

Tanulmányainkkal a húsz éve létrejött nyelvészeti kutatóállomásokat és a Termini Magyar Nyelvi Kutatóhálózatot köszöntjük. A nyelvi irodák sokoldalú nyelvészeti tevékenységéből jelenleg kizárólag a lexikológiai, lexikográfiai és terminológiai kutatások egy-egy részeredményét mutatjuk be. A közös témakör és a többszerzős cikkek létrejötte a külső régiók folyamatos együttműködését is fémjelzi.

Benő Attila
egyetemi tanár
Babeş-Bolyai Tudományegyetem
Magyar és Általános Nyelvészeti Tanszék
<https://orcid.org/0000-0002-6931-2999>

A Termini magyar–magyar szótár és adatbázis szerkezeti és szerkesztési sajátosságai

1. Bevezetés

A Termini magyar–magyar szótár és adatbázis¹ a külső régiók magyar nyelvhasználatát dokumentálja. A lexikográfiai adattár nemcsak kontaktusjelenséget adol, széles körben használt közvetlen és közvetett kölcsönszavakat, hanem olyan megőrzött régiégeket is, amelyek korábban az egész magyar nyelvterületen standard jellegűek voltak. A közhasználatú lexikális elemek mellett a szótár regionális köznyelvi, valamint népnyelvi szavakat is tartalmaz. Mivel tehát a szótározott nyelvi formák nem kötődnek csak bizonyos tájnyelvekhez, társadalmi dialektusokhoz vagy regiszterekhez, a szótár ilyen szempontból a mai nyelv általános szótárának tekinthető. Ugyanakkor a szótár régebbi nyelvi formák dokumentálásával jól tükrözi a nyelvi hagyományt, a neologizmusok bemutatásával pedig a nyelvi változást, a megújulást.

Mint adatbázis, mindegyik címszó használatának szemléltetésére lehetőség szerint előnyelvi példamondatokat, szövegrészleteket tartalmaz, lehetővé téve, hogy a szó vagy szerkezet használati sajátosságait grammatikai és pragmatikai kontextusban vizsgáljuk.

A szótár szerkesztői a Termini Magyar Nyelvi Kutatóhálózat tagjai.² Az online szerkesztés lehetővé teszi, hogy a szótár nyolc országban folyó tevékenység alapján – akár egy időben – bővüljön.

¹ <http://termini.nytud.hu/htonline/htlista.php?action=firstpage>

² Jelenleg az erdélyi vonatkozású szavak gyűjtését és lexikográfiai kidolgozását Benő Attila, a felvidékiekét Lanstyák István és P. Márkus Katalin, a vajdaságiakét Molnár Csikós László, a kárpátaljaiakét Márku Anita, a horvátországiakét Lehocki-Samardžić Anna, a muravidékiekét Kolláth Anna és Gaál Péter, az örvidékiekét pedig Szoták Szilvia és Sági Zsuzsa végzi. A megjelenítendő szavak, kifejezések magyarországi nyelvváltozatokban való előfordulását, illetve újabb magyarországi neologizmusok feltérképezését Nádor Orsolya, a szótár informatikai fejlesztését és a felmerülő technikai problémák megoldását a kolozsvári Szabó T. Attila Nyelvi Intézet munkatársa, Juhász Tihamér végzi. Ezenkívül Kitlei Ibolya végez folyamatosan karbantartási munkákat. A munkaközösség szótári tevékenységét az eddigiekben folyamatosan támogatta a Domus Hungarica Scientiarum et Artium akadémiai ösztöndíjprogram.

A szótárról több mint egy évtizede jelennek meg tanulmányok. Ezek részben általánosabb jellegű szótári bemutatások, leírások (Benő 2017, 2019; Benő–Lanstyák–Juhász 2020; Lanstyák 2009c, Lanstyák–Benő–Juhász 2010; M. Pintér 2017b, 2019a, 2019b), másrészt a szótár anyaga alapján készült olyan publikációk, amelyek a lexicográfiai kutatások eredményeit, terminológiai kérdéseit ismertetik (Benő 2014, 2015; Benő–Lanstyák 2019; Kolláth 2011, 2012; Molnár Csikós 2015, 2017; Szoták 2018, 2019).

Tanulmányunkban a szótár újszerűségeit, sajátosságait szeretnénk részletesebben áttekinteni a következő témaköröket érintve: szótári szerkezet, informatikai háttér, adatbázisjelleg, multimédiás megjelenítés, illetve az alkalmazott minősítési rendszer.

2. Szerkezeti kérdések

A Termini-szótár számos tekintetben eltér az egynyelvű magyar szótáraknál megszokott szótárszerkezeti megoldásoktól (például a mikrostruktúra részletessége, kép- és hanganyagok becsatornázása).

A szócikkek mikrostruktúrája³ három – a szótárban egymástól tipográfiaileg is elkülönülő – részből áll: szócikkfej, szócikktörzs, szócikkkláb.

A *szócikkfej* elemei két kategóriába sorolhatók: azok az elemek, amelyeket a szótár minden szócikke tartalmaz (*kötelező elemek*: címszó, szófaj, grammatikai kategóriák / fontosabb nyelvtani alakok), illetve azok az elemek, amelyek csak bizonyos típusú szócikkeknél szerepelnek (ott, ahol azok feltüntetése nyelvészetileg indokolt; *fakultatív elemek*, például az írásváltozat vagy a kiejtés).

A szócikkek összetettségét a szócikktörzs információgazdagsága bizonyítja. Ez két, egymástól elkülönülő információcsoportot, a címszó értelmezéséhez szükséges információkat és a használatot bemutató, illetve jelentést pontosító példamondatot tartalmaz. A címszó értelmezését a jelentések száma, a fogalomkörü minősítések, a földrajzi nyelvváltozat megnevezése, illetve a stílusminősítések segítik.

A szótár – magyar nyelvváltozatokat feldolgozó szótárak között egyedülként – a példamondatok eredetével kapcsolatosan sajátos tipológiát alkalmaz, amely kitér az írott-beszélt distinkcióra, a példamondat származására (valós vagy konstruált), valamint a beszélő nyelvi tudatosságára (nyelvész vagy nem nyelvész beszélő). A szótár példamondatai nemcsak a címszó denotatív jelentéseinek és jelentésárnyalatainak pontosabb bemutatását szolgálják, hanem – főleg ott, ahol nagyobb számú példamondat található – segítségükkel viszonylag megbízhatóan megítélhető a címszó stilisztikai, nyelvhasználati értéke.

A Termini-szótár példamondatainak nagy száma, illetve jellegük szerinti részletes minősítése jól mutatja a szócikkek online megjelenítésében rejlő eddig kiaknáztatlan lehetőségeket.

³ A Termini-szótár szócikkeinek többsége önálló szócikk, azaz megtalálható benne a szócikkfej, szócikktörzs és a szócikkkláb is. A kisebb számban előforduló utaló szócikk a címszón kívül csupán utalást tartalmaz arra az önálló szócikkre, amelyben a címszóra vonatkozó információk megtalálhatók.

3. Informatikai háttér

A szótár és szerkesztési eljárásai számos tekintetben eltérnek a hagyományos értelemben vett szótárszerkesztési eljárásoktól – a szótár kizárólag online felületen létezik, egy időben nyolc országból szerkesztik, a változtatások pedig azonnal látványosan a lekérdezésben, amely értelmezhető akár egy papírszótár oldalainak is. A többfelhasználós platform alapja egy relációs adatbázis, amelyhez a bemeneti (szerkesztőfelület) és a kimeneti (lekérdezőfelület) oldalt egy PHP-nyelven írt többfunkciós felület biztosítja (a lekérdezés szabványos HTML-oldalon jelenik meg). Ez az összetett technológiai megoldás, amelyet a mai lexikográfiai és terminográfiai munkák már széles körben alkalmaznak, a Termini-szótár készítésének kezdeti fázisában még újdonságnak számított.

Mind a relációs adatbázis, mind az online felület lehetőséget biztosít statisztikák készítésére és információk kinyerésére adatbányászati technikák felhasználásával; papíralapú szótáraknál, illetve csak egy-egy szócikk lekérdezésére alkalmas felületnél erre nincs lehetőség. Ezenkívül a relációs adatbázis és az online felület összekapcsolásának köszönhetően a szócikkekhez képek és hangfájlok is kapcsolhatók. A szótár mögötti adatbázis nyelvészeti, elsősorban lexikográfiai szempontból rendkívül gazdag szerkezetű – igyekszik minél árnyaltabb képet adni a címszó mögötti lexikális és konceptuális tartalmakról.

A szótár használata kétféle scenárió mentén történik: a nem regisztrált felhasználók (a nagyközönség) által használható szótár lekérdezőfelülete (*Keresés* mező) kevesebb keresési lehetőséget biztosít, mint a regisztrált felhasználók számára biztosított felület. Ugyanakkor a regisztrált felhasználók lekérdezőfelülete sem tartalmazza az adatbázis tábláinak összes mezőjét, csupán azokat a tulajdonságokat, amelyek nyelvileg fontos adatot tartalmaznak: a lekérdezés folyamán lesz adatból információ, amely lexikográfiai, lexikológiai vagy adatbányászati szempontból is érdekes lehet.

A technológia meghatározza a szótár szerkesztését: a papíralapú szótárak lineáris (akár XML-alapú) szerkesztési folyamatát felváltotta a technológiailag fejlettebb, adatbázis-alapú munkafolyamat, ahol a szerkesztés folyamán létrehozott változások azonnal láthatók (a szerkesztés folyamatáról a szerzők e-mailben értesíthetők), a szerkesztés folyamán használt felület (nézet) azonban jelentősen eltér a véglegesen megjelenített szócikkek vizualizációjától.

Az osztott adatbázis lehetővé teszi, hogy a szócikkeket helytől függetlenül lehessen szerkeszteni. A szerkesztőfelület nem jelent nagy kihívást a használók és a szerkesztők számára sem grafikailag, sem technikailag, mivel könnyen és egyszerűen kezelhető. A szerkesztők jogosultak bármelyik szócikket szerkeszteni vagy a szócikkekhez (csak a szerkesztők számára hozzáférhető) megjegyzést fűzni.

4. Interaktív szótár

A szótár nem pusztán lexikai gyűjtemény, hanem a nyelvhasználat, nyelvváltozatok lexikai és morfológiai adatbázisa – mint ilyen, adatbányászatra (lásd M. Pintér 2017a, b), valamint a tárolt tartalom különféle formában, rendezési elv alapján történő megjelenítésre is alkalmas (akár szótárformában is).

A háttérben működő adatbázisból különféle lehetséges utakat bejárva a felhasználó a lekérdezésnek megfelelő struktúrájú szócikkeket kap. A lexikai adatot és információt keresve három felület közül választhat, így juthat el a saját maga által keresett információt tartalmazó „saját szótár” megjelenítéséig (a szótár készítésének gyorsaságát leginkább az internetkapcsolat, a felhasználó számítógépének, illetve böngészőjének a sebessége, azaz a szerkesztéstől független folyamatok befolyásolják). A nézetek (azaz a felhasználó számára biztosított lekérdezőfelületek, valamint a megjelenített szótár) elsősorban a regisztrációhoz kötött jogosultságoktól függenek: a regisztrált és különféle jogosultságokkal rendelkező felhasználók bővebben is ki tudják használni az adatbázist. A honlap nyilvános felülete⁴ felhasználóbarát-módon csupán az alapfunkciókat felkínáló *Keresés* mezőt tartalmazza, ahol a felhasználó beállíthatja, hogy a keresett kifejezés a szócikk mely részében, illetve a keresett részegység melyik részére (eleje, közepe, vége vagy teljes egésze) illeszkedjen, emellett beállíthatja a megjelenítendő példamondatok számát is (egy sem, kettő, mind).

A szótár elsősorban nyelvészek által használt része a *Szójegyzékek* menü, amely *lexikográfiai adatbányászat*ként működik. A kereső/lekérdező a szócikkstruktúra bármelyik elemében kereshet (a keresést legördülő menük segítik).

A *szójegyzékek generálása* menüben többféle szójegyzék is előállítható, amelyek létrehozását négy alaptípusba sorolt kategóriaegyüttes teszi lehetővé: így lehetséges a szerkesztés időszakára vonatkozó szócikkek (*Időszak*), a szótárral foglalkozó nyelvészek lexikográfiai kutatásaihoz szükséges szócikkek (*Kutatási listák*), a szótári mikrostruktúra bizonyos részeit érintő, illetve a lexikográfiai kutatásokat segítő szócikkek (*Fejlesztési listák*), valamint a fejlesztők munkásságát, illetve az aktuális állapotát tükröző listák (*Dokumentációs listák*) előállítására. A listák, illetve szócikkek megjelenítése a szócikk mikrostruktúrája és metaadatai mentén paraméterezhető.

A nézetek sajátos típusa a *Megjegyzések* rovat/elem, amelybe a szerkesztés folyamatán a szócikkekhez megjegyzések fűzhetők (a szerkesztéshez szükséges kommunikáció is online történik), a „félkész” állapotban lévő szócikkek pedig *rejtett* szócikké alakíthatók, amelyek így csak a szerkesztők számára jelennek meg, a regisztráció nélküli használók pedig nem láthatják őket.

A lexikográfiai, lexikológiai munkák pontosítására, valamint a határon túli magyar nyelvvaltozatok elemeinek behatóbb vizsgálatára különféle paraméterekkel rendelkező listák is szolgálnak. A kvantitatív kutatások adatigényét kiszolgáló listák a szójegyzékekhez hasonlóan igény szerint alakíthatók, segítve vagy adatokkal ellátva a szerkesztőket.

A különféle nézetek adta lehetőségnek köszönhetően az olvasó vagy szerkesztő csak azokat az információkat látja, amelyekhez hozzá szeretne férni, amelyeket látni szeretne. A szócikkek olvasását a rövidítéseket feloldó tooltip (helyi súgó)⁵ is segíti.

A Termini magyar–magyar szótár és adatbázis a maga nemében sajátos vállalkozás a magyar lexikográfiában. Sajátos a lefedett nyelvvaltozatok tekintetében, sajátos a mikrostruktúra kidolgozottságában, illetve sajátos a készítés folyamata alapján is, mivel olyan, nyelvész szakemberek által szerkesztett adatbázis-alapú szótárról

⁴ <http://termini.nyud.hu/htonline/htlista.php>

⁵ A rövidítések felett megjelenő, a rövidítések feloldását tartalmazó kis szövegdoboz.

van szó, amely elsődlegesen az online megjelenésre van paraméterezve (lásd például a példamondatok számát vagy a szócikkekhez kapcsolódó fényképeket).

A Termini-szótár az egyetlen olyan komplex szótár, amely a magyar nyelv határon túli változatainak sajátos lexikális elemeit, jelentéseit tartalmazza – mindezt meglehetősen finomhangolt lexikográfiai és lexikológiai apparátussal, valamint infokommunikációs technológiával.

5. A szótár multimédiás elemei

Napjainkban a digitális szótárak egyre növekvő népszerűsége vitathatatlan. Legnagyobb előnyük az információhoz történő könnyű és gyors hozzáférés, amely mindennapi életünkben bárhol, bármikor lehetséges. Az elektronikus tárolás ezen túl a multimédiás eszközök használatának széles skáláját is lehetővé teszi, így a szótárhasználók nem kizárólag a papírszótárban megszokott, „hagyományos” szócikkeket olvashatnak, hanem lehetőség nyílik vizuális és akusztikus megoldások alkalmazására is. A kép jelentésleírást támogató elemként történő megjelenítése a szótártörténetben már hagyományosnak tekinthető, gondoljunk csak a képes szótárakra. Ha csak a magyar szótárak történetében tekintünk vissza, akkor több száz évvel korábban, már 1653-ban megtalálhatjuk Comenius képes „szótárát”, amely a vizuális lehetőségek alkalmazásának nagyszerű előfutára (P. Márkus 2019), valamint megemlíthetjük még az az Értelmező kéziszótár első kiadását is, amely egyes címszavakat a definíción túl képek segítségével magyarázott. A képek a szótárban enciklopédikus tartalomnak minősülnek: nem a nyelvi jelet, hanem a világot láttatják az olvasóval, segítségükkel a jelentések pontosíthatók, egyfajta megerősítést adnak az olvasónak, valamint a szöveges definícióknál sokkal több, a címszóval kapcsolatos tipikus jellemzőt tudnak közvetíteni (Svensén 2009).

A Termini-szótárban képeket – amelyek a szócikkekben az egyes jelentésekhez külön-külön is beilleszthetők – legtöbbször főnevek, konkrét tárgyak megjelenítésére alkalmazunk. Az illusztrációk alatt a képaláírás és a forrás is olvasható. A képaláírások nem feltétlenül csak megismélik a címszót, hanem gyakran leíró jellegűek, és egyfajta magyarázatként is szolgálnak. Ez láthatjuk a *blokk* szócikk Er, Fv, Va, Ka, Hv, Mv, Öv '(lakótelepi) panelház' jelentésénél, ahol a következő képaláírás áll: „Jellegetes szocializmus korabeli blokk Szereden”.

Hupka (1989) felosztása alapján a vizuális elemek többféle típusba sorolhatók. A Termini-szótár három fő típust alkalmaz ezek közül. Az első csoportba azon képek tartoznak, amelyek kizárólag a tárgy jelenik meg. Ezekben az esetekben igyekezni kell a legtipikusabb elem(ek) kiválasztására. Jó példa erre a típusra a Fv, Ka *baklászán* 'padlizsán' címszó illusztrációja, hiszen itt a zöldség formájának, alakjának és színének a megjelenése elegendő az értelmezéshez; de az élet egyéb területein is találunk hétköznapi tárgyakat: Er, Fv, Va, Ka, Hv, Öv *kalendár* 'naptár'; Fv *bandaszka* 'benzin, olaj szállítására használt marmonkanna; víz szállítására használt műanyag kanna'; Fv, Va, Ka, Hv, Mv, Öv *batri* 'szárazelem', ahol hasonlóan nem volt szükség tágabb környezetre a jelentés bemutatásához. A szótárban találunk még márkanéveket is, amelyek szintén ide sorolhatók, hiszen a vizuális elemek márkához

kapcsolódó termékeket ábrázolnak, így a képek a termékeket egyszerűen be tudják mutatni (pl. Fv *ticsinki* 'ropi'; Fv *chrumki* 'mogyorós pufi', ahol a termékek csomagolása, illetve kinézete önmagáért beszél).

A következő csoportban a tárgyat már környezetében kell bemutatni, és ez a típus rengeteg plusz információval gazdagíthatja az olvasót. Megemlíthetjük például a Fv *bábovka* 'kuglóf' szócikket, ahol a sütemény ízlésesen tálalva jelenik meg az olvasó előtt; a Fv *kopacski* 'stoplis futballcipő' képén már a futball-labda és a futballpálya is látható, tehát a tágabb környezet segít az értelmezésben; a Fv *dodávka* ' furgon' esetében a járművet egy raktárbejárat előtt, a kocsibeállón láthatjuk, illusztrálva ezzel a haszongépjármű-jellegét.

A harmadik csoportba tartozó képek a tárgyat használat közben mutatják be, hiszen az értelmezést itt a működés, a funkció bemutatása illusztrálja a legjobban. Szemléltetésül megnézhetjük a Fv *kolobeska* 'roller' képén megjelenő rollerező fiataalt vagy a Fv *bublifik* 'szappanbuborék-fújó gyermekjáték' alatt szereplő gyereket, aki éppen buborékot fúj, de a Fv, Va *bazén* 'medence' képe is pontosan a használatot és a tárgy funkcióját illusztrálja.

Esetenként melléknevekhez, igékhez is lehet képet csatolni, ilyenek például az Fv, Öv *autoritativ* 'tekintélyelvű, tekintélyt parancsoló'; Fv, Va, Ka, Hv, Mv, Öv *degusztál* 'ételt, italt (meg)kóstol' szócikkek, ahol a melléknevet mint emberi tulajdonságot egy kifejező mozdulattal, valamint szemléletes környezettel lehet ábrázolni; az ige pedig a cselekvés végzése közben mutatható be.

A Termini-szótár tipológiai sajátosságai miatt nem alkalmaz egyéb, a Hupka-féle felosztás szerinti képtípusokat, mint például a fölrendelt fogalom (hiperonima) szócikkében megjelenő alárendelt fogalmak (hiponimák) illusztrációi (pl. a sport szócikkben található kép tartalmazza a különböző sportokat: kosárlabda, tollaslabda, tenisz stb.). Az említett típus főként a tanulószótárakban gyakori, hiszen ezek a szótanulást nagymértékben támogatják. Szintén nincs szükség olyan folyamatot szemléltető képekre, amelyek inkább az enciklopédiákban megszokottak, és olyan változásokat tudnak bemutatni, mint például a rovarok fejlődésének a szakaszai.

A vizuális tartalom mellett a szótárban hanganyagok is megjelennek, hiszen a szócikkek egyes példamondatai beszélt nyelvi használatot tükröznek. A hangfelvételek régiónkénti önálló gyűjtésből származnak, amelyekből a szerkesztők kiválasztják az adott címszót legjobban szemléltető szövegrészletet, amely írott formában példamondatként olvasható, és ehhez kapcsolódóan hanganyagként meghallgatható. A szótár viszonylag kevés hanganyagot tartalmaz, mivel ezen terület új fejlesztésnek számít, és egyelőre kísérleti fázisban van. A jövőben így a hanganyagok segítségével sokoldalúan bemutatható a régiónkénti autentikus beszélt nyelvhasználat is.

6. A szótár stílusminősítési rendszere

Az új stílusminősítési rendszer kidolgozásában a szerkesztők a Magyar értelmező kéziszótár minősítéseiből indultak ki (ÉKsz. 1972 és ÉKsz.² 2003), arra törekedve, hogy ok nélkül ne térjenek el tőlük. Az egyes szókészleti elemek, illetve jelentések stílusértékét különböző, szociolingvisztikailag meghatározható kategóriák, úgyne-

vezett dimenziók mentén állapítják meg. A többdimenziós megközelítés fő előnye, hogy – szemben a hagyományos minősítési rendszerrel – a minősítés különféle szempontjai nem keverednek,⁶ ezáltal pedig lehetővé válik, hogy a minősítés során minden lehetséges szempont külön-külön legyen mérlegelve.

- (1) A dialektális dimenzió a földrajzi és a társadalmi dialektusokra vonatkozó minősítéseket tartalmazza. Ezek a minősítések a lexémák abból adódó stílusértékére utalnak, hogy használatuk bizonyos társadalmi csoportokra jellemző (míg más földrajzi és társadalmi csoportokra nem vagy nem annyira). A földrajzi dialektusokra a „tájnyelvi” minősítés utal (pl. Er *csimpolya* ’fűrulya’). A társadalmi dialektusok közül az iskolázottságin az „iskolázott” (pl. Öv *förderung* ’támogatás’) és a „népnyelvi” (pl. Fv, Ka *horcsica* ’mustár’) minősítés található, a kétnyelvűség az „idegenes” (pl. Ka *kiesik a hó* ’leesik a hó’) minősítés, a korosztályin pedig a „gyermeknyelvi” (Ka *pupu* ’köldök’) és az „ifjúsági nyelvi” (pl. Mv *dongó* ’erősebb testalkatú, kövér nő’) minősítés. Gyakorlati okokból semleges kategória csak az iskolázottsági dimenzió van megkülönböztetve: ez az „általános” minősítés (pl. Fv *alobal* ’alufólia’), amely a legtöbb vagy minden földrajzi és társadalmi csoportban elterjedt szavakat, illetve szójelentéseket jelöli.
- (2) A regiszterbeli dimenzió mentén javarészt a bizonyos specializált tevékenységekhez kötődő stílusértékek helyezkednek el. Ide tartozik a „szaknyelvi” (pl. Fv, Va, Hv, Mv *ekológia* ’környezettan, ökológia’), a „közéleti” (pl. Er *félnorma* ’félmunkaidő’), a „sajtónyelvi” (pl. Ka *üzembe állít* ’üzembe helyez’), a „szépirodalmi” (Hu *arranzsál* ’rendez, elrendez’) és a „szleng” (pl. Fv *humusz* ’kupi’) minősítés. Ha egy szó – a látszat ellenére – nem szaknyelvi, a „hétköznapi” minősítést kapja (pl. Öv *slipsz* ’slicc’). Arra a tényre, hogy egy szó(jelentés) nem kötődik semmilyen specializált regiszterhez, a „közhasználatú” (pl. Fv, Va, Öv *exponát* ’kiállítási tárgy’) minősítés utal.⁷
- (3) A stílusváltozati dimenziót a beszédhelyzet formalitása által meghatározott regiszterek alkotják. Az egyes szavak, illetve jelentések stílusértékét itt az határozza meg, hogy formális vagy informális beszédhelyzetekben használják-e őket jellemző módon. Ide tartozik a „választékos” stílusminősítés (pl.

⁶ A hagyományos stílusminősítési rendszer legnagyobb problémája a minősítések egy részének többértelműsége (amit nyilván a szótár közhasználatúságának igénye magyaráz). Például a „választékos” és a „bizalmas” minősítések nemcsak a lexémák beszédhelyzeti kötöttségére utalhatnak, hanem a beszélők társadalmi háttérére, sőt a lexémák expresszív stílusértékére is. A Termini-szótár stílusminősítési rendszerének hátránya viszont, hogy részletessége miatt a laikus felhasználó számára nehezen követhető. A tapasztalatok azonban azt mutatják, hogy a stílusminősítéseket a laikus felhasználók más szótárakban is javarészt figyelmen kívül hagyják, így azt mondhatjuk – kissé leegyszerűsítve –, hogy a szótárban alkalmazott stílusminősítések inkább a szakembereknek szólnak.

⁷ A „szleng” és a „hétköznapi” minősítés kilóg a többi közül annyiban, hogy nem feltétlenül kötődik specializált tevékenységhez. De akár kötődhetnek is, hiszen a szleng jellegzetes válfaja a szakmai szleng, a szakemberek pedig nem „hivatalos” *terminus technicus* jellegű szakszavakat is használnak.

Fv *sikanéria* 'hadseregben, kollégiumban fiatalabbak, gyengébbek zaklatása') és annak két szélső változata: egyfelől a „kissé választékos” (pl. Fv *borókafenyő* 'karácsonyfának való erdeifenyő'), másfelől az „emelkedett” (Ka *dicsérettábla* 'dicsőségtábla'), továbbá a „bizalmas” stílusminősítés (Er *pilája van* 'protekcója van') és annak két szélső változata, a „kissé bizalmas” (Va *fazon* 'poén, csattanó'), valamint a „lezser” (Hv *event* 'esemény'). Azok a szavak, amelyek nem kötődnek egy-egy stílusváltozathoz, hanem bármilyen beszédhelyzetben használhatók, a „közömbös” (pl. Er *bevásárlószekér* 'bevásárlókocsi') minősítést kapják.

A rendszerezés egyik előnye, hogy e dimenzió minősítései valóban *tisztán* stílusválogatási minősítések, közvetve sem utalnak a beszélők társadalmi háttéréből adódó stílusértékre. Ezért lehet gond nélkül jelölni például az iskolázatlan beszélők által használt „választékos” stílusértékű, nemegyszer hiperkorrekt indíttatású szavakat („népnyelvi” és „választékos”, pl. Fv, Va, Hv, Mv *garancia* 'garancia') vagy az iskolázott beszélők által használt „bizalmas” stílusértékű szavakat („iskolázott” és „bizalmas”, pl. Fv, Va *szkripta* 'sokszorosított egyetemi, főiskolai jegyzet”).

- (4) Az időbeliség dimenzióján jellemző módon a nyelv korábbi állapotából fennmaradt elemeknek a mai nyelvben meglévő sajátos stílusértéke fejeződik ki. E dimenzió mentén helyezkednek el a „régies” (pl. Er *bordély* 'viskó, kunyhó') és az annak enyhébb fokozatát képviselő „kissé régies” (pl. Mv *modli* 'süteményszaggató'), az „elavult” (Er *élelemporció* 'élelemadag'), a „történeti” (pl. Fv *járasi nemzeti bizottság* 'járasi szintű igazgatási szervezet a rendszerváltás előtt') és az „új” (pl. Mv *hitriteszt* 'gyorsteszt') minősítések. Mivel csak viszonylag kevés lexémának van időbeliségre vonatkozó stílusértéke, az egyszerűség kedvéért az időbeliség síkján sajátos stílusértéket nem hordozó szavaknak nincs speciális jelölése.
- (5) A dialektusokhoz, regiszterekhez és kronolektusokhoz kötődő sajátos stílusminősítések mellett a rendszerezésben „igazi” stílusminősítések is szerepelnek, amelyekben a beszélők érzelmei fejeződnek ki, illetve a szavak által jelölt valósághoz való pozitív vagy negatív viszonyulásuk. Ezen az érzelmi dimenzió helyezzük el az „eufemisztikus” (Va *malverzáció* 'hütlén kezelés'), a „kedveskedő” (Öv *basztlizik* 'barkácsol'), a „tréfás” (pl. Va *visak* 'többletsúly'), a „pejoratív” (Er *spága* 'csúszópéncz'), illetve ennek enyhébb fokozata, a „kissé pejoratív” (Hv *siptár* 'albán'), az „ironikus” (Va *hegyi tango* 'popzenével vegyített szerb népi zene') és ennek enyhébb válfaja, a „kissé ironikus” (Öv *wurst* 'mindegy'), a „durva” (Fv *jebák* 'pattanás') és a „kissé durva” (Mv *frájer* 'szoknyapeccér, kurvahajcsár'), valamint a „trágár” (–).

Az érzelmi dimenzióhoz sorolható a lexémából magából, annak hangalakjából és/vagy jelentéséből adódó sajátos stílusérték is: a hangalakjuk miatt expresszív szavak sajátos stílusértékére az „expresszív” (pl. Fv *babracska* 'babramunka'), a jelentésük miatt expresszív szavak (metaforák) sajátos stílusértékére pedig az „átvitt” (Fv *gigant* 'óriás') minősítés vonatkozik.

Az „expresszív” minősítés utal az érzelmileg telített, de a fentebbi kategóriákba nehezen besorolható szavakra is (pl. Er *bába* ’öregasszony’).

- (6) A teljesség kedvéért meg kell említeni, hogy a fenti dimenziókon túl lehetőség van a lexémák gyakoriságának jelölésére a „ritka” (pl. Er *braszka* ’béka’) minősítéssel, ezenkívül rendszeresen jelezve van a szavak és szókapcsolatok régióbeli hovatartozása az Er, Fv, Va, Ka, Hv, Mv, Öv, továbbá a Hu (Magyarország) és Km (Kárpát-medence, illetve közmagyar) felségjelzésekkel. Ezek azonban rendszeresen nem forrásai az érintett lexémák sajátos stílusértékének, kivéve olyankor, amikor nyelven belüli kódváltásként más állami nyelvváltozatba kerülnek be.⁸

A Termini-szótár stílusminősítési rendszerében látszólag sok az új kategória, ám figyelembe kell venni, hogy az újnak látszó minősítések jelentős része lappangva eddig is megvolt szótárainkban. Például a minősítetlenség utalt a szó „általános”, „közhasználatú” és „közömbös” stílusértékére, az „argó” az ifjúsági nyelvire, nemegyszer a „választékos” az „iskolázott”-ra, a különféle fogalomkörü minősítések és részben a „hivatalos” minősítés a „szaknyelvi”-re. A cél soha nem a minősítések számának öncélú növelése, hanem a minősítések egyértelműsítése volt.

7. Összegzés

A *Termini magyar–magyar szótár és adatbázis* sok tekintetben egyedülálló a magyar lexikográfia területén. Erre utal többek között a szócikkek sajátos szerkezete, összetettsége, a szócikktörzs információgazdagsága, a szövegkontextusban idézett példamondatok eredetének jelzése. A szótár relációs adatbázisként lehetőséget biztosít többek között statisztikák készítésére és adatbányászatra, ugyanakkor biztosítja a klasszikus szótár lineáris olvasását is. A technikai háttér, az online jelleg lehetővé teszi a szócikkek helytől független szerkesztését. Interaktív szótárként a felhasználó beállíthatja, hogy a lekérdezés alapján milyen struktúrájú szócikket kapjon, és a regisztrált felhasználók is megjegyzéseket fűzhetnek az egyes szócikkekhez. A lexikográfiai adattár sajátos a megjelenített nyelvváltozatok tekintetében, sajátos a mikrostruktúra kidolgozottságában, illetve sajátos a készítés folyamata tekintetében is, mivel olyan, nyelvész szakemberek által szerkesztett adatbázis-alapú szótárról van szó, amely elsődlegesen az online megjelenésre van paraméterezve.

A szótár multimédiás jellege tovább erősíti egyedülállóságát. Ezt a szócikkeket kísérő kép-, és legújabb fejlesztésként, hanganyag teszi lehetővé.

A szótár által alkalmazott új stílusminősítési rendszer szociolingvisztikailag meghatározott kategóriákon alapul: dialektális, regiszterbeli, stílusváltozati, időbeliségi, érzelmi viszonyulásra utaló dimenziót különböztet meg, kizárva a hagyományos minősítési rendszerben előforduló minősítési szempontok keveredését.

A szótár bemutatott sajátosságai egyben az adatok kutatási célra való felhasználhatóságát is jelzik.

⁸ A Termini-szótár stílusminősítési rendszerére lásd még Lanstyák 2009a, 2009b.

SZAKIRODALOM

- ÉKsz. 1972. *Magyar értelmező kéziszótár*. Szerk. Juhász József – Szőke István – O. Nagy Gábor – Kovalovszky Miklós. Akadémiai Kiadó, Budapest.
- ÉKSz.² 2003. *Magyar értelmező kéziszótár*. Pusztai Ferenc (főszerk.). Akadémiai Kiadó, Budapest (2., átdolgozott kiadás.)
- Benő Attila 2014. *Kontaktusjelenségek az erdélyi magyar nyelvváltozatokban*. Erdélyi Múzeum-Egyesület, Kolozsvár. <https://doi.org/10.36240/etf-281>
- Benő Attila 2015. Többirányú lexikai átvételek a mai magyar nyelvváltozatokban. In: Benő Attila – Fazakas Emese – Zsemlyei Borbála (szerk.): *Többnyelvűség és kommunikáció Kelet-Közép-Európában*. Erdélyi Múzeum-Egyesület, Kolozsvár, 248–56.
- Benő Attila 2017. Termini – az internetes magyar–magyar szótár és adatbázis. In: Fábíán Zsuzsanna (szerk.): *Szótárírás a Kárpát-medencében. A magyar és a szomszédos országok többségi nyelveinek kétnyelvű szótárjai*. Tinta Könyvkiadó, Budapest, 209–27.
- Benő Attila 2019. Kontaktusjelenségek vizsgálata a külső régiók magyar nyelvváltozataiban és a Termini magyar–magyar szótár. In: Pletl Rita – Kovács Gabriella (szerk.): *Soknyelvűség és többnyelvűség Európában*. Scientia Kiadó, Kolozsvár, 15–26.
- Benő Attila – Lanstyák István 2019. Megőrzött régiség és kontaktushatás. In: É. Kiss Katalin – Hegedűs Attila – Pintér Lilla (szerk.): *Nyelvelmélet és kontaktológia 4*. Pázmány Péter Katolikus Egyetem, Bölcsészettudományi Kar, Piliscsaba, 13–26.
- Benő Attila – Lanstyák István – Juhász Tihamér 2020. A Termini „határtalan” szótára. *Magyar Tudomány 2*: 153–63. <https://doi.org/10.1556/2065.181.2020.2.2>
- Hupka, Werner 1989. *Wort und Bild: Die Illustrationen in Wörterbüchern und Enzyklopädien. Lexicographica*. Series Maior 22. Max Niemeyer Verlag, Tübingen. <https://doi.org/10.1515/9783111347776>
- Kolláth Anna 2011. Idegen szók és gyakoriságuk a muravidéki magyar nyelvváltozatokban. Kozmács István – Vančonek Kremmer Ildikó (szerk.): *A csitári hegyek alatt... Írások Sándor Anna tiszteletére*. Europica varietas. Arany A. László Társulás – Konstantin Filozófus Egyetem, Közép-Európai Tanulmányok Kara, Nyitra, 123–31.
- Kolláth Anna 2012. *A szlovéniai magyar nyelv a többnyelvűség kontextusában*. Mednarodna založba Oddelka za slovenske jezike in književnosti, Filozofska fakulteta. Mednarodna knjižna zbirka Zora, Maribor.
- Lanstyák István 2009a. A sajátos határon túli magyar szókincs változatosságának dialektális, regiszterbeli, időbeli és stilisztikai aspektusáról. In: Borbély Anna – Vančonek Kremmer Ildikó – Hattyár Helga (szerk.): *Nyelvideológiák, attitűdök és sztereotípiák*. Budapest – Dunaszerdahely – Nyitra, MTA Nyelvtudományi Intézet – Gramma Nyelvi Iroda – Konstantin Filozófus Egyetem Közép-európai Tanulmányok Kara, Tinta Könyvkiadó. 17–24. – Újraközlés: Benő Attila – Péntek János (szerk.): *A Termini Magyar Nyelvi Kutatóhálózat tíz éve. Tanulmányok, beszámolók, kutatási programok*. Gramma Nyelvi Iroda – Szabó T. Attila Nyelvi Intézet, Dunaszerdahely–Kolozsvár, 331–9.
- Lanstyák István 2009b. A platni botránya. (Egy új stílusminősítési rendszer felé.) *Fórum Társadalomtudományi Szemle 11/2*: 25–40.
- Lanstyák István 2009c. Termini magyar–magyar szótár – minden Kárpát-medencei magyar szótára. In: Lanstyák István – Menyhárt József – Szabó Mihály Gizella (szerk.): *Tanulmányok a kétnyelvűségről V*. Gramma, Dunaszerdahely, 208–18.
- Lanstyák István – Benő Attila – Juhász Tihamér 2010. A Termini magyar–magyar szótár és adatbázis. *Régió 4*: 37–58.

- Molnár Csikós László 2015. A vajdasági magyar regionális köznyelv. In: Gasparics Judit – Ruda Gábor (szerk.): *Adalékok a délvidéki magyarok nyelvéhez és kultúrájához*. Muravidék Baráti Kör Kulturális Egyesület, Pilisvörösvár, 68–74.
- Molnár Csikós László 2017. Regionális idegen szavak a vajdasági magyar nyelvhasználatban. In: Gróf Annamária – Szoták Szilvia (szerk.): *Szélrózsa. A Termini Magyar Nyelvi Kutatóhálózat 2014. évi lendvai konferenciájának anyaga*. Termini Egyesület, Budapest, 61–71.
- P. Márkus Katalin 2019. Szótárak: Kezdetektől napjainkig. In: Furkó Péter – Szathmári Éva (szerk.): *Népszerű tudomány, tudománynépszerűsítés: A Károli Gáspár Református Egyetem 2019-es évkönyve*. Károli Gáspár Református Egyetem, Budapest, 146–55.
- M. Pintér Tibor 2017a. Számszerűsített kétnyelvűség. Adatbányászati technikák használat a nyelvészetben. In: Misad Katalin (szerk.): *Kétnyelvűség térben és időben*. Szenczi Molnár Albert Egyesület, Pozsony, 125–57.
- M. Pintér Tibor 2017b. Adatbányászat a kétnyelvűség mögött. In: Benő Attila – Gúti Erika – Juhász Dezső – Szoták Szilvia – Terbe Erika – Trócsányi András (szerk.): *Tudományköziség és magyarságtudomány a nyelvi dimenziók tükrében. A VIII. Nemzetközi Hungarológiai Kongresszus (Pécs, 2016) három szimpóziumának előadásai*. Termini Egyesület, Budapest, 158–67.
- M. Pintér Tibor 2019a. Magyar nyelv a határon. A Termini magyar–magyar szótár. 2. rész. *Magyar Nyelv* 115/4: 473–9. <https://doi.org/10.18349/MagyarNyelv.2019.4.473>
- M. Pintér Tibor 2019b. Magyar nyelv a határon. A Termini magyar–magyar szótár. 1. rész. *Magyar Nyelv* 115/3: 333–47. <https://doi.org/10.18349/MagyarNyelv.2019.3.333>
- Svensén, Bo 2009. *A Handbook of Lexicography: The Theory and Practice of Dictionary-Making*. Cambridge University Press, New York.
- Szoták Szilvia 2018. A Termini Magyar Nyelvi Kutatóhálózat terminológiai tevékenysége – A Termini magyar–magyar szótár jogi vonatkozásai. *Magyar Jogi Nyelv*, 25–9.
- Szoták Szilvia 2019. Értjük egymást? Oktatásterminológiai különbségek a magyar–magyar kommunikációban. In: Szoták Szilvia (szerk.): *Diszciplínák találkozása – nyelvi közvetítés a 21. században*. OFFI Zrt., Budapest, 247–59.

Benő Attila
egyetemi tanár

Babeş-Bolyai Tudományegyetem
Magyar és Általános Nyelvészeti Tanszék
<https://orcid.org/0000-0002-6931-2999>

Lanstyák István
egyetemi tanár

Comenius Egyetem
Magyar Nyelv és Irodalom Tanszék
<https://orcid.org/0000-0002-8237-9258>

P. Márkus Katalin
egyetemi docens

Károli Gáspár Református Egyetem
Angol Nyelvészeti Tanszék
<https://orcid.org/0000-0001-6826-7631>

M. Pintér Tibor
egyetemi docens

Károli Gáspár Református Egyetem
Magyar Nyelvtudományi Tanszék
<https://orcid.org/0000-0002-5212-4107>

SUMMARY

Benő, Attila; Lanstyák, István; P. Márkus, Katalin; M. Pintér, Tibor

Structural and editorial features of the Termini Hungarian–Hungarian dictionary and database

The *Termini Online Hungarian–Hungarian Dictionary and Database* describes the lexicon of the Hungarian language as spoken in the countries surrounding Hungary. It is considered to be a general dictionary of Present-day Hungarian. Each entry contains authentic sample sentences to illustrate the use of the headword, making it possible to examine the special use of a word or construction in a grammatical and pragmatic context. The lexicographical database is edited online in eight countries. The editors of the dictionary are members of the *Termini Hungarian Language Research Network*. Online editing makes it possible for the dictionary to expand – even simultaneously – as a result of activity in eight countries. In the present study, the authors review the novelties and peculiarities of the dictionary in some detail, touching on the following topics: dictionary structure, IT support, database character, multimedia elements, and labelling system.

Keywords: online dictionary, lexicographical database, Hungarian language variants, language contacts, style, etymology

Az oktatás terminológiája a Termini-szótárban – délszláv nyelvi kontaktusban**1. Bevezetés**

A vajdasági (Va), a drávaszögi (Hv) és a muravidéki (Mv)¹ magyar közösség a 20. század kilencvenes éveinek elejéig, Jugoszlávia széthullásáig egy országban élt, a közvetlen, fizikai érintkezés különösen a drávaszögi és a vajdasági közösség esetében volt mindennapos, de a három régió tekintetében is folyamatos kapcsolattartás volt jellemző, sorsuk alakulásában ez a tényező is fontos szerepet játszott. Jugoszlávia megalakulásától a horvát és a szerb nyelvi standard egyfajta keverékeként a szerbhorvát, illetve horvátszerb vált az állam hivatalos nyelvvé. Ahogyan a többi nemzetiség esetében, úgy a magyar anyanyelvű lakosság vonatkozásában is jellemzőek lettek/voltak a nyelvi kontaktusjelenségek. A Jugoszláviában biztosított kisebbségi nyelvi jogok lehetővé tették a kisebbségi nyelven működő oktatási rendszer kialakítását is, ugyanakkor az oktatásszervezési folyamatok – objektív okok miatt – nem követték a Magyarországon aktuális oktatási terminológiát. Így egyrészt sok régies kifejezés is megmaradt az oktatással kapcsolatos magyar nyelvhasználatában, másrészt a kialakult kétnyelvűség jegyei is nyomon követhetők az oktatási terminológiában. A fontos szociológiai tényezőként számontartott nyelvi jogok oktatásszervezési

¹ A tanulmányban a Va a vajdasági, a Hv a horvátországi, az Mv pedig a muravidéki korpuszt (annak lexikai egységeit) jelöli ugyanúgy, ahogy teszi a Termini-szótár is.