

## Összegzés

Hagyomány és kísérlet ötvöződik a konferencián előadott és írásban közzétett tanulmányokban. A konklúziókra, megállapításokra vonatkozó minősítésnek három fokozata lehet: egészen hiteles, részben elfogadható és kétséges – tehát nem minden tekintetben meggyőző. A kötet tanulmányainak eklektikus jellege ellenére dicsérendő érdemei vannak például a mögötte rejlő szakmai tudást tükrözi, a kutatómunkához szükséges elmélyedést és azt a fogékonyságot, amely feltétele az „új” felfedezésének. Az „új” jelen van a kötetben akár mint módszer, korpusz, elméleti háttér vagy megállapítás, akár mint a lehetséges további vizsgálatok prognózisa.

Elismeréssel szólhatunk a támogatott kutatás részeredményeiről, eredményeiről, a vezetőről és a munkatársakról, akik ezt a tartalmas szellemi terméket létrehozták.

*Rozgonyiné dr. Molnár Emma*

kandidátus, ny. főiskolai tanár  
magister emerita

**Bollettino. Dizionari e ricerca filologica.** Atti della Giornata di Studi in memoria di Valentina Pollidori, Supplemento III, Opera del Vocabolario Italiano, 2012. 232 lap

Mára jelentős mértékben elterjedtek a nyelvészeti kutatásokban használt különféle számítógépes alkalmazások, eszközök. Mindig különleges helyzetben volt és van e tekintetben a nyelvtörténet, hiszen a gépi szabályszerűséget a történeti szövegek egyediségével kell ötvözni. Ha nincsenek illúzióink, akkor a számítógép ebben az ambivalensnek tűnő helyzetben is nagyon sokat tud segíteni. Ennek az optimális munkamegosztásnak az egyik kiváló példája az olaszországi kutatóintézet-hálózat (CNR) firenzei székhelyű szótári központjában (Opera del Vocabolario Italiano = OVI) folyó munka. Az itt dolgozó munkatársak évente számot adnak a szakma és a nagyközönség számára a legújabb eredményeikről, szótárkészítői feladataikról, aktuális projekteikről, fejlesztéseikről. A 2012-es tanulmánykötet különös jelentőségű: egyfelől a kutatók bemutatják, összegzik és értékelik az egyes területeken elért eddigi eredményeiket, másfelől ez az összegzés egyben tiszteletadás is a korábban elhunyt és az intézményben sokáig meghatározó és nélkülözhetetlen munkát végző kiváló kutató, Valentina Pollidori számára is.

A kötet tíz tanulmányt tartalmaz, amelynek fele az intézeti kutatóktól származik, akik a szótári munka egy-egy részterületéről adnak számot. Ezekben a tanulmányokban javarészt konkrét példákon bemutatva ismerhetjük meg az elvi és módszertani hátteret. A tanulmányok felépítése arányos, megfogalmazásukban lényegre törőek, tömörek, rendkívül tartalmasak és informatívak. Szakcikként önállóan is megállják a helyüket, de egybeszerkesztve komplex képet tárnak elénk a mai modern szótárkészítésről. A tanulmányok másik fele olyan szakemberektől származik, akik tudományos munkáikban felhasználták, illetve továbbgondolták a történeti korpusz és a szótár adta lehetőségeket. Ezek a tanulmányok tematikailag és megfogalmazásukban már jóval heterogénebb írások. Az ismertetésben nem a kötetbeli sorrendet követem, hanem e két csoport alapján mutatom be a tanulmánygyűjteményt.

Pietro Beltrami a kötet előszavában méltatja elhunyt kollégája munkásságát, és egyben összegzi az intézet eddig elért eredményeit. Pär Larson és Elena Artale a korpuszalapú történeti szótár (TLIO) bemutatása, majd a korpuszépítés kezdeti szakaszának vázolása után a közelmúlt és a jelen állapotát tárja elénk. Részletesebben is szó esik a történeti korpuszok állandó nehézségéről, a filológiai problémákról, például a kéziratok különleges karaktereinek a rögzítéséről, az eltérő szövegkiadások filológiai hitelességének a megállapításáról, a kézirattal való összevetésekről, a szövegkiadásokban lévő hibák javításairól és a gép számára feldolgozható egységes forma kialakításáról. A szerzők a lemmatizálás kérdéseit is taglalják, és az eredmények mellett a problematikus területeket sem hallgatják el, például a nyelvjárási változatok egységes címszó alá rendezésének a kérdését. A grammatikai szemléletváltozások után mára már letisztult az elméleti háttér: szoftveres segítséggel immár a mai címszavak alapján lekérdezhető korpusz áll a felhasználók rendelkezésére. A szerzők összevetették

a lemmatizált korpusz korábbi és jelenlegi számszerűsíthető adatait a feldolgozott szövegek, szólófordulások, szóalakok, címszavak, lemmatizálatlan formák tekintetében. A kapott eredményekből egyértelműen látszik, hogy az elvi változtatások mellett az általuk fejlesztett félautomata lemmatizáló szoftver is segíti, javítja és gyorsítja a munkát.

Az informatikai háttér alaposabb feltárásához Domenico Ioro-Fili és Andrea Boccellari tanulmányai nyújtanak segítséget. Előbbi szerző a lemmatizálást segítő, speciálisan e munka számára készült szoftver működését mutatja be. A cikk részletesen taglalja a szoftver lehetséges alkalmazásainak a széles körét. A lemmatizálás félig automatikussá tétele komoly előrelépés nemcsak a szótári munka egységesebbé, pontosabbá és gyorsabbá tétele érdekében, hanem a történeti szövegek számítógépes feldolgozásában is, hiszen ez az egyik legproblematisabb és leginkább időigényes terület. Az eljárás alkalmazza részint az automatikus lemmatizálók módszereit, részint a már meglévő elemzett korpusz leírásait, tanulságait, de az emberi ellenőrzés sem maradhat el. A cikk illusztratív ábrák segítségével részletesen tárgyalja a szoftver működési elveit, valamint az alkalmazott algoritmusokat is. Andrea Boccellari a szócikkíráshoz használt számítógépes támogatásról és a szócikkek interneten való megjelenítési módjáról számol be. Részletesen bemutatja az e területen végzett fejlesztéseket, amelyek nemcsak felgyorsítják a szócikkírói munkát, hanem az ellenőrzési, javítási feladatokat is nagymértékben redukálják. Ezek az alkalmazások lexikográfusok és számítástechnikai szakemberek kooperációjában jöttek létre. A technikai fejlesztések alapvetően a lexikológiai munka támogatását szolgálják, nem kerülnek túlsúlyba, ily módon a pontos, hiteles filológiai feldolgozás az elsődleges feladat. E két tanulmány bár az informatikai hátteret taglalja, de a területen kevésbé jártas olvasó is könnyen megértheti a fő koncepciót.

Paolo Squillacioti írásának elején áttekinti a TLIO szótár történetét, majd összegzést ad a szótár jellegéről, típusáról és a szótárkészítés alapjául szolgáló történeti korpuszról is. Szócikkek bemutatásával vizsgálja a szócikkírás főbb részeit, esetleges problémáit. Bemutatja a szerkesztési folyamatot és az ezt támogató saját fejlesztésű informatikai segédeszközöket is. Megismerhetjük a munkacsoport felépítését és a más egyetemekkel kialakított együttműködések; a címszavak feldolgozásának módszerét; az online megjelenés velejáróit; a jövőbeni terveket. A tanulmány írója végezetül kiemeli a TLIO szótár és a reprezentatív történeti korpusz fontosságát és egymásrautaltságát az európai történeti lexikográfiában.

Rossella Mosti tematikailag szorosan az előző tanulmányhoz kapcsolódó írásában az előszerkesztési munkálatokat mutatja be. A tanulmány szerzője a lemmatizálás és a tényleges szócikkírás közti szakaszban felmerülő problémákat tárja fel. Konkrét példák kapcsán láthatjuk, hogy az alaposabb etimológiai vizsgálatok, akár a homográfok hangsúlybeli eltérései is, módosíthatják a korábbi lemmatizálás eredményeit. Az előszerkesztés ráirányíthatja a figyelmet szófajbeli, szónembeli specifikumokra, amelyek a szokványostól való eltérést jelölik, és ezáltal módosítják a jelentésleírást is. A cikk írója azokat a területeket mutatja be egy-egy illusztratív példa segítségével, amelyek alapos filológiai kutatómunkát igényelnek: például a szövegkiadások értelemszertező hibáinak a javítását a kéziratok alapján; szócikkek egybeolvasztását, illetve kettéválasztását az etimológia és a jelentés összefüggései alapján; a szövegtörzsetben ugyan nem szereplő, de a szótárba felveendő címszavak listázását; az első előfordulások dátumának ellenőrzését, javítását. A tanulmány szerzője felhívja a figyelmet arra, hogy a TLIO szótárban több esetben korábbra datálódik a szavak első előfordulása más szótárakhoz képest, és ezért a TLIO a többi szótár számára is hivatkozási alapként is szolgál. Eredményeiből kiderül, hogy a szócikkek készítése során több korábbi megválaszolatlan, feltáratlan történeti, filológiai problémára is választ kaphatunk.

A kötet szerkesztőinek érdeme, hogy nemcsak szótárkészítői, hanem felhasználói aspektusból is közölnek tanulmányokat. Ezek az írások egyrészt újabb szempontjaikkal egészítik ki az eddigieket, másrészt rávilágítanak a szótári munka eredményeinek széles körű alkalmazhatóságára is a nyelvtudományban. Edgar Radtke az olasz nyelvtörténet elhanyagolt területével foglalkozik: a diskurzus elemeivel a diakróniában. Nyelvtörténeti szempontból nehezen kutatható területről van szó, hiszen a langue szintjén nem kötelezően megjelenő elemeket vizsgál. A beszélő szubjektivitásából táplálkozó diskurzuspartikulák kutatása speciális módszert igényel, hiszen a régi nyelvezetet csak írásos szövegek alapján lehet kutatni. A szerző a történeti szótár szócikkeit vizsgálva megállapítja, hogy a régi írott szövegekben jól megfigyelhető a diskurzusok szerkezeti felépítése, amelynek a megismerése hozzájárul a nyelvtörténet alaposabb feltárásához. Felvázolja a szövegek elemzésére

vonatkozó elképzeléseit: a szintaktikai pozíció meghatározásától a kommunikatív funkció definiálásáig. Kiemeli a szövegkorpusz jelentőségét abban a tekintetben, hogy az adatok mennyiségi értékelése is lehetővé válik. A szótári teljes feldolgozásban a szavak szemantikai jellemzői mellett a kommunikációban betöltött szerepet is feldolgozandónak tartja.

Hogy a történeti korpusz nemcsak a szótárkészítés szövegforrásául szolgálhat, hanem más nyelvtörténeti kutatás alapjául is szolgálhat, kiderül Roberta Cella enklitikus elemekről írott tanulmányából is. A szerző az 1300-as évek Firenzéjének névmási szerkezetét vizsgálja a maguk összetettségében a TLIO szövegkorpusza alapján. Az összetett névmási szerkezeteket típusonként elemzi, tanulmányában részletesen taglalja ezek változásait, megjelenési formáit. A korpusznak köszönhetően statisztikai elemzést is készít. Munkája jelentős részét a teljes példaanyag rendszerezése teszi ki, így a bő terjedelem inkább bemutatás, mint alaposabb összefüggések feltárása.

A számítógépes szövegfeldolgozás bár új távlatokat nyit a források közzétételében, ehhez eddig nem vagy csak részben feltárt és megoldott filológiai problémák társulnak. Új módszerek kidolgozása válik szükségessé, amelyek még jobbra a tervezés fázisában járnak. E terület néhány lényeges kérdését taglalja a következő két írás is. Lino Leonardi a sienai egyetem munkatársaként a filológia és a lexikográfia közt húzódo sajátos kapocsról ír. Kiemeli a metaadatok fontosságát: az adott szóhoz kapcsolódó kronológiai, nyelvföldrajzi, filológiai információkat, és javasolja, hogy a problémás esetekben tanácsos ezeket az adatokat a kézirat adataival helyettesíteni. A bibliográfiai adatokat olyan adatbázisban rögzítené, ahol a kiadások, a kéziratok, a kronológiai, nyelvföldrajzi a adatok együtt szerepelnek minden szöveg és minden szerző esetében. Elismerően szól a TLIO filológiai adatbázisáról, és utal arra, hogy a kéziratos szövegek digitalizálása, összekapcsolása és a hozzájuk rendelt bibliográfiai adatok rendszerszerű kezelése jelenti a modern filológia és lexikográfia útját. Lényeglátóan fogalmazza meg a történeti szövegek digitalizálása során felmerülő problémákat, amelyek megválaszolása más hasonló munkálatoknak is tanulságul szolgálhat.

Alessandro Pancheri tanulmányában a TLIO a *Studi di Filologia Italiana (SFI)* című folyóirat közti kapcsolatról ír. A szótárban mintegy 40 szöveg és azok bibliográfiai hivatkozásai származnak az említett folyóiratból. A szerző a folyóirat és a szótár munkatársainak, munkáinak összefonódását taglalja. Méltatja a szövegek alapján definiált lexikográfiai kincseket, az első előfordulásokat, a hapaxokat, ugyanakkor néhány jelentéshöz nyelvtörténeti megjegyzéseket fűz, és egyúttal elismeri a nyelvtörténet kezdeti időszakából származó nyelvelmékek esetében felmerülő jelentésleírással kapcsolatos bizonytalanságokat.

A tanulmánykötet végigolvasásával aligha marad hiányérzetünk a számítógépes történeti lexikográfia újdonságaiból. A kötet jól felépített, a tanulmányokban nincsenek sorozatosan visszatérő információk, így kerek egészet alkotva egységes kép rajzolódik ki belőlük. Ritka az ehhez hasonló összefoglaló tanulmánykötet, amelyben a nyelvtörténészek, a filológusok, a lexikográfusok és a számítógépes szakemberek közös munkájukat bemutatva, de egyúttal a saját tudományterületükön elért újdonságokat is közreadva tárják elének a legfrissebb eredményeiket. Az írások révén az olvasó hiteles képet kap a mai modern szótárkészítés főbb fázisairól, és egyúttal felismerhetővé válik, hogy a szócikkek elkészítéséhez összetett tudásra és komplex képességekre van szükség: a filológiai érzékenység, a nyelvtörténeti ismeretek, a rendszerlátás, az absztrahálás, a számítástechnika adta lehetőségek ismerete, mind-mind szükségeltetnek ehhez. A kötet egyértelműen példázza, hogy a számítógépes szótárkészítés sokkal többet tud nyújtani egy-egy szócikk bemutatásánál: a létrehozott korpusz, a különféle adatbázisok, a szerkezetileg rögzített tartalmak kialakítása stb., mind eddig feltáratlan összefüggéseket mutatnak. A cikkek rávilágítanak a történeti szótár és a korpusz egymásrautaltságára. Amellett, hogy a történeti szövegkorpusz létrehozása önmagában is jelentős eredmény, hiszen filológiai kérdések rendszerszerű feldolgozása társul hozzá, számos, eddig el nem végzett vagy el nem végezhető kutatás alapja is egyben.

Kiss Margit

tudományos munkatárs

MTA Irodalomtudományi Intézet