

HUN-REN Nyelvtudományi Kutatóközpont, Eszterházy Károly Katolikus Egyetem
ludanyi.zsobia@nytud.hun-ren.hu
<https://orcid.org/0000-0002-8506-5684>

Ludányi Zsófia: A korpuszhasználat szerepe a nyelvi tanácsadói munkában
Alkalmazott Nyelvtudomány, XXIII. évfolyam, 2023/2. szám, 54–70.
doi:<http://dx.doi.org/10.18460/ANY.2023.2.004>

A korpuszhasználat szerepe a nyelvi tanácsadói munkában

The role of corpora in language consulting

One way of managing language problems in producing and proofreading written texts is for language users to contact language consulting institutions. Through answering questions received from members of the public, language consulting services offer guidance to inquirers in simple language management processes by elaborating an adjustment design (Beneš et al., 2018). In this way, institutional language consulting plays an essential role in bridging the gap between organized and straightforward language management processes, i.e., between the micro and macro levels of management (Kimura & Fairbrother, 2020).

In developing an adjustment design, language consultants evaluate the linguistic form in question from several points of view, providing information on its range of use, typical associated text types, stylistic value, and prevalence. Their assessment must be sensitive to the linguistic background and, depending on the kind of question, may also require language usage guides, dictionaries, and corpora.

Relying on Language Management Theory (Jernudd & Neustupný, 1987), the paper presents the use of corpus data in language consulting (cf. Van de Velde & Zenner, 2010) through the work of the Language Consulting Service (LCS) of the HUN-REN Hungarian Research Centre for Linguistics. In the formation of language advice, it is essential to consider the linguistic background, apply a data-driven approach, and, in some cases, provide access to linguistic data related to specific language problems. The research is based on the LCS database of approximately 10,000 emails and examines questions answered by language consultants informed by a corpus-based investigation (124 inquiries).

The results show that using corpus query interfaces (Sass, 2022) is well-suited for answering complex linguistic questions and describing particular expressions. On the other hand, corpus data are rarely needed to answer questions on orthography, which constitute the bulk of inquiries that the LCS receives. In answering spelling questions, the primary aim is to communicate the spelling norm; thus, examining usage is only necessary when spelling codification is absent or inconsistent. The paper presents a case study on a specific language use issue from the database, highlighting the strategy used to develop language advice in detail, with a particular focus on the appropriate use of corpus query interfaces.

Keywords: language consulting, corpora, corpus querying, language use problems, spelling problems

1. Bevezetés

A szövegekkel való professzionális foglalkozás során a nyelvhasználók számos összetettebb nyelvi problémával szembesülnek, amelyek megoldásában nyelvi tájékoztatásra van szükségük. A felmerülő nyelvi problémák kezelésének egyik módja, ha a nyelvhasználók nyelvi tanácsadással foglalkozó intézményekhez fordulnak.

A tanulmány a HUN-REN Nyelvtudományi Kutatóközpont (NYTK) nyelvi közönségszolgálatának adatai alapján arra a kérdésre keresi a választ, hogy

melyek azok a nyelvi problémák, amelyek kezelésében szükséges a nyelvészeti szövegkeresők, korpuszlekérdező eszközök alkalmazása, azaz milyen problémák esetén támaszkodtak korpuszadatokra a nyelvi tanácsadók a válasz kidolgozása során.

1.1. Elméleti háttér és alapfogalmak

Mivel a nyelvi tanácsadás a beszélők konkrét nyelvi problémáinak kezelésére irányuló tevékenység, a tanulmányban az intézményes nyelvi tanácsadás leírására a nyelvmenedzselés-elméletet (Jernudd & Neustupný, 1987) alkalmazom, melynek kulcsfogalmai a nyelvi problémák.

A dolgozatban nyelvmenedzselésnek tekintem egyrészt a nyelvi jelenségekre való reflektálást, másrészt az annak nyomán megvalósuló nyelvalkító tevékenységeket is (Ludányi et al., 2022: 75). A nyelvmenedzselésnek létezik egyszerű és szervezett formája, az utóbbit intézmények végzik, az egyszerű nyelvmenedzseléshez képest nagyobb hangsúlyt fektetve a nyelvvideológiai reflektáltságra (Nekvapil, 2012). Sajátos nyelvmenedzselési folyamatnak tekinthető a nyelvi közönségszolgálatokhoz forduló nyelvhasználók és a tanácsadók közti interakció (Beneš et al., 2018), amely kapcsolatot teremt a nyelvmenedzselés egyszerű és szervezett formái között (Kimura & Fairbrother, 2020).

A tanulmányban azokat a nyelvi jelenségeket tekintem nyelvi problémának, amelyek a beszélőből, a beszédpartnerből, de akár egy, az adott interakciót kívülről szemlélő harmadik személyből valamilyen reflexiót váltanak ki (Ludányi et al., 2022: 76). A nyelvi probléma tágabb értelmezése azért is szükséges, mivel a levélírók nemcsak a problematikusnak ítélt nyelvi jelenségekkel kapcsolatban kérnek tanácsot, hanem az általuk érdekesnek ítélt nyelvi formákról, helyzetekről kérnek tájékoztatást.

A dolgozatban Lanstyák (2014: 327–328, 2018: 31) munkái alapján elkülönítem egymástól a konkrét élethelyzetekben jelentkező, helyhez és időhöz, konkrét emberekhez köthető eseti problémákat („problémapéldányokat”), illetve az eseti problémák általánosításával létrejövő elvont kategóriát, a típusproblémákat (metaproblémákat). Noha a *probléma* szó igei kollokációi közül a *Mazsola* (Sass, 2009) adatai alapján a *megold* a leggyakoribb, Lanstyák (2018: 159–202) alapján a megoldást csak a problémakezelési módok egyik válfajának tekintem, így a továbbiakban tehát következetesen „a nyelvi problémák kezelése” terminust alkalmazom.

1.2. Segédeszközök a nyelvi tanácsadásban

Eőry (2007) tanulmányában részletes áttekintést nyújt azokról a szakmunkákról, amelyek vagy a nyelvhasználó közvetlen felhasználására készültek, vagy a nyelvi tanácsadó szolgálatok munkatársai alkalmazzák munkájuk során. Az NYTK közönségszolgálati könyvtárának nyomtatott segédeszközei mindenekelőtt az

alapvető helyesírási szakmunkák, az értelmező és egyéb szótárak, leíró nyelvtanok és a nyelvhasználati kézikönyvek.

A kérdések megválaszolásához sok esetben valós nyelvi adatokra van szükség. Ezek forrása egyrészt lehet a Google (ennek hátrányairól l. Sass, 2016: 798–799) vagy az *Arcanum.hu* (W1) adatbázisa, másrészt a nyelvészeti szövegkeresők, korpuszok. A dolgozatban a korpusz fogalmán az olyan, valamilyen szempont szerint kiválogatott, ténylegesen előforduló írott vagy lejegyzett beszélt nyelvi adatok gyűjteményét értem, amelyek tartalmazzák a szövegek bibliográfiai adatait, továbbá valamilyen annotációval rendelkeznek (MNSz.²). A korpuszadatok egyrészt közvetlen automatikus szövegfeldolgozással, másrészt ennél felhasználóbarátabb módon, korpuszlekérdező eszközök segítségével nyerhetők ki (Sass, 2022). Az utóbbiak olyan felületek, melyek segítségével nagyméretű szöveggyűjteményekben adott tulajdonságú szavak, szókapcsolatok keresésére van lehetőség. A keresési feltételek különféle nyelvészeti szempontok alapján állíthatók be (Sass, 2016). A korpuszlekérdező eszközök használata így nemcsak a nyelvtudományi kutatásokban meghatározó szerepű, hanem az olyan, tudományosan megalapozott gyakorlati munkákban is, mint a nyelvi tanácsadás (Van de Velde & Zenner, 2010).

1.3. Hipotézisek

A tanulmány az NYTK közönségszolgálati munkatársainak nézőpontjából, önreflektív módon vizsgálja azt a kérdést, hogy az elmúlt tíz évben milyen esetekben fordultak a munkatársak korpuszadatokhoz a kérdés megválaszolásakor. A kutatást az motiválta, hogy az eddigi korpuszhasználati szokások feltérképezése segíti és reflektálttá teszi a nyelvi tanácsadás folyamatát. A kutatási kérdés megválaszolásához az alábbi két hipotézist állítottam fel.

(H1) Az NYTK nyelvi közönségszolgálatának munkatársai a nyelvhasználati kérdések megválaszolásakor gyakrabban támaszkodnak ténylegesen előforduló nyelvi adatokra, mint a helyesírással kapcsolatos kérdéseknél.

(H2) A helyesírási kérdések megválaszolásakor a tanácsadók egyrészt az ellentmondásos esetekben, másrészt pedig a helyesírási kodifikáció hiánya esetén vizsgálják az írásgyakorlatot korpuszadatok segítségével.

2. Az elemzés anyaga és módszere

Az NYTK-ban az intézmény jogelődjének, az MTA Nyelvtudományi Intézetének megalapítása óta működik nyelvi közönségszolgálat. A tanulmányban bemutatott elemzés anyagát a közönségszolgálatához e-mailben érkezett nyelvhasználói kérdések és a rájuk adott válaszok képezik. A jelenleg több mint 10000 e-mailből álló nyelvi anyagból anonimizált adatbázis épül, amely 2011 szeptembere óta érkezett e-maileket tartalmaz (Domonkosi & Ludányi, 2023a: 24–25).

Az e-mailek elsődleges, a nyelvi probléma típusát körülhatároló annotálása már a levél beérkezésekor, legkésőbb annak megválaszolásakor megtörténik egy strukturált, dinamikus címkerendszer segítségével. Az elsődleges kódok a teljes levélváltásra (threadre) kerülnek rá, funkciójuk a nyelvi probléma típusának alapvető körülhatárolása, a kérdéses helyesírási vagy nyelvi jelenség azonosítása. Az e-mailek annotálása részben manuálisan, részben automatikusan történik. Az automatikus annotálás a levelezőrendszer szűrőfeltételei segítségével állítható be, így ha az adott levél megfelel a szűrőfeltételnek (pl. tartalmaz bizonyos kulcsszavakat), automatikusan rákerül a vonatkozó címke. A címkehierarchia legtetején a két fő problémátípus, a helyesírással és az egyéb, nem helyesírással kapcsolatos kérdések címkéje helyezkedik el. A közönségszolgálati kérdések 80-90%-a a magyar helyesírás szabályaival kapcsolatos (Ludányi et al., 2022). A kérdések fennmaradó 10-20%-a egyaránt tartalmaz nyelvhasználati tanácsot kérő e-maileket, valamint nem nyelvi problémák megoldására irányuló, hanem nyelvi tájékoztatást kérő megkereséseket is (pl. az egyes szavak jelentése, eredete, első előfordulásuk).

A nyelvi probléma típusát körülhatároló elsődleges annotálás mellett az e-mailek utólagos feldolgozása során további, a probléma típusától függő annotálási módok kialakításán dolgozunk (Domonkosi & Ludányi, 2023b). Ezek a kódok egyrészt vonatkozhatnak az e-mailek egyes, kisebb tartalmi egységeire is (pl. a probléma észlelési körülményeire), másrészt arra, hogy a levélben konkrét tanácskérés, javaslat jelenik-e meg (vö. Jakab, 2023). Az e-maileket abból a szempontból is annotáljuk, hogy megválaszolásukhoz a tanácsadó támaszkodott-e korpuszadatokra.

Az elemzéshez az alkalmazott kódrendszer segítségével azokat az e-maileket válogattam ki, amelyeken előzetesen a „korpuszhasználat/MNSz.”, valamint a „korpuszhasználat/MTSz.” annotáció szerepelt. A címkék a válaszadás során leggyakrabban használt két korpuszra, az NYTK munkatársai által fejlesztett *Magyar nemzeti szövegtár 2. verziójára* (MNSz.², Oravecz et al., 2014) és a *Magyar történeti szövegtárra* (MTSz., Sass, 2017) utalnak.¹ A levelek korpuszhasználat szerinti annotálása nagyrészt automatikusan történik: minden olyan válaszlevélben, amelyben a közönségszolgálati munkatárs reflektáltan hivatkozik a korpuszhasználatra, automatikusan rákerül a „korpuszhasználat/MNSz.” vagy „korpuszhasználat/MTSz.” kód. Amennyiben a tanácsadó kifejtetlenül, reflektálatlanul támaszkodik korpuszadatokra (ez általában nem jellemző), abban az esetben a korpuszhasználatra utaló címkével utólag, a válasz elküldésével párhuzamosan, manuálisan jelöli meg a levelezést.

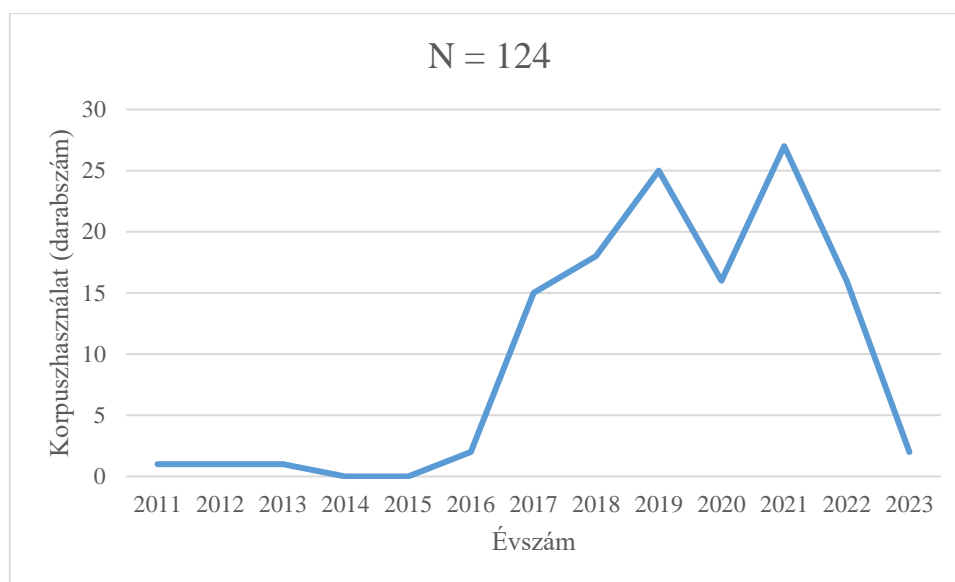
A két korpuszlekérdező rendszer használati gyakoriságát jól mutatják az elemzéshez kiválogatott levelek arányai: a teljes, több mint 10000 e-mailt

¹ A korpuszok nevének írásmódja ingadozást mutat, sokszor még az azonos intézmény által fejlesztett adatbázisok írásmódja sem egységes: a korpuszok keresőfelületén *Magyar Nemzeti Szövegtár*, illetve *Magyar történeti szövegtár* írásmóddal szerepelnek. A tanulmányban egységesen az egyedicím-szerű *Magyar nemzeti szövegtár*, *Magyar történeti szövegtár* írásmódot alkalmazom (Ludányi, 2019).

tartalmazó anyagból összesen 118 levél megválaszolása során támaszkodtak a munkatársak a magyar nyelv reprezentatív korpuszára, az MNSz.²-re, és nagyságrendekkel ritkábban, csupán 13 esetben az MTSz.-re. Hét esetben a tanácsadó – a több korpusz egyidejű használatának alapelvét (Sass, 2022: 606–608) érvényesítve – mindkét korpuszt segítségül hívta.

A korpuszlekérdező eszközök használatának gyakorlatában növekvő tendencia figyelhető meg. Míg a 2010-es évek első felében a korpuszhasználat ritkán fordult elő, addig 2015-től kezdve fokozatosan gyakoribbá vált, ahogy az 1. ábra is mutatja. (A kutatási anyag utolsó levele 2023. májusi, ez indokolja a 2023-as év kevés adatát.)

1. ábra. A korpuszadatokra támaszkodó nyelvi tanácsadói válaszok megoszlása évenkénti bontásban (2011. november – 2023. május)



Az előzetes szűréssel kapott, összesen 124 kérdés-válasz párból állítottam össze az elemzés anyagát, amelyet további szempontok alapján annotáltam. A levelezéseket a kérdező személyes adatait eltávolítva Excel-táblázatba rendeztem, minden egyes levelezést egyedi azonosítóval ellátva. Az így összeállított vizsgálati korpusz tartalmazza a kérdés beérkezésének dátumát, a nyelvi probléma tömör leírását (pl. *ékei ~ ékjei*), a válaszadó munkatárs monogramját, a kérdés és a válasz teljes szövegét, a felhasznált korpuszt és az esetleges megjegyzéseket. A táblázat tartalmaz továbbá egy „Helyesírás / Nem helyesírás” nevű attribútumot is, amely a „Helyesírás” és a „Nem helyesírás” értékek közül csak az egyik értéket veheti fel. Az elhatárolás néhány esetben nehézséget okozott, mivel a nyelvi probléma mindkét típusba besorolható volt. Tipikusan ilyen átmeneti kategóriájúak azok az esetek, amikor a nyelvi problémát az okozta, hogy a *Helyesiras.mta.hu* portál (Váradi et al., 2014) *Helyes-e így?* eszköze ismeretlennek ítélte a kérdező által megadott, tipikusan ritka, kisebb

csoport által használt szóalakot, elbizonytalanítva ezzel a felhasználót akár magának a szónak a létezésében is. Ilyen határesetre példa, amikor a levélíró a (feltehetőleg kisebb gyakorlatközösség által használt) *málházat* 'tűzoltó járműre felrakott felszerelés' szóval kapcsolatban kereste meg a közönségszolgálatot, jelezve, hogy a *Helyes-e így?* eszköz ismeretlennek minősítette a szót (#mnsz83).² Mivel azonban a levélből nem derült ki a felhasználó valódi célja (hogy ti. a *málházat* szónak az akadémiai helyesírási normát követő írásmódjára kíváncsi, vagy arra, hogy létezik-e ilyen szó), a nyelvi adat „Helyesírás / Nem helyesírás” szerinti kategorizációja nehézségekbe ütközött. Feltételezve, hogy inkább az utóbbira gondolhatott a levélíró, az ilyen eseteket a „Nem helyesírás” típusba soroltam be.

Az egyes konkrét nyelvi problémákat előzetesen áttekintve elvégeztem az eseti problémák általánosítását, létrehozva a főbb típusproblémák (metaproblémák) kategóriáit. Ha legalább két hasonló eseti probléma előfordult, azokat egy típusba soroltam be. Például amikor a nyelvi problémát a *mokaszin* főnév többes száma (#mnsz60) és a *nokedli* főnév *-val/-vel* ragos alakja megalkotásának (#mnsz80) nehézsége okozta, a problémát általánosítva külön kóddal jelöltem a vegyes hangrendű, elöl képzett, nem ajakkerekítéses magánhangzóra (*e, é, i, í*) végződő főnevek toldalékolási nehézségeit mint típusproblémák között. Az eseti problémák sokfélesége miatt – különösen a „Nem helyesírás” kategóriában – szükséges volt egy „Egyéb” típust is létrehozni azokra az esetekre, ahol az eseti problémák típusproblémává való összevonása nem volt elvégezhető.

3. Az elemzés eredményei

Az 1. táblázat a kapott eredmények összesített megoszlását mutatja be.

1. táblázat. A korpuszadatokra támaszkodó nyelvi tanácsadói válaszok megoszlása

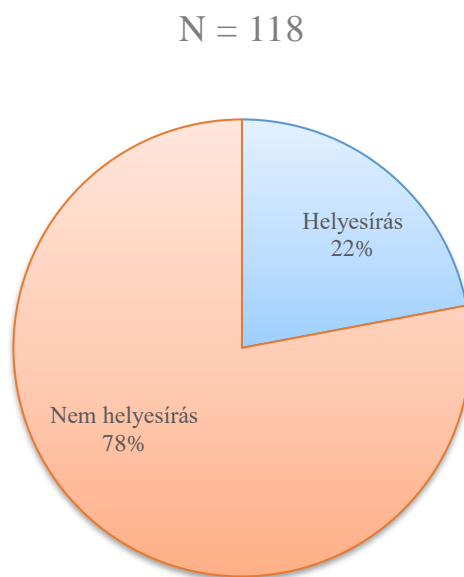
	Helyesírási kérdés	Nem helyesírási kérdés	Összesen
MNSz. ² (beleértve a más korpuszokkal történő együttes használatot)	26	92	118
Csak MTSz.	2	3	5
MTSz. és Tmk.	0	1	1
Összesen	28	96	124
%	~8000	~2000	~10000
	~0,35%	4,8%	~1,2%

² A zárójeles, kettőskereszttel kezdődő karaktersorozat a nyelvi probléma egyedi azonosítóját jelenti.

3.1. A Magyar nemzeti szövegtár használata

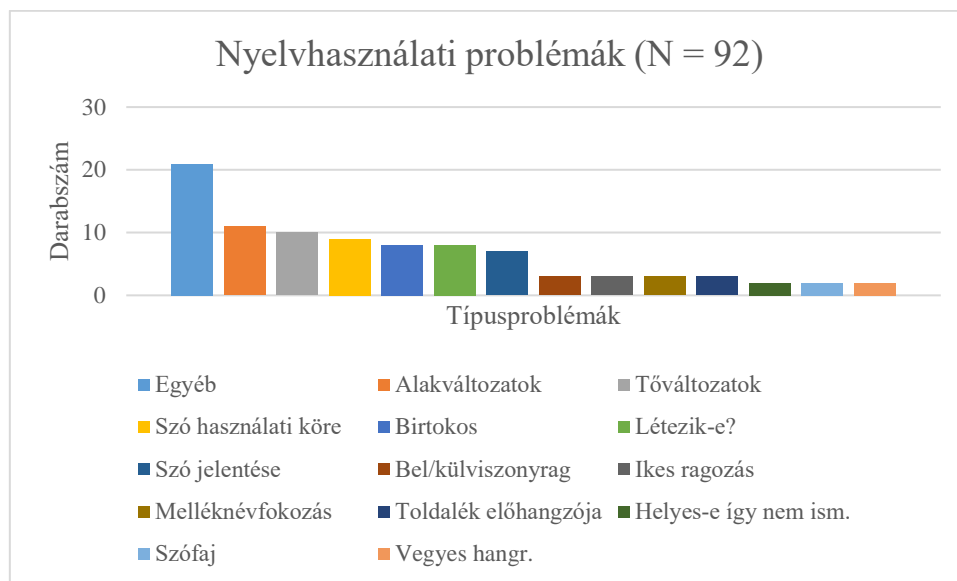
A magyar nyelv reprezentatív korpusza, a *Magyar nemzeti szövegtár* a leggyakrabban használt korpusz a közönségszolgálat munkájában. Az MNSz.² adatain alapuló 118 válaszlevélből a tanácsadók 26 esetben (22%) helyesírással kapcsolatos probléma kezelésében támaszkodtak korpuszadatokra, míg 92 alkalommal (78%) nem helyesírási (tipikusan nyelvhasználati) kérdésben hívták segítségül a korpuszlekérdező eszközt. A helyesírási és nem helyesírási kérdések megoszlását mutatja a 2. ábra.

2. ábra. Az MNSz.² adataira támaszkodó nyelvi tanácsadói válaszok megoszlása: helyesírással kapcsolatos és nem helyesírással kapcsolatos kérdések



Az egyes konkrét eseti problémák általánosításaként létrejött típusproblémák megoszlását a 3. és a 4. ábra illusztrálja. A 3. ábrán a nyelvhasználati típusproblémák számadatai láthatók.

3. ábra. A nyelvhasználattal kapcsolatos, az MNSz.² adatai alapján megválaszolt típusproblémák megoszlása. Az Egyéb kategóriában kizárólag egyszer előforduló (egyedi) problémák szerepelnek



A nyelvi variativitás az összes nyelvi szinten megjelenik, azaz elkülöníthetünk fonéma-, morféma-, lexéma-, szintaktikai és szövegvariánsokat is (Lőrincz, 2011: 353). Az elemzés adatai azt mutatják, hogy a közönségszolgálat munkatársai a nyelvhasználati problémák közül legtöbbször lexémavariánsokkal kapcsolatos kérdések megválaszolásában támaszkodnak korpuszadatokra (11 kérdés). Az annotálás során Lőrincz (2007: 100–101) alapján lexémaszintű variánsoknak, alakváltozatoknak egy adott szó paradigmáján belül előforduló, nagymértékű alaki azonosságot és kismértékű alaki eltérést mutató, denotatív jelentéseikben megegyező variánsokat tekintettem, amelyek jelentésbeli eltérései csak az egyes változatok pragmatikai jelentéseiben megengedettek. Ilyen alakváltozatok például az *irattározás* ~ *irattározás* (#mnsz12), *felülvizsgálat* ~ *fölvizsgálat* (#mnsz15).

A variativitás a morféma szintjén is megjelenik. Típusproblémának tekinthetők a tőváltozatokkal kapcsolatos kérdések: *olajos magok* ~ *magvak* (#mnsz23) (10 kérdés), a toldalékmorfémák szintjén pedig különösen a főnevek birtokos személyjeles vagy birtoktöbbsesítő jeles alakjainak megalkotásával kapcsolatban merülnek fel nehézségek (8 levél): *hotele* ~ *hotelje* (#mnsz116). A morfémaszintű variativitás további, ritkább, de folyamatosan jelen lévő típusproblémái: az ikes igeragozás kérdései (*nem kések el* ~ *nem késem el*, #mnsz97), a bel- és külviszonyragok használata, különösen a településnevek esetén (*Balmazújvárosban* ~ *Balmazújvároson*, #mnsz29), valamint az előhangzós és előhangzó nélküli toldalékok váltakozása (*orrszarvúk* ~ *orrszarvúak*, #mnsz106).

A fonémaszintű variativitás a leggyakrabban azokban az esetekben okoz nyelvi problémát, amikor a vegyes hangrendű, elől képzett, nem ajakkerekítéses magánhangzóra (*e, é, i, í*) végződő főnevekhez kell toldalékot kapcsolni. Ezek a

magánhangzók ugyanis bizonyos esetekben a harmónia szempontjából nem elülsőként, hanem semlegesen, áttetszően viselkednek, ilyenkor a toldalékolásban ingadozás mutatkozik (Markó, 2017: 105–106, részletesen l. Forró, 2013): *nokedlival ~ nokedlivel* (#mnsz80).

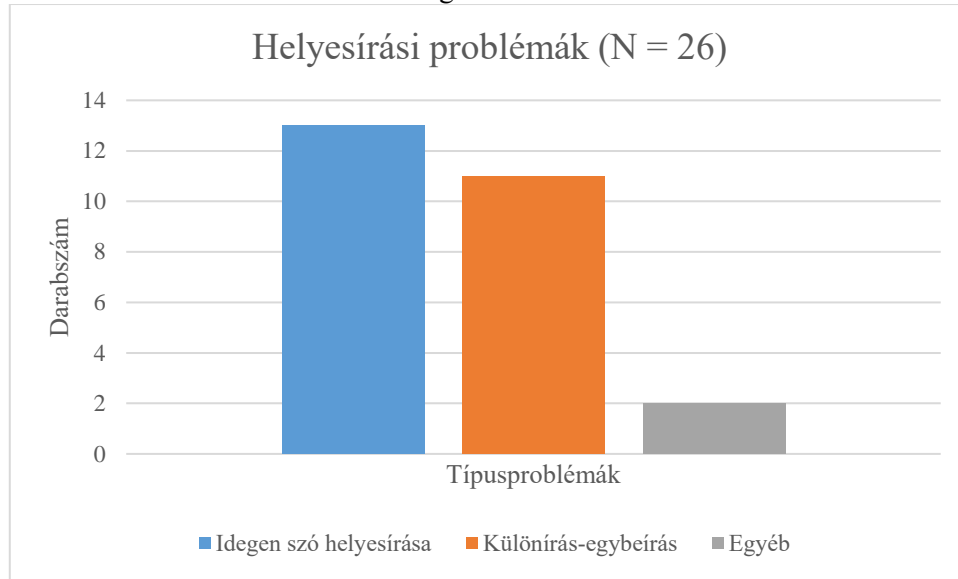
A nyelvhasználatra vonatkozó, az MNSz.² adatai alapján megválaszolt típusproblémák másik része nem a nyelvi variativitással kapcsolatos. Ezek közül az egyik leggyakoribb problémátípus az, amikor egyes lexémák (ritkán frazémák) használati köréről (9 levél), jelentéséről (7 levél), létezéséről (8 levél) kérnek tájékoztatást a kérdezők. Egy rovarirtással foglalkozó cég munkatársa azt szeretne volna tudni, hogy a gyakorlati szakmai nyelvhasználatukban elterjedt *büdösbogár* kifejezés használható-e az ügyfelekkel való kommunikációban is az ázsiai márványos poloska megnevezésére (#mnsz99, szó használati köre mint típusprobléma). Egy levélíró abban kérte a közönségszolgálat segítségét, hogy lehet-e egy lédús sütemény jelzője a *szaftos* – tekintve, hogy a *szaft* tipikusan 'zsíros, fűszeres húslé' jelentésben szokásos inkább a nyelvhasználatban (#mnsz113, szó jelentése mint típusprobléma). A „létezik-e ez a szó” típusú kérdések megválaszolása során is gyakran támaszkodnak a tanácsadók az MNSz.² adataira. Ebben a kategóriában egyrészt tipikusan ritka és/vagy újszerű, kisebb gyakorlatközösségek által használt szavak létezésével kapcsolatban érdeklődnek a levélírók (*blőr* 'fénykép egy részének elhomályosítása', #mnsz87), másrészt a helyesírás-ellenőrző eszközök által aláhúzott, ismeretlennek minősített szavak létezésére kérdeznek rá (*unkonvencionális*, #mnsz47).

Ugyancsak morfémaszintű, de nem a variativitással kapcsolatos típusprobléma a melléknevek fokozhatósága (3 levél), erre példa az *alapvető* felsőfokára vonatkozó kérdés (#mnsz66). Ritka típusprobléma az, amikor egyes lexémák szófajával kapcsolatban kérnek tájékoztatást a levélírók (pl. egy édesanya a 3. osztályos gyermeke nyelvtanfeladata kapcsán érdeklődött a *haladó* szófajáról, #mnsz63).

A nem általánosítható, eseti problémák kerültek az összevont „Egyéb” kategóriába, ezek a problémák nagy változatosságot mutatnak: mikorra adatható a *zöldbáró* főnév első előfordulása (#mnsz1); mit jelent a *fiaiéért* szóalak (#mnsz40), miért van többes számban a *mű* főnév az *Arany János összes művei* szerkezetben (#mnsz62) stb.

A közönségszolgálat munkatársai a nyelvhasználati problémák mellett ritkábban helyesírási kérdések megválaszolása során is segítségül hívják az MNSz.² adatait. A 4. ábra a helyesírással kapcsolatos típusproblémák megoszlását mutatja be.

4. ábra. A helyesírással kapcsolatos, az MNSz.² adatai alapján megválaszolt típusproblémák megoszlása



A nyelvhasználati problémákhoz képest a helyesírási típusproblémák nem mutatnak nagy változatosságot. A két fő típusprobléma, melynek kezelése során a nyelvi tanácsadók korpuszadatokra támaszkodnak: az idegen szavak helyesírása (13 levél) és a különírás-egybeírás területe (11 levél).

Az újonnan átvett kölcsönzők helyesírása tipikusan olyan helyesírási probléma, amelynél elengedhetetlennek tűnik az írásgyakorlat feltérképezése. A problémát minden esetben a forrásnyelvet követő idegenes és a kiejtést követő magyaros írásmód közti választás okozza. A magyar helyesírási szabályok szerinti a kölcsönzők írásmódja háromféle lehet: 1) az átadó nyelv írásmódját követő írásmód (pl. *déjà vu*, *flow*), 2) a magyar kiejtést követő (pl. *randevú* [*< rendez-vous*]), 3) vagylagos írásmód, szótározott helyesírási alakváltozatok (*chat* ~ *cset*). A helyesírási szótárakban mindhárom típusra akadnak példák. Ha azonban az újonnan kölcsönzött neologizmusok még nem rendelkeznek kodifikált írásmóddal, nehézséget okoz annak eldöntése, mikor írandók idegenesen, magyarosan, illetve vagylagosan. Az elemzett anyagban számos ilyen kérdéses esetre akad példa: *avatar* ~ *avatár* (#mnsz28), *gamer* ~ *gémer* 'számítógépes játékot a hobbiszintűnél komolyabban űző személy' (#mnsz45).

A különírás-egybeírás a magyar helyesírás egyik legproblematisabb területe, ezt a tényt a közönségszolgálatához érkező nagyszámú erre vonatkozó kérdés is bizonyítja. A különírással-egybeírással kapcsolatos kérdések egy részében a tanácsadók korpuszadatokra is támaszkodnak. Ilyen pl. a *testreszab* ~ *testre szab* írásmódja 'a felhasználó igényeihez igazít' jelentésben (#mnsz92).

3.2. A Magyar történeti szövegtár használata

Az 1772–2010 közötti szövegeket tartalmazó MTSz. adataira lényegesen ritkábban támaszkodnak a közönségszolgálat munkatársai a válasz kidolgozása során. Az ebbe a csoportba tartozó 13 levél közül az MTSz.-t hét esetben az MNSz.² mellett, kiegészítésként használták a nyelvi tanácsadók, egy alkalommal pedig az MTSz. mellett a Tmk. adataira is támaszkodtak. Az elemzett anyagban öt olyan nyelvi probléma található, amelyek kezelésében a tanácsadók csak az MTSz.-ből gyűjtött adatokat használták fel. Ezekben az esetekben egyedi, konkrét problémákról van szó, amelyek nem vonhatók össze típusproblémákká. A helyesírási és nem helyesírási kérdések arányát tekintve azt mondhatjuk, hogy az MTSz.-t – a más korpuszok melletti kiegészítő használatot is beleszámolva – 11 esetben nem helyesírási, 2 esetben pedig helyesírási kérdésben vették igénybe a nyelvi közönségszolgálat munkatársai.

4. Esettanulmány: a *tervez* ige segédigés szerkezetben

Egy pedagógus abban kérte a nyelvi közönségszolgálat véleményét, javítsa-e volt tanítványának nyelvhasználatát, aki – miután meghívta egy rendezvényre egykori tanárát – a következőképpen érdeklődött: [*Keresztnév*] bácsi *el tervez jönni*?

A kérdező levelében magyartalannak minősítette a *tervez* ige segédigei használatát, ugyanakkor reflektált arra is, hogy diákjának ez a fajta grammatikai szerkezet valószínűleg ugyanolyan természetes, mint a jövő idejű igealak használata: *el fogok menni*.

A tanácsadói válaszban megerősítettük, hogy az *el tervez menni* szerkezet valóban hasonlít a *fog* szóalakteremtő segédigével, analitikusan kifejezett jövő idejű igealakhoz. A tanácsadói stratégia első lépéseként a leíró nyelvtanokat tekintettük át. A *Magyar grammatika* (Lengyel, 2000: 258) ehhez hasonló aspektuális, modális és pragmatikai jelentésű szóalakteremtő segédigéknek tekinti a főnévi igenévi alakokkal megjelenő *szokott*, *tetszik* és *talál* igéket. A *tervez* ige is hangsúlytalanul követi a főnévi igenevet semleges pozitív kijelentő mondatban (*én el tervezek menni*), továbbá funkciója a realitástól való eltávolítás jelzése (Imrényi, 2017: 707), vagyis az, hogy a beszélő a jövőben elvégzendő cselekvést szeretné ugyan elvégezni, de nem ígéri biztosra, ezért az *el tervez menni* egyértelműen segédigés szerkezetként azonosítható.

A tanácsadói stratégia második lépése az *el tervez menni* szerkezethez hasonló korpuszadatok keresése: olyan valós nyelvi adatokra van szükség, ahol a *tervez* egy igekötős igébe ékelődik be. Ahhoz, hogy megfelelő találatokat kapjunk, elsőként meg kell határozni a vizsgálni kívánt nyelvi jelenség legsajátosabb felszíni elemét (Sass, 2022: 603), ez jelen esetben maga a *tervez* ige. Az MNSz.²-ben egy háromlépéses keresést végezve jutottunk el a kívánt nyelvi adatokig. A *tervez* igére szótóként rákeresve, a kapott konkordanciát szűrővel két lépésben leszűkítve nyertük ki a megfelelő találatokat (Sass, 2022: 609–610). A *tervez* szótőre kapott találatokat először úgy szűrtük meg, hogy a keresési

ablakot -1 , -1 -re állítva, a „Részletes keresés” funkciót használva a szófajt igekötőre állítottuk. Azaz egy olyan lekérdezést fogalmaztunk meg, hogy a találati szótól (*tervez* és ragozott alakjai) balra, pontosan egyszónyi távolságra igekötő szerepeljen. A második szűrésnél a keresési ablakot $1,1$ -re, a szófajt főnévi igenévre beállítva megfogalmaztuk azt a keresési feltételt, hogy a találati szótól jobbra, pontosan egyszónyi távolságra főnévi igenév szerepeljen.

Annak a hipotézisnek a bizonyítására, miszerint új keletű nyelvi jelenségről van szó, a korábbi szövegeket tartalmazó MTSz.-ben is keresést végeztünk. Az MTSz. az MNSz.²-vel ellentétben morfológiai (és más szempontból) elemzetlen korpusz, azaz a szöveganyagra egyszerűen szóalakok sorozataként tekinthetünk, így a releváns nyelvi adatokat is más módon kell kinyerni (Sass, 2017). A nyelvi adatok gyűjtéséhez összetettebb keresőkifejezés megfogalmazására volt szükség, reguláris kifejezésekkel (Kocsis, 2022: 45–51), az arra épülő Corpus Query Language (CQL) formális lekérdezőnyelven. Első lépéseként itt is a *tervez* igére kerestünk rá, de a morfológiai annotáció hiánya miatt a ragozott alakok kinyeréséhez "tervez.*" formában adtuk meg azt. Ezt követően a keresési ablakot -1 , -1 -re szűkítve beállítottuk, hogy a találati szótól balra, egyszónyi távolságban a leggyakoribb igekötők álljanak: "(el|ki|be|fel|le|át|meg)". Az így kapott öt találaton újabb szűrést végezve megadtuk, hogy a találati szótól jobbra, egyszónyi távolságban főnévi igenév szerepeljen, melynek legfőbb ismertetőjele a *-ni* végződés: ".*ni". Mivel a keresés erre már nem adott találatot, vagyis az 1772–2010 közti szövegekben nem található meg a *tervez* ige segédige szerkezetben, valóban új keletű nyelvi jelenségről lehet szó.

A tanácsadási stratégia következő lépéseként az MNSz.²-ből kinyert korpuszadatokból levontuk a következtetéseket. A keresést tovább finomítva, a téves találatokat manuálisan kiszűrve, alkorpuszok szerinti kereséssel megállapítottuk, hogy a *tervez* segédigei használata a sajtónyelvben a leggyakoribb (43-ból 23 találat), kisebb nagyságrendben jelen van a tudományos nyelvhasználatban (6 találat) és a közösségi oldalak személyes közléseiben is (5 találat), a többi stílusrétegben alig vagy egyáltalán nem fordul elő.

Néhány adatolt példa a szövegtárból:

- (1) Szintén **el tervezik adni** a [...] Hajógyári Sziget Kft. 43 százalékos állami üzletrészt (sajtónyelvi)
- (2) Az osztrák uralom alatt [...] a királysírokat **át tervezték szállítani** a krakkói Szent Péter és Pál templomba (tudományos)
- (3) **Le tervezem venni** a kari borító műanyagát (személyes-közösségi)

Összességében tehát a korpuszadatokból az a következtetés vonható le, hogy a közönségszolgálathoz forduló pedagógus jól érzékelt: a *tervez* segédigei használata nem része a választékosabb nyelvhasználatnak, mivel gyakrabban

jelenik meg lazább stílusú sajtónyelvi megnyilatkozásokban. Ugyanakkor mivel a tudományos stílusúhoz közel álló szövegekben (Wikipédia-szócikkekben) is előfordul, elképzelhető, hogy formálisabb szövegtípusokban is egyre megszokottabb lesz a használata.

5. Következtetések

Az eredmények alapján mindkét hipotézist sikerült igazolni.

(H1) Az NYTK közönségszolgálati levelezését szisztematikusan végigelemezve megállapítható, hogy a korpuszokban való keresés tipikusan a nyelvhasználattal kapcsolatos kérdések válaszadási stratégiájának a része. Az egyébként nagy számban érkező helyesírási kérdések megválaszolásában kevésbé jellemző a korpuszlekérdező eszközök használata. Ennek oka, hogy a helyesírási és a nyelvhasználati problémák kezelésének célja különbözik egymástól. A helyesírási problémák esetén a kérdezők szándéka az akadémiai helyesírási normához való igazodás, a tanácsadók is a normát közvetítik, így a létező írásgyakorlat korpuszalapú vizsgálatára nincs szükség. Bár a kérdezők nyelvszemléletében erőteljesen érvényesül a nyelvhasználati problémák „helyes-helytelen” minősítése, a tanácsadók célja az ilyen jellegű kérdések megválaszolásakor túlmutat a normaközvetítésen. Azokban az esetekben, amikor a nyelvhasználati kérdés feltevője egy nyelvi forma általános érvényű helyességéről érdeklődik, a válasz minden esetben reflektál arra, hogy a helyesség-helytelenség kérdésénél összetettebb helyzetről van szó. A nyelvhasználati kérdések esetén a tanácsadók célja a kérdéses nyelvi jelenség használati értékének adatalapú, kutatásokra, korpuszokra építő bemutatása. Az ilyen kérdések megválaszolása a kapcsolódó nyelvi adatokra építve, azokat láthatóvá téve történik, figyelembe véve a regiszter- és kontextusfüggő változatosságot és a beszélők nézőpontját, reflektálva a nyelvi tanács kidolgozásában érvényesített ideológiákra (Domonkosi & Ludányi, 2023a). Az a tény azonban, hogy a közönségszolgálati adatbázis összes nyelvhasználati kérdéséből (~2000 levél) csak 124 esetben történt korpuszhasználat, azt mutatja, hogy a nyelvhasználati problémák kezelésében sem látszik feltétlenül szükségesnek a korpuszadatokra való támaszkodás. Léteznek ugyanis olyan nyelvhasználati kérdések, amelyek megválaszolásához elegendőnek bizonyulnak a szótári adatok, minősítések, a leíró nyelvtanok és a nyelvhasználati kézikönyvek vonatkozó fejezetei és/vagy a kutatási eredmények. A nyelvi udvariasságra vonatkozó kérdések megválaszolásakor például a korpuszhasználat nem képezi a tanácsadói stratégia részét (Domonkosi & Ludányi, 2023b).

Az elemzés eredményei arra is rámutattak, hogy azok a nyelvhasználati problémák, amelyek kezeléséhez a közönségszolgálat munkatársai korpuszkeresést végeztek, két csoportba sorolhatók. A kérdések egyik része a nyelvi variativitással kapcsolatos, míg másik részük egyes lexémák, frazémák használati értékére, jelentésére, létezésére vonatkozik.

(H2) Az eredmények azt mutatják, hogy a helyesírási problémák kezelési stratégiájának csak abban az esetben része a korpuszhasználat, ha ellentmondásos vagy szabályozatlan helyesírási kérdéstről van szó. Korpuszhasználatra egyrészt azoknak az eseteknek egy részében volt szükség, amikor a helyesírási problémát a kodifikáció és a kialakult írásgyakorlat közötti ellentmondás okozta. Ezek a problémák a különírás-egybeírás területén jelentkeztek: a kodifikált helyesírási norma a különírás volt, a szaknyelvi gyakorlatban azonban az egybeírás volt a jellemző (*testre szab ~ testreszab, nagy mintás ~ nagymintás, jó gyakorlat ~ jógyakorlat*). Másrészt az elemzett adatok alapján a korpuszhasználat a helyesírási kodifikáció hiánya esetén is része a tanácsadási stratégiának. Különösen az újszerű kölcsönszók helyesírásához kapcsolódó kérdéseknél tipikus az írásgyakorlat feltérképezése korpuszadatok segítségével. Ennek az az oka, hogy a szótári kodifikációval nem rendelkező idegen átvételek helyesírására nincs általános érvényű szabály arra nézve, hogy ezeket a neologizmusokat mikor írjuk idegenesen vagy magyarosan, és mi alapján kerülnek be a szótárba vagylagos írásmóddal. A kodifikált írásmóddal nem rendelkező idegen neologizmusok írásmódját illetően általános alapelvnek tűnik, hogy minden egyes esetet egyedileg, többféle szempont alapján kell megvizsgálni. Az egyik legfőbb szempont az idegenes vagy magyaros írásmód gyakorisága, elterjedtsége, a kölcsönszó meghonosodottsága (idegen szó vagy jövevényszó)³, amely rámutathat, hogy a nyelvhasználók „mennyire érzik magyarnak” a kifejezést. Az ilyen esetekben a korpuszhasználat elengedhetetlen része a tanácsadási stratégiának.

6. Összegzés, kitekintés

A tanulmány az NYTK nyelvi közönségszolgálatához érkezett kérdésekre adott tanácsadói válaszokat elemezte abból a célból, hogy megvizsgálja, melyek azok a problémátípusok, amelyek kezelésében a korpuszlekérdező eszközök használata hatékonyan bizonyult. Az eredmények azt mutatják, hogy a korpuszhasználat az összetettebb nyelvhasználati kérdések megválaszolásában jelent hatékony segítséget. A lényegesen nagyobb számban érkező helyesírási kérdések megválaszolásában csak ritkán van szükség korpuszadatokra. A helyesírási kérdések megválaszolása esetén ugyanis a normaközvetítés az alapvető cél, így az írásgyakorlat vizsgálatára csak a helyesírási kodifikáció hiányakor vagy ellentmondásos helyesírási helyzetekben van szükség. A tanulmány a feldolgozott adatbázisból egy konkrét nyelvhasználati kérdést esettanulmányként kiemelve részletesen bemutatta a tanácsadásban érvényesített stratégiát, különös tekintettel a korpuszlekérdező eszközök használatára.

A kérdéstípusok számarányai azonban azt mutatják, hogy a válaszadás során a korpuszadatok segítségül hívása összességében viszonylag kevés alkalommal

³ Az idegen szavak és a jövevényszavak közti határ nem éles, sőt a kettő közti különbség valójában tudományosan megragadhatatlan (Nádasdy, 2003).

történt meg (124 eset), azaz a nyelvhasználati kérdések ~4,8%-ában, az összes kérdésnek megközelítőleg ~1,2%-ában. Az adatok alapján a nyelvhasználattal kapcsolatos kérdések ~95%-ában a korpuszhasználat nem képezte a nyelvi tanácsadási stratégia részét, így fontos jövőbeni feladatnak mutatkozik annak szisztematikus vizsgálata, hogy a különféle nyelvhasználati kérdések megválaszolásához tipikusan milyen segédeszközöket vesznek igénybe a nyelvi tanácsadók. A bemutatott kutatás folytatásaként tervezem alaposabban megvizsgálni azokat a nyelvhasználati problémákat, amelyek kezelésében nem történt korábban korpuszhasználat, abból a szempontból, hogy nem lett volna-e mégis szükséges és/vagy hasznos korpuszadatok gyűjtése és elemzése. Tekintettel ugyanis a korpuszhasználati szokásoknak az évek során történő változására (1. ábra), nagy valószínűséggel léteznek ilyen esetek is.

Irodalom

- Beneš, M., Prošek, M., Smejkalová, K., Štěpanová, V.** (2018). Interaction between language users and a language consulting center: Challenges for language management theory research. In Fairbrother, L., Nekvapil, J. & Sloboda, M. (eds), *The Language Management Approach: A Focus on Research Methodology* (119–140). Berlin: Peter Lang.
- Domonkosi Ágnes & Ludányi Zsófia** (2023a). Nyelvhasználói nézőpontok, adatalapúság, nyelvi tanácsadás. Hogyan beszélhetünk ma nyelvművelésről? In Tolcsvai Nagy Gábor, Laczkó Krisztina & Tátrai Szilárd (szerk.), *Nyelv, kultúra és tudomány. Köszöntő kötet a Magyar Nyelvőr alapításának 150. évfordulójára* (9–36). Budapest: ELTE Eötvös Kiadó. doi: <https://doi.org/10.21862./MNyr150fest2023>
- Domonkosi Ágnes & Ludányi Zsófia** (2023b). „Mi a helyes megszólítás?” Nyelvi udvariassági problémák a nyelvi tanácsadói gyakorlatban. In Albert Ágnes, Bóna Judit, Borbás Gabriella Dóra, Brdar-Szabó Rita, Csizér Kata & Vladár Zsuzsa (szerk.), *Fejezetek az alkalmazott nyelvészet területéről*. Budapest: Akadémiai Kiadó. Letöltés: https://mersz.hu/hivatkozas/m1062faant_2
- Dömötör Adrienne, Gugán Katalin, Novák Attila & Varga Mónika** (2017). Kiútkeresés a morfológiai labirintusból – korpuszépítés ó- és középmagyar kori magánéleti szövegekből. *Nyelvtudományi Közlemények*, 113, 85–110. doi: <https://doi.org/10.15776/NYK.2017.113.3>
- Eőry Vilma** (2007). A nyelvművelés segédeszközei: szótárak, nyelvtanok, nyelvészeti szakirodalom. In Domonkosi Ágnes, Lanstyák István & Posgay Ildikó (szerk.), *Műhelytanulmányok a nyelvművelésről* (87–94). Budapest & Dunaszerdahely: Gramma Nyelvi Iroda & Tinta Könyvkiadó.
- Forró Orsolya** (2013). *Ingadozás a magyar előlségi harmóniában. Szempontok a variabilitás szinkroniájának és diakroniájának feltárásához és értelmezéséhez*. Doktori értekezés. Piliscsaba: Pázmány Péter Katolikus Egyetem Nyelvtudományi Doktori Iskola. doi: <https://doi.org/10.15774/PPKE.BTK.2013.016>
- Imrényi András** (2017). Mondattan. In Tolcsvai Nagy Gábor (szerk.), *Nyelvtan* (661–895). Budapest: Osiris Kiadó.
- Jakab Dorottya** (2023). A beszédcselekvés-érték jelölésének lehetséges gyakorlatai az intézményes nyelvi tanácsadásban. *Alkalmazott Nyelvtudomány*, 2023/2, 19–35.
- Jernudd, B. H. & Neustupný, J. V.** (1987). Language planning: for whom? In Laforge, L. (ed.), *Proceedings of the International Colloquium on Language Planning* (69–84). Québec: Les Press de L'Université Laval.
- Kimura, G. C. & Fairbrother, L.** (2020). Reconsidering the language management approach in light of the micro-macro continuum. In Kimura, G. C. & Fairbrother, L. (eds), *Language Management Approach to Language Problems: Integrating Macro and Micro Dimensions* (255–267). Amsterdam (Philadelphia): John Benjamins. doi: <https://doi.org/10.1075/wlp.7.13kim>

- Kocsis Ágnes** (2022). *Informatikai eszközök a nyelvészeti kutatásban: Segédanyag az informatikai eszközök használatához kognitív nyelvészeknek* (45–51). Budapest: ELTE BTK Nyelvtudományi Doktori Iskola & ELTE Eötvös Kiadó. Letöltés: https://www.eltereader.hu/media/2022/11/KocsisAgnes_Informatikai_eszkozok_segedanyag.pdf
- Lanstyák István** (2014). On the process of language problem management. *Slovo a slovesnost*, 75(4), 325–351.
- Lanstyák István** (2018). *Nyelvalakítás és nyelvi problémák*. Somorja: Fórum Kisebbségkutató Intézet & Gramma Nyelvi Iroda.
- Lengyel Klára** (2000). A segédigék és származékaik. In Keszler Borbála (szerk.), *Magyar grammatika* (254–260). Budapest: Nemzeti Tankönyvkiadó.
- Lőrincz Julianna** (2007). Azonosságok és különbségek a lexikológiai jelentések körében. *Acta Academiae Paedagogicae Agriensis. Sectio Linguistica Hungarica*, 34, 98–104.
- Lőrincz Julianna** (2011). A nyelvi variánsok helye és szerepe az anyanyelvoktatásban. In Boda István Károly & Mónos Katalin (szerk.), *MANYE XX. Az Alkalmazott nyelvészet ma: innováció, technológia, tradíció. I. kötet* (352–359). Budapest & Debrecen: MANYE & Debreceni Egyetem.
- Ludányi Zsófia** (2019). Korpuszneveink helyesírásáról. In *Szabályok, normák, nyelvszokás. Tanulmányok a köznyelvi és szaknyelvi helyesírás és nyelvalakítás köréből* (62–68). Eger: Liceum Kiadó.
- Ludányi Zsófia, Domonkosi Ágnes, Kocsis Ágnes & Jakab Dorottya** (2022). A nyelvi menedzselés szemlélete és a nyelvi tanácsadás. In Deme Andrea, Kuna Ágnes (szerk.), *Tanulmányok a nyelvészet alkalmazásainak területéről* (73–107). Budapest: ELTE Eötvös Kiadó.
- Markó Alexandra** (2017). Hangtan. In Tolcsvai Nagy Gábor (szerk.), *Nyelvtan* (75–203). Budapest: Osiris Kiadó.
- MNSz.² = A Magyar nemzeti szövegtár keresőfelülete. <http://clara.nytud.hu/mnsz2-dev/> (Letöltve: 2023. december 29.)
- MTSz. = A Magyar történeti szövegtár keresőfelülete. <http://clara.nytud.hu/mtsz/> (Letöltve: 2023. december 29.)
- Nádasdy Ádám** (2003). Használjunk minél több idegen szót! In Nádasdy Ádám: *Ízlések és szabályok* (234–236). Budapest: Magvető.
- Nekvapil, J.** (2012). Some thoughts on “noting” in language management theory and beyond. *Journal of Asian Pacific Communication*, 22(2), 160–173.
- Oravecz Csaba, Váradi Tamás & Sass Bálint** (2014). The Hungarian Gigaword Corpus. In Calzolari, N., Choukri, K., Declerck, T., Loftsson, H., Maegaard, B., Mariani, J., Moreno, A., Odijk, J. & Piperidis, S. (eds.), *Proceedings of Ninth International Conference on Language Resources and Evaluation (LREC 2014)* (1719–1723). Reykjavik: ELRA.
- Sass Bálint** (2009). „Mazsola” – eszköz a magyar igék bővítményszerkezetének vizsgálatára. In Váradi Tamás (szerk.), *Válogatás az I. Alkalmazott Nyelvészeti Doktoranduszkonferencia előadásaiból* (117–129). Budapest: MTA Nyelvtudományi Intézet.
- Sass Bálint** (2016). Nyelvészeti szövegkeresők, Nemzeti Korpuszportál. *Magyar Tudomány*, 7, 798–808.
- Sass Bálint** (2017). A kibővített Magyar Történeti Szövegtár új keresőfelülete. In Forgács Tamás, Németh Miklós & Sinkovics Balázs (szerk.), *A nyelvtörténeti kutatások újabb eredményei IX.* (267–277). Szeged: SZTE Magyar Nyelvészeti Tanszék.
- Sass Bálint** (2022). Principles of corpus querying: A discussion note. *Acta Linguistica Academica*, 69(4), 599–614. doi: <https://doi.org/10.1556/2062.2022.00581>
- Tmk. = A Történeti magánéleti korpusz keresőfelülete. <https://tmk.nytud.hu/> (Letöltve: 2023. december 29.)
- Tolcsvai Nagy Gábor** (szerk.) (2017). *Nyelvtan*. Budapest: Osiris Kiadó.
- Van de Velde, F. & Zenner, E.** (2010). „Pimp my Lexis”: het nut van corpusonderzoek in normatief taaladvies. In Hendrickx, E., Hendrickx, K., Martin, W., Smessaert, H., Van Belle, W. & Van der Horst, J. (eds.), *Liever meer of juist minder? Over normen en variatie in taal* (51–68). Gent: Academia Press.

Váradi Tamás, Ludányi Zsófia & Kovács Réka (2014) Géppel segített helyesírás.
A helyesírás.mta.hu portál készítéséről. *Modern Nyelvoktatás*, 20(1–2), 43–58.
W1 = Arcanum.hu. <https://adt.arcanum.com/hu> (Letöltve: 2023. július 22.)