtudomány néhány tudományelméleti-módszertani kérdése. In: Ruzsa Imre (szerk.): *Tertium non datur*. Osiris, Budapest, 2000. 127–128. Kapcsolódó módszertani kérdéseket, problémákat tárgyal: Gervain Judit–Zemplén Gábor: Az elefántcsonttorony átépítése. *Világosság*, 2003. 5–6. 219–224. [<http://www.vilagosság.hu/pdf/20030702151506.pdf>].
- 16 A korpusz összetételére vonatkozó pontos adatok (*Bevezető*), a regisztrációs űrlap, a felhasználói fórum és a lekérdezésekkel kapcsolatos legfontosabb tudnivalók (*Súgó*) ugyanezen a honlapon olvashatók, a szerkesztési munkálatokban részt vevő munkatársak felsorolásával együtt. A gyűjtemény összeállításáról lásd még Váradi Tamás: Szótár, korpusz – Magyar Nemzeti Szövegtár. In: Gecső Tamás (szerk.): *Lexikális jelentés, aktuális jelentés*. Tinta Könyvkiadó, Budapest, 2000. 263–270. A reprezentativitás fogalmáról, valamint egy teljes beszédközösség nyelvhasználatát bemutató korpusz szerkesztésének gyakorlati és elvi nehézségeiről Váradi Tamás: A nyelvhasználat empirikus vizsgálatáról. In: Andor József–Szűcs Tibor–Terts István (szerk.): *Színes eszmék nem alszanak... Szépe György 70. születésnapjára*. Lingua Franca Csoport, Pécs, 2001. 1285–1291. A Magyar Nemzeti Szövegtár és más nyelvi erőforrások nyújtotta lehetőségekre mutat rá egyszerű példákkal a következő írás: Bottyán Gergely: A korpuszhasználatban rejlő kiaknázatlan lehetőségekről. In: Klaudy Kinga–Dobos Csilla (szerk.): *A XV. Magyar Alkalmazott Nyelvészeti (MANYE) Kongresszus előadásai*. Miskolci Egyetem, Miskolc, 2005. I. kötet. 271–275. [http://www.nytud.hu/oszt/korpusz/resources/bg_manye2005.pdf]. A dolgozatban a korpuszok összeállításával és a korpuszalapú kutatásokkal kapcsolatos magyar nyelvű szakirodalom közel teljes tematikus bibliográfiája is megtalálható. A témában megjelent első magyar nyelvű monográfia elsősorban a korpuszok nyelvtanulásban és nyelvtanításban való felhasználásához ad tippeket: Szirmai Mónika: *Bevezetés a korpusznyelvészetbe*. Tinta Könyvkiadó, Budapest, 2005. A könyv számtalan idegen nyelvű korpusz webhelyét és a legfontosabb angol nyelvű szakirodalmat is megadja.

- 17 A listát a következő cikkből vettük: Gábor Kata–Héja Enikő–Mészáros Ágnes: Kötőszók korpuszalapú vizsgálata. In: Alexin Zoltán–Csendes Dóra (szerk.): *MSzNy 2003 – I. Magyar Számítógépes Nyelvészeti Konferencia*. Szegedi Tudományegyetem, Szeged, 2003. 305–306.
- 18 É. Kiss Katalin: A személyragos alaptagú főnévi igeneves kifejezés. In: É. Kiss Katalin–Kiefer Ferenc–Sip-tár Péter: *Új magyar nyelvtan*. 3. kiadás. Osiris, Budapest, 2003. 120–123.
- 19 Vö. Bottyán Gergely–Sass Bálint: Conjugated infinitives in the Hungarian National Corpus. In: Garabik, Radovan (szerk.): *Computer Treatment of Slavic and East European Languages*. Szlovák Tudományos Akadémia, Pozsony, 2005. 27–30. [http://www.nytud.hu/oszt/korpusz/resources/bg_sb_slovko2005.pdf].

Summary

GERGELY BOTTYÁN
HUNGARIAN INFINITIVAL CONSTRUCTIONS:
THE PLACE OF A RESEARCH TOPIC IN (COMPUTER-AIDED)
LINGUISTICS

The present paper consists of two main parts. The first part surveys the directions of interdisciplinary research carried out in the Research Institute for Linguistics in Budapest. Five departments (out of ten) in the institute are presently engaged in interdisciplinary research. These are, in alphabetical order, the Department of Language Technology (called Department of Corpus Linguistics until 2006), the Department of Phonetics, the Department of Psycholinguistics and Neurolinguistics, the Department of Sociolinguistics and the Department of Theoretical Linguistics. The survey shows that the research activities in the institute are dominated by the two main international trends detectable in linguistics since the second half of the 20th century, namely specialization in the core areas of modern linguistics and increasing collaboration with related fields. Among the disciplines that have had a lasting impact on or have collaborated with linguistics not only social sciences, but also cognitive, mathematical and technical sciences can be found.

The second part of the paper presents the aims, the methodology and the accomplished results of the author's own research on Hungarian infinitival constructions. A computer-aided systematic data collection procedure is reported on which uncovered linguistic data not yet described in the theoretical linguistic literature. This procedure was applied to the Hungarian National Corpus, a representative collection of present-day Hungarian written texts. The classification of the newly found data, which remains to be done in the future, is hoped to make a significant contribution to the understanding of the use of Hungarian infinitival constructions.