

ETTÉL MÁR SÜLT HERMELINT HRANOLKIVAL? A Termini-szótárról dióhéjban¹

M. PINTÉR TIBOR, P. MÁRKUS KATALIN

Károli Gáspár Református Egyetem, Bölcsészet- és Társadalomtudományi Kar

m.pinter.tibor@kre.hu | p.markus.kata@kre.hu | DOI: 10.18135/PG70.2024.10

Prószék Gáborral először egy majdnem tízórás vonatút során találkoztam. Szepsiszentgyörgyre utaztunk (talán még emlékszik is rá), ahol előadást tartott. Később, egyik első munkahelyem épp a MorphoLogic Kft. lett (ti. MPT-nek), amely inspiráló volt – a mai napig szeretettel gondolok arra az időszakra és Gábor mindenkori közvetlenségére, vidámságára, világlátására. Lexikológiai, lexikográfiai előadásaiban mindig kitért arra, hogy a kis nyelvek szótárait online kellene működtetni (ehhez képest a MorphoLogic Kft. végezte el angol–magyar kétnyelvű szótárain XML-esítését és online térbe helyezését – PMK). Hát, az évek múlnak, mi pedig épp egy online szótár munkálataiban is részt veszünk. Egy szótáréban, amelyet épp a Nyelvtudományi Kutatóközpont hoztolt. Álljon itt egy rövid értekezés arról, hogy is gondoljuk a magyar–magyar online szótárat – bizonyosságként a magyar nyelv többközpontúságára.

Bevezetés

A Termini magyar–magyar szótár és adatbázis (termini.nytud.hu/htonline/) a külső régiók magyar nyelvhasználatát dokumentálja, így egyedülálló szótártípusnak számít a hazai szótárpalettán. A magyar nyelv leírása szempontjából különösen fontos, hiszen a szótárba azok a magyarországiaktól eltérő, idegen eredetű szavak kerülnek, amelyeket a Kárpát-medencében, hazánk határain túl élő magyarok használnak. A szótár hét, Magyarországot körülvevő ország magyarlakta régióinak szavait tartalmazza: Erdély, Felvidék, Vajdaság, Kárpátalja, Horvátország, Muravidék, Órvidék. A gyűjtemény 2007 óta online is elérhető. A szótár szerkesztőinek elsődleges célja, hogy lehetővé tegyék más magyar nyelvváltozatok sajátos szavainak és szójelentéseinek (pl. *hermelin*, *hranolki*, *kofola*, *nanuk*) megismerését, másrészt, hogy az említett régiók beszélői is megismerhessék e szavak közmagyar megfelelőit (ha vannak ilyenek). A Termini-szótár „nyitott” szótárként működik, a közösség is beleszólhat a szerkesztésébe: ha az olvasók új szócikket, esetleg meglévő szócikkhez új jelentést szeretnének javasolni, írhatnak a szerkesztőknek. Mindegyik címszó használatának szemléltetésére lehetőség szerint élőnyelvi példamondatot, szövegrészletet tartalmaz (a forrás pontos megjelölésével), lehetővé téve, hogy a szó vagy szerkezet használati sajátosságait grammatikai és pragmatikai kontextusban vizsgáljuk. Tanulmá-

¹ Jelen tanulmány alapjául az alábbi írás szolgált: Benő et al., 2021.

nyunkban a szótár újszerűségeit, valamint érdekességeit szeretnénk részletesebben áttekinteni.

A Termini-szótár és adatbázis a maga nemében sajátos vállalkozás a magyar lexikográfiában. Sajátos a lefedett 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.

Informatikai háttér

A szótár kizárólag online felületen létezik, egy időben nyolc országból szerkesztik. A változtatások azonnal látszanak a találati listában, amely értelmezhető akár egy papírszótár oldalainak is. A többfelhasználós platform alapja egy relációs adatbázis, melyhez a bemeneti (szerkesztőfelület) és kimeneti (lekérdezőfelület) oldal tartozik. Ez a technológiai megoldás, amelyet a mai lexikográfiában 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, illetve az online megjelenés folyamatos lehetőséget biztosít képek és hangfájlok közzétételéhez és azok folyamatos revíziójához. A szótár mögötti adatbázis nyelvészeti, elsősorban lexikográfiai szempontból rendkívül gazdag szerkezetű – igyekszik minél árnyaltabb képet adni a címszó mögötti lexikális és konceptuális tartalmakról. Az egyes lekérdezések folyamán lesz adatból információ, amely lexikográfiai, lexikológiai vagy adatbányászati szempontból is érdekes.

A technológia meghatározza a szótár szerkesztését: a papíralapú szótárak lineáris 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), azonban a szerkesztés folyamán használt felület (nézet) 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.

Interaktív szótár

A szótár nem pusztán lexikai gyűjtemény, hanem a nyelvhasználat, nyelvváltozatok, lexikai és morfológiai adatbázisa – adatbányászatra, valamint a tárolt tartalom különféle formában, különböző rendezési elvek 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 honlap nyilvános felülete – felhasználóbarát módon – csupán az alapfunkciókat felkináló *Keresés* mezőt tartalmazza. A szótár elsősorban nyelvészek által használt része a *Szójegyzékek* menü, mely *lexikográfiai adatbányászatként* működik. A kereső/lekérdező a szócikkstruktúra bármelyik elemében kereshet.

A *szójegyzékek generálása* menüben többféle szójegyzék is előállítható, melyek létrehozását négy alaptípusba sorolt kategóriaegyüttes teszi lehetővé: így a szerkesztés időszakára vonatkozó szócikkek (*Időszak*), a szótárral foglalkozó nyelvészek lexikográfiai kutatásaihoz szükséges szócikkek (*Kutatási listák*), a szótári mikrostruktúra bizonyos részeit érintő és a lexikográfiai kutatásokat segítő szócikkek (*Fejlesztési listák*), valamint a fejlesztők munkásságát vagy az aktuális állapotát tükröző listák (*Dokumentációs listák*) előállítása egyaránt lehetséges. A listák és 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 „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 nyelvváltozatok elemeinek behatóbb vizsgálatára különféle paraméterekkel rendelkező listák is szolgálnak. A kvantitatív kutatások adatigényét kiszolgáló listák a szójegyzékekhez hasonlóan igény szerint alakíthatóak, 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.

A Termini-szótár az egyetlen olyan komplex szótár, mely speciálisan a magyar nyelv határon túli változatainak sajátos lexikális elemeit, jelentéseit tartalmazza.

A szótár multimédiás elemei

Napjainkban a digitális szótárak egyre növekvő népszerűsége vitathatatlan. Legnagyobb előnyük a könnyű és gyors információhozzáférés, 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 képjelentésleírást támogató elemként történő megjelenítése a szótártörténetben már hagyományosnak tekinthető, gondoljunk csak a képes szótárakra. Ha csak a magyar szótárak történetben tekintünk vissza, akkor több száz évvel korábban, már 1653-ban megtalálhatjuk Comenius képes „szótárát”, amely a vizuális lehetőségek alkalmazásának előfutára, valamint megemlíthetjük még az Értelmező kézisztár 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.

A Termini-szótár képei, amelyek a szócikkekben az egyes jelentésekhez külön-külön is beilleszthetők, legtöbbször főneveket, konkrét tárgyakat jelölnek. Az illusztrációk alatt a forrás és a képaláírás is olvasható, amelyek leíró jellegük miatt egyfajta magyarázatként is szolgálnak. Ezt láthatjuk a *blokk* szócikk '(lakótelepi) panelház' jelentésénél, ahol a következő képaláírás áll: „Jellegzetes 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, amelyekben kizárólag a tárgy jelenik meg: igyekezni kell a legtipikusabb elem(ek)et kiválasztani. Jó példa erre a típusra a Fv, Ka *baklazsán* 'padlizsán' címszó illusztrációja, hiszen itt a zöltség formájának, alakjának és színének a megjelenése elegendő az értelmezéshez; de az élet egyéb területein is találunk hétköznapi tárgyakat: *kalendár* 'naptár', *bandaszka* 'benzin, olaj szállítására használt marmonkanna; víz szállítására használt műanyag kanna', *batri* 'szárazelem', ahol hasonlóan nem volt szükség tágabb környezetre a jelentés bemutatásához. A szótárban találunk még márkanevet is, amelyek szintén ide sorolhatók, hiszen a vizuális elemek márkához kapcsolódó termékeket ábrázolnak, így a képek a termékeket egyszerűen be tudják mutatni (pl. *ticsinki* 'ropi', *chrumki* 'mogyorós pufi', ahol a termékek csomagolása vagy kinézete önmagáért beszél).

A következő csoportban már a tárgyat környezetében kell bemutatni, ez a típus rengeteg plusz információval gazdagíthatja az olvasót. Megemlíthetjük például a *bábovka* 'kuglóf' szócikket, ahol a sütemény izlésesen tálalva jelenik meg az olvasó előtt; a *kopacski* 'stoplis futballcipő' képén már a futball-labda és a futballpálya is látható, tehát a tágabb környezet segít az értelmezésben; a *dodávka* 'furgon' esetében a járművet egy raktárbejárat előtt, a kocsibeállón láthatjuk, illusztrálva ezzel 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 *kolobeska* 'roller' képén megjelenő rollerező fiatal vagy a *bublifuk* 'szappanbuborék-fújó gyermekjáték' alatt szereplő gyereket, aki éppen buborékot fúj, de a *bazén* 'medence' képe is pontosan a használatot és a tárgy funkcióját illusztrálja.

Esetenként melléknevekhez, igékhez is lehet képet csatolni, ilyenek például az *autoritativ* 'tekintélyelvű, tekintélyt parancsoló', a *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 bemutatható.

A vizuális tartalom mellett a szótárban hanganyagok is megjelennek: mivel a példamondatok egy része előszóban is elhangzott. A hanganyagból viszonylag kevés van, mivel ezen terület új fejlesztésnek számít.

Összegzés

A Termini magyar–magyar szótár és adatbázis egyedülálló vállalkozás a magyar nyelv lexikográfiai feldolgozásában. Ezt látszik alátámasztani a szócikkszerkezet összetettsége, információgazdagsága, vagy akár a szövegkontextusban idézett példamondatok eredetének jelzése. Olyan interaktív online szótár, ahol a szócikk adatstruktúráját a lekérdező a saját „információéhségéhez” tudja alakítani. Jobban megnézve még az is megtudható, hogyan lehet hranolkit (sült krumpli, hasáburgonya) és hermelint (szlovákiai gyártású camembert) enni.

Irodalom

- Benő A., Lanstyák I., P. Márkus K. & M. Pintér T. (2021). A Termini magyar–magyar szótár és adatbázis szerkezeti és szerkesztési sajátosságai. *Magyar Nyelvőr*, 145(4) (pp. 378–389).
- Hupka, W. (1989). *Wort und Bild: Die Illustrationen in Wörterbüchern und Enzyklopädien*. Tübingen: Max Niemeyer Verlag (Lexicographica, Series Maior 22).