

NYELVTECHNOLÓGIAI ESZKÖZÖKKEL (NEM) KEZELHETŐ HELYESÍRÁSI PROBLÉMÁK

LUDÁNYI ZSÓFIA | HUN-REN Nyelvtudományi Kutatóközpont
ludanyi.zsafia@nytud.hun-ren.hu | DOI: 10.18135/PG70.2024.9

1. BEVEZETÉS

Prószéky Gábort a Budapesti Műszaki és Gazdaságtudományi Egyetem másodéves mérnök informatikus (BSc) szakos hallgatójaként ismertem meg 2010 környékén. Adigra már megszereztem magyar szakos bölcsész és tanári diplomám, végzős hallgatója voltam az ELTE Nyelvtudományi Doktori Iskolájának, és azon gondolkoztam, hogyan tudnám a kétféle érdeklődési területemet, a helyesírást és az informatikát „összeházasítani”. Így ismerkedtem meg a nyelvtechnológia, a számítógépes nyelvészet területével, ahonnan egyenes út vezetett Prószéky Gáborhoz. Később, 2012-ben, utolsó éves műegyetemistaként hivatalosan is hallgatója lehettem, mivel két féléven keresztül a Pázmány Péter Katolikus Egyetem Információs Technológiai Karának vendéghallgatójaként jártam hozzá nyelvtechnológiai alapozó kurzusokra. Nagy megtiszteltetés és öröm volt számomra, hogy külsős témavezetőként segítette a BME-s önálló laborom, később BSc-s szakdolgozatom elkészítését.

Ahogy e rövid visszaemlékezés is mutatja, Prószéky Gábort születésnapja alkalmából köszöntő írásomhoz nem kellett sokáig keresnem a témát: hiszen az, hogy miként segíthetik a nyelvtechnológiai eszközök az írott szövegek helyesírásának az akadémiai normához való igazítását, az ünnepelt munkásságának egyik meghatározó területe (pl. Prószéky, 1998; Prószéky & Kis, 1999: 81–104). Köszöntő írásomban ehhez a témakörhöz kapcsolódva, a HUN-REN Nyelvtudományi Kutatóközpont nyelvi tanácsadással foglalkozó munkatársaként azt mutatom be, hogyan oszthatóak a magyar beszélőközösségben élő helyesírási problémák a gépi kezelhetőség szempontjából. A tanulmány címében szereplő kerek zárójel az értelmező szótári hagyományokhoz igazodva vagylagosságot jelent, tehát a szöveg a zárójeles rész hozzáolvasásával és anélkül is érvényes. Ebből adódóan a jelen írásban egyaránt foglalkozom azokkal a helyesírási problémákkal, amelyek kezelésében hatékony segítséget, megoldást jelentenek a nyelvtechnológiai eszközök, valamint azokkal is, amelyek különféle okokból kívül esnek a gépi kezelhetőség területén.

2. A NYELVI ÉS HELYESÍRÁSI PROBLÉMÁK KEZELÉSE

A tanulmány középpontjában – a nyelvmenedzselés-elméletre (Jernudd & Neustupný, 1987) támaszkodva – a nyelvi problémák állnak. Nyelvi problémáknak a beszélőből, a

beszédpartnerből vagy az adott interakciót kívülről szemlélő harmadik személyből valamilyen reflexiót kiváltó nyelvi jelenségeket tekintem (Ludányi et al., 2022: 76). Itt jegyzem meg, hogy a továbbiakban következetesen a nyelvi problémák *kezeléséről*, nem pedig *megoldásáról* fogok beszélni. Habár a *probléma* főnév igei kollokációi közül a *Ma-zsola* (Sass, 2009) adatai alapján a *megold* a leggyakoribb, Lanstyák (2018: 159–202) alapján a megoldást csak a problémakezelési módok egyik válfajának tekintem, hiszen a problémát például el is lehet kerülni vagy letagadni. Mindez fokozottan igaz azokra a problémákra, amelyekkel a nyelvhasználók az írott szövegek létrehozása és/vagy nyelvi ellenőrzése közben szembesülnek. Például ha egy korrektor nem tudja eldönteni, hosszú vagy rövid *i*-vel írja-e a helyesírási kodifikációval nem rendelkező *elh[íj]vatás*¹ szót, elkerülheti a nyelvi problémát azáltal, hogy kicseréli azt egy szinonimával, átfogalmazza a szöveget. Maga a helyesírási probléma – hogy ti. a szónak nincs helyesírási szótárban rögzített írásmódja – azonban továbbra is megoldatlan marad (Ludányi, 2023a: 40–41).

A helyesírási problémák kezelésének az alapvető célja tehát a mindenkori akadémiai helyesírási normához való igazodás. Ezt a célt szolgálja a nyelvtechnológiai eszközökkel támogatott helyesírás-ellenőrzés, illetve – lehetőség szerint a nyelvhasználói nézőpontokat, adott esetben a sajátos (szakmai) írásgyakorlatot is figyelembe véve – az emberi erővel végzett helyesírási tanácsadás is (l. 2.2 fejezet).

2.1 Gépesített helyesírás-ellenőrzés és helyesírási tanácsadás

A nyelvtechnológiai eszközökkel támogatott helyesírás-ellenőrzés megjelenése előtt a helyesírási problémák kezelése a helyesírási szabályzat és szótárak segítségével történt. A helyesírási problémák újabb, korszerű és egyszerű kezelési lehetőségei a szövegszerkesztőkbe épített helyesírás-ellenőrző programok, amelyek a 20. század utolsó évtizedében jelentek meg, majd terjedtek el (Prószéky, 1993). A kezdetleges helyesírás-ellenőrző eszközök használata azonban sok esetben újabb nyelvi problémákat generált a felhasználók számára, mivel csupán az ortográfiai szavak (Fóris, 2020: 20) szintjén tudták felismerni a normától eltérő alakokat, vagyis két betűköz vagy központozási jel között elhelyezkedő karaktorsorokkal dolgoztak. A mai környezetérzékeny helyesírás-ellenőrzők azonban statisztikai alapon többé-kevésbé képesek kiszűrni az efféle tévedéseket (Prószéky, 2019: 270–271).

A szövegszerkesztőkbe, később a böngészőkbe integrált programok mellett internetes weboldalak is segíthetik a nyelvhasználókat egyszerűbb helyesírási kérdéseik megválaszolásában. Ilyen például a *Helyesiras.mta.hu* tanácsadó portál. A HUN-REN Nyelvtudományi Kutatóközpont jogelődje, az MTA Nyelvtudományi Intézet által fejlesztett online helyesírási tanácsadó weboldal (Miháltz et al., 2013) hétféle helyesírási területen kínál gyors és automatikus választ: külön- és egybeírás, szószintű helyesírás-ellenőrző (*Helyes-e így?* eszköz), elválasztás, tulajdonnevek írása, számnevek helyesírása, keltezés, betűrendbe sorolás.

¹ A kurzívval szedett karaktorsorozat reguláris kifejezésként értelmezendő.

2.2 A helyesírási tanácsadó portál együttműködése a „humán” helyesírási tanácsadóval

A HUN-REN Nyelvtudományi Kutatóközpontban a jogelőd intézmény megalapítása óta működik nyelvi közönségszolgálat, amely e-mailben és telefonon válaszol az érkező helyesírási, nyelvhasználati kérdésekre (Ludányi et al., 2022). A több mint tíz éve működő *Helyesiras.mta.hu* portál tapasztalatai azt mutatják, hogy a helyesírási tanácsadás régi és új formái egymással folyamatos kölcsönhatásban vannak, a közönségszolgálat és a helyesírási portál egymást kiegészítve működik. Egyrészt azért, mert a felhasználói visszajelzéseket, javaslatokat idővel beépítjük a portál eszközeibe. Másrészt pedig azért, mert a helyesírási portál a felmerülő helyesírási problémáknak csak egy részét oldja meg. A tapasztalatok alapján a kérdezők egy része, mielőtt megkeresi a közönségszolgálatot, a helyesírási portál eszközeivel kísérli meg helyesírási problémája kezelését. Ha az automatikusan generált választ nem találják megfelelőnek, a közönségszolgálathoz fordulnak. A portállal kapcsolatos visszajelzések a következőképpen csoportosíthatók (Ludányi, 2023b):

- a) A portál által nyújtott javaslat megerősítésének kérése a közönségszolgálattól.
- b) A portál javaslatának elutasítása, egyet nem értés (ezen belül: egyet nem értés helyes válasz esetén, egyet nem értés hibás válasz esetén, a portál javaslatának részleges elfogadása, vagyis kiegészítési javaslat nyújtása az automatikus válaszhoz).
- c) Ellentmondás észlelése a különféle segédeszközök javaslataiban (a portál javaslata eltér a hagyományos, papíralapú szótárétól).
- d) Segítségkérés a közönségszolgálattól a portál által felajánlott többféle helyesírási javaslat közti választásban.
- e) A portál nem tudott segíteni, mert nem adott választ.
- f) A felhasználó a portálon nem a megfelelő eszközhöz fordult, ezért nem oldódott meg a nyelvi problémája.
- g) A felhasználó a portálon nem a megfelelő formátumban adta meg a bemenetet a specifikus, megszorított bemenetet váró eszközöknél, ezért nem oldódott meg a nyelvi problémája.
- h) Egyéb, a *Helyesiras.mta.hu* által előidézett egyéb nyelvi problémák esetén.

3. A HELYESÍRÁSI PROBLÉMÁK OSZTÁLYOZÁSA

A fentiek alapján az írott szövegek alkotása és/vagy nyelvi ellenőrzése során a nyelvhasználók a helyesírási problémáknak többféle típusával szembesülhetnek (Ludányi, 2023a). A nyelvtechnológiai kezelhetőség szempontjából (is) alapvető, hogy külön típusként kell számontartani a helyesírási kodifikáció által „lefedett” és „le nem fedett” helyesírási problémákat.

3.1 A helyesírási kodifikáció által lefedett helyesírási problémák

3.1.1 Egyszerű esetek

Kezelhetőségük szempontjából azokat a helyesírási problémákat tekintem egyszerűnek, amelyek megoldhatók az akadémiai helyesírási szabályzat, a helyesírási kézikönyvek és a szótárak segítségével. Ebben az értelemben tehát „egyszerű” eset az is, amikor maga a helyesírási szabály összetett (pl. szótagszámlálás szabálya), de egyértelmű útmutatást ad a nyelvhasználónak az írásmódot illetően. Azaz a nyelvhasználó a szabály elolvasása és értelmezése után le fogja tudni írni a problematikusnak ítélt alakulatot a mindenkori helyesírási normának megfelelően.

Azokban az esetekben, amikor a helyesírási szabályok leírhatók algoritmusokkal, és/vagy lexikális adatbázisokkal, kivétellistákkal kezelhetők, a *Helyesiras.mta.hu* portál eszközei és a helyesírás-ellenőrző szoftverek is hatékony segítséget jelenthetnek a felhasználónak. A helyesírási portál szószintű helyesírás-ellenőrzője, a *Helyes-e így?*, valamint a tulajdonnevek (főként földrajzi nevek) helyesírási kérdéseiben segítő *Névkereső* eszköz használata lényegében a helyesírási szótárban való utánanézést váltja ki, a szótári keresés gyorsabb változatának feleltethető meg. A weboldal további, összetettebb működésű eszközei, mint például a környezetfüggetlen, attribútum-érték struktúrás formális nyelvtanon alapuló *Külön vagy egybe?* (Ludányi et al., 2013), valamint a *Számok* és a *Dátumok* eszközök pedig a helyesírási szabályok azonosítását és alkalmazását végzik el a nyelvhasználó helyett.

Léteznek azonban olyan területei is a magyar helyesírásnak, amelyeket a helyesírási portál eszközei nem képesek kezelni. Különösen a különírás-egybeírás területén fordulnak elő ilyen szabályok. A helyesírási portál „Hogyan működik?” menüpontja² részletesen tájékoztat ezekről a helyesírási területekről. Ilyenek például a mellérendelő szerkezetek, illetve összetételek helyesírásával kapcsolatban felmerülő külön- vagy egybeírás, illetve kötőjelezési problémák (pl. *testi-lelki*, *rüggkapál*, *napról napra*). A *Külön vagy egybe?* eszköz az esetekben sem tud segíteni, amikor egy többszörös szókapcsolatban két tag egyetlen egységként kapcsolódik egy harmadik elemhez (pl. *kutya-macska barátság*, *réz-arany ötvözet*). Ehhez ugyanis a programnak ismernie kellene az alkotó tagok közti szintaktikai és jelentéstani viszonyokat, ezzel a tudással azonban jelenleg nem rendelkezik. A helyesírási portál fejlesztés alatt lévő, 2.0-s verziójában (Prószék & Várad, 2023) a mesterséges intelligenciát használó, a kontextus pontosabb felismerését lehetővé tévő interaktív technológia azonban várhatóan ki fogja tágítani a gépi módszerekkel kezelhető helyesírási problémák körét.

3.1.2 Problematikus esetek

Azt is látni kell azonban, hogy maga a helyesírási kodifikáció sem tökéletes: előfordulnak benne ellentmondások, következetlenségek. Ezeknek az eseteknek a gépi kezelése pontosan azért ütközik nehézségekbe, mert nem létezik rájuk egyértelmű megoldás.

² <https://helyesiras.mta.hu/helyesiras/default/howitworks> (2024. január 13.)

Ilyen például a *humán* szóval létrejött alakulatok írásmódja. A probléma oka az, hogy a szótárban található példák írásmódjából nem következtethető ki egyértelműen az az alapelv, hogy mi alapján íródik egybe, illetve külön a *humán* előtag az azt követő szóval.³ A szótári példákat áttekintve annyi látszik biztosnak, hogy a *humán* 'társadalomtudomány(okat tanító)' jelentésben különírandó az azt követő szótól: *humán tagozat*, *humán tudomány*, *humán gimnázium*. A *humánnak* létezik azonban az 'emberi, emberrel kapcsolatos' jelentése is, ezekben az esetekben már nem tűnik következetesnek a kodifikáció (részletesen I. Ludányi, 2023c). Kiindulva abból, hogy a gépesített helyesírási tanácsadás célja is a normaközvetítés, kérdésként merül fel, mit közvetítenek az eszközök abban az esetben, ha például a szótárban nem szereplő *humán* ()*faktor*, *humán*()*egészségügy* alakulatok lejegyzése okozza a helyesírási problémát. A *humán faktor* esete a jelentésviszonyokat tekintve inkább a helyesírási szótárban különírvá rögzített *humán erőforrás*, *humán szféra* szerkezetekhez hasonló, a *humán* + *egészségügy* esetében a *humánpolitika* analógiája miatt elképzelhető az egybeírás, de a szótári kodifikáció ellentmondásossága miatt tulajdonképpen mindkét írásmód elfogadható: *humánegészségügy*, *humán egészségügy*. A *humán* szóval létrejött alakulatok írásmódja tehát azért nem írható le algoritmusokkal, mert maga a „szabály” nem egyértelmű. (Számos további ellentmondásos esetről I. Ludányi, 2023a: 37–40).

Az ellentmondás nemcsak magában a kodifikációban lehet, hanem a kodifikáció és egy kisebb-nagyobb közösség írásgyakorlata között is. Ilyen például a *kis*()*csoport* külön- vagy egybeírása. A *Magyar helyesírási szótár* (Tóth, 2017) csak óvodai kontextusban írta egybe a *kiscsoportot*, 'kis létszámú csoport' jelentésben a különírást javasolja. Az egyes szakmák írásgyakorlata (pl. szociológia, társadalomtudományok) azonban ellentmond ennek, mivel következetesen az egybeírást alkalmazza. Ez a gyakorlat indokolható is, mivel ezekben az esetekben nem pusztán kis létszámú csoportról van szó, hanem egy bizonyos, konkrétan meghatározott szám alatti csoportról. Kérdéses azonban, hogy miként járjon el egy helyesírási tanácsadó program ezekben az esetekben: szorítkozzon pusztán a kodifikált helyesírási norma közvetítésére, és 'kis létszámú csoport' jelentésben írassa külön (*kis csoport*), vagy pedig figyelembe véve a szakmai írásgyakorlatot és a jelentésváltozást, bírálja felül a kodifikált helyesírási normát, és írassa egybe (*kiscsoport*)?

3.2 A helyesírási kodifikáció által le nem fedett helyesírási problémák

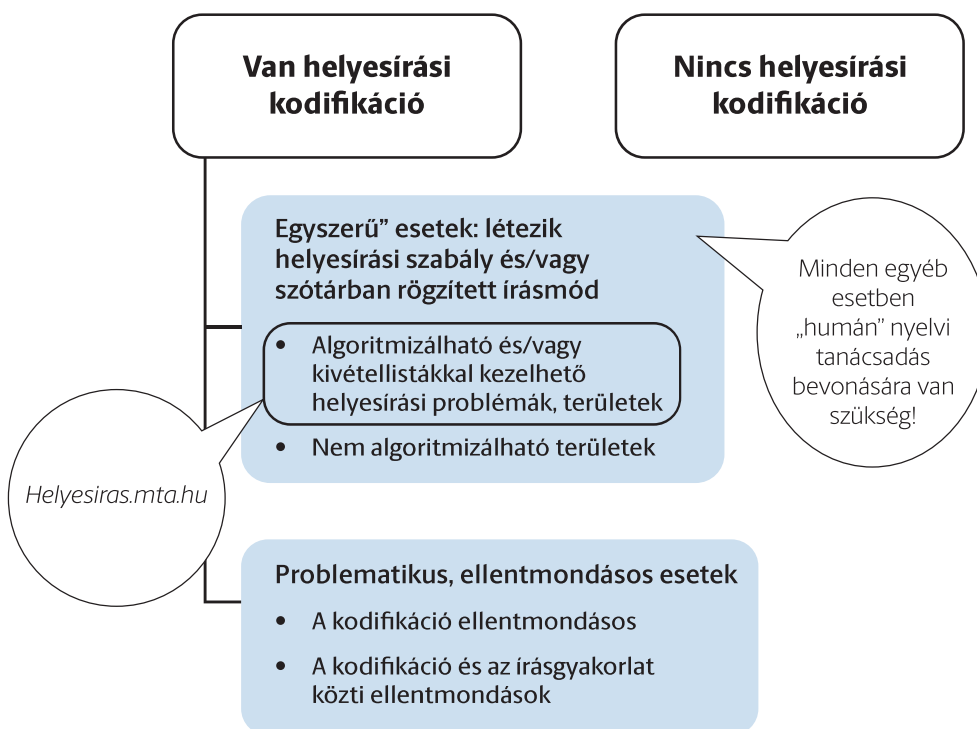
A helyesírási problémák gyakori forrása a helyesírási kodifikáció hiánya is, amely több szinten jelentkezhet. A kodifikáció hiánya vonatkozhat egyrészt egy-egy adott szóalakra (l. az említett *elh[ij]vadás*-t, valamint számos újonnan átvett kölcsönzót, pl. *stream* [sztrím?], *gender* [dzsender?]). Másrészt egész helyesírási területek, például azonos fogalomkörü csoportokhoz, kategóriához tartozó szavak helyesírása is „lefedetlen” lehet a kodifikáció szempontjából. Például nincs szabályozva az olyan újabb jelenségek körök

³ Nem minden esetben kérdéses azonban a *humán* egybe- vagy különírása az utána következő szóval. A 2.2 fejezet címében például egyértelműen különírandó a *humán*, mivel egy minőségjelzős szerkezet újabb minőségjelzője

helyesírása, mint a tanösvények, a számítógépes keretrendszerek stb. Ezekben az esetekben a helyesírási probléma gépi eszközökkel történő megoldása, vagyis a norma közvetítése azért nem lehetséges, mert – nincs mit közvetíteni. Emellett azonban a nyelvtechnológiai eszközök abban a tekintetben közvetve mégis a segítségünkre lehetnek, hogy a nagyméretű korpuszok, illetve a korpuszadatok kinyerésére szolgáló korpuszlekérdező eszközök (Sass, 2022) segítségével feltérképezhető az írásgyakorlat, így a közönségszolgálathoz (vagy akár egy majdan létrehozandó intelligens helyesírási segédeszközhöz) forduló kérdezőnek valós nyelvi adatokon alapuló helyesírási tanácsot lehet adni. Természetesen ez nem jelenti azt, hogy a kodifikáció által le nem fedett esetekben egyértelműen a kialakult írásgyakorlat szolgálja követése jelenti a megoldást, de a helyesírási problémák egy részénél, különösen az idegen nyelvekből kölcsönzött neologizmusok idegenes és magyaros írásmódja közti választás esetében hatékony segítség, sőt döntő jelentőségű lehet az úzus vizsgálata.

4. ÖSSZEGRÉS

A magyar beszélőközösségben élő helyesírási problémák gépi kezelhetőség szempontjából (is) történő csoportosítását az 1. ábra foglalja össze.



1. ábra. A helyesírási problémák osztályozása

Jól látható, hogy a helyesírási problémáknak csak egy igen kis szelete kezelhető gépi eszközök segítségével. Az ebbe a szűk halmazba tartozó helyesírási problémák esetén valóban igen gyors és hatékony megoldást jelentenek a nyelvtechnológiai eszközök, ezt egyértelműen mutatják is a Kutatóközpont közönségszolgálatához érkező, a *Helyesiras.mta.hu* portállal kapcsolatos pozitív visszajelzések.

A helyesírási tanácsadás számítógépes nyelvészeti eszközökkel történő támogatása tehát mindenképpen fontos és szükséges, de önmagában nem elégséges megoldás a létező helyesírási problémák kezelésére. A jelen tanulmány is amellettt érvel, hogy elengedhetetlen a gépesített helyesírási tanácsadás mellett annak hagyományos, „humán” formájának működtetése is. Nem csupán azért, mert a gépesített helyesírási tanácsadás a helyesírási problémáknak csak a kodifikált írásmóddal/szabályokkal rendelkező, algoritmusokkal leírható és/vagy kivétellistákkal kezelhető eseteiben képes segítséget nyújtani. A helyesírási problémák ezen részalmazának ugyanis növelhető a mérete: a neurális hálók, a mesterséges intelligencia nyelvtechnológiai alkalmazásai minden bizonnyal jelentősen ki fogják bővíteni az automatikusan kezelhető helyesírási problémák körét. Továbbá egy mesterséges intelligencián alapuló csetbotalkalmazás várhatóan a felhasználókat is segíteni fogja akár a *Helyesiras.mta.hu* megfelelő eszközének használatában, akár abban, hogy a jelentéstől, a kontextustól függően ajánljon írásmódot a hozzá forduló felhasználónak. Annak azonban, hogy a kérdező az ellentmondásos, illetve szabályozatlan helyesírási esetekben is rugalmas, az ő (és gyakorlatközössége) nézőpontját, helyzetét figyelembe vevő választ kapjon, a tisztán gépi eszközökkel történő megvalósíthatósága erősen kétséges. A helyesírási tanácsadó portál jelenlegi és majdani megvalósítandó 2.0-s verziójának is alapvető feladata ugyanis a helyesírási norma közvetítése, így legyen bármilyen korszerű nyelvtechnológiai alapokra helyezve, akkor sem tud teret nyitni az akadémiai helyesírási szabályozást rugalmasabban, az egyes gyakorlatközösségek igényeihez igazítani szándékozó törekvéseknek.

Egy korszerű helyesírási tanácsadó segédeszköz fejlesztésekor azonban mindenképpen szükséges tisztázni olyan elvi kérdéseket, mint hogy:

- Felülbírálnak-e egy helyesírási tanácsadó program a mindenkor kodifikált helyesírási normát?
- Helyesírási kodifikáció hiányában hozhat-e egy helyesírási-ellenőrző eszköz kodifikációs döntést? (A kérdés tulajdonképpen költői, hiszen a program válasza nagy valószínűséggel kodifikációs döntésként, „a helyes írásmódként” fog értelmeződni a beszélők körében, még ha nem is ez a fejlesztők szándéka.)

Mindemellett pedig nem lehet figyelmen kívül hagyni azt a tényt sem, hogy a nyelvhasználók is – még ha el is fogadják a gépi eszközök helyesírási javaslatait – alapvetően jobban bíznak a „humán” nyelvi tanácsadásban. Ahogy ezt az egyik levélíró is megfogalmazta a közönségszolgálat munkatársai által készített kérdőívben (Ludányi et al., 2022: 81–83): „Jó volt igazi emberektől, szakértőktől kapni direkt, egyenes és érthető választ”.

Ezekkel a tanulságokkal zárom Prószéky Gábornak szánt írásomat, akinek ezúton, 70. születésnapján köszöntve mondok köszönetet a nyelvi tanácsadás, a kutatóközpontbeli közönségszolgálat ügyeinek mindenkorai támogatásáért.

Irodalom

- Fóris Á. (2020). *Lexikológiai és lexikográfiai ismeretek magyar (magyar mint idegen nyelv) tanároknak*. Budapest: L'Harmattan Kiadó, Károli Gáspár Református Egyetem.
- Jernudd, B. H., & Neustupný, J. V. (1987). Language planning: for whom? In Laforge, L. (Ed.), *Proceedings of the International Colloquium on Language Planning* (pp. 69–84). Québec: Les Press de L'Université Laval.
- Lanstyák I. (2018). *Nyelvalakítás és nyelvi problémák*. Somorja: Fórum Kisebbségkutató Intézet, Gramma Nyelvi Iroda.
- Ludányi Zs. (2023a). A helyesírási problémák kezelési stratégiái a nyelvi tanácsadásban. *Alkalmazott Nyelvészeti Közlemények*, 16(2) (pp. 29–46).
- Ludányi Zs. (2023b). *A nyelvi tanácsadási nemzetközi és hazai gyakorlatai, különös tekintettel a Nyelvtudományi Kutatóközpontra*. Habilitációs értekezés. Budapest: ELTE Nyelvtudományi Doktori Iskola.
- Ludányi Zs. (2023c). A nyelvi tanácsadó szolgálat postaládájából 14. *Amega*, 30(3) (pp. 36–37).
- Ludányi Zs., Miháltz M. & Hussami P. (2013). Különírás-egybeírás – automatikusan. In Váradi T. (szerk.), *AlkNyelv-Dok7. Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből* (pp. 116–130). Budapest: MTA Nyelvtudományi Intézet.
- Ludányi Zs., Domonkos Á., Kocsis Á., & Jakab D. (2022). A nyelvi menedzselés szemlélete és a nyelvi tanácsadás. In Deme A. & Kuna Á. (szerk.), *Tanulmányok a nyelvészet alkalmazásainak területéről* (pp. 73–107). Budapest: ELTE Eötvös Kiadó.
- Miháltz M., Hussami P., Ludányi Zs., Mittelholcz I., Nagy Á., Oravec Cs., Pintér T., & Takács D. (2013). Helyesírás.hu – Nyelvtechnológiai megoldások automatikus helyesírási tanácsadó rendszerben. In Tanács A. & Vincze V. (szerk.), *MSZNY 2013. Magyar Számítógépes Nyelvészeti Konferencia* (pp. 135–147). Szeged: Szegedi Tudományegyetem.
- Prószéky G., & Kis B. (1999). *Számítógéppel emberi nyelven. Intelligens szövegkezelés számítógéppel*. Bicske: SZAK Kiadó.
- Prószéky G. (1993). Nyelvművelés számítógéppel? (A számítógépes helyesírás-ellenőrzés új útjai). *Magyar Nyelvőr*, 117(4), 509–511.
- Prószéky G. (1998). Nyelvhelyesség és számítógépes szövegírás. *Magyar Tudomány*, 105 (Új folyam 43)(4), 491–493.
- Prószéky G. (2019). Nyelv, nyelvhasználat és a magyar helyesírás szabályainak 12. kiadása. In Bozsik G. & Ludányi Zs. (szerk.), *Szabályzat, oktatás, gyakorlat. Helyesírásról sokszínűen. A 2015-ös és a 2017-es Nagy J. Béla helyesírási verseny előadásai, feladatai és egyéb tanulmányok* (pp. 265–571). Eger: Linceum Kiadó.
- Prószéky G., & Váradi T. (2023). A magyar nyelv digitális fenntarthatóságának támogatása. *Magyar Nyelv*, 119(4), 482–489. <https://doi.org/10.18349/MagyarNyelv.2023.4.482>
- Sass B. (2009). „Mazsola” – eszköz a magyar igék bővítményszerkezetének vizsgálatára. In Váradi T. (szerk.), *Válogatás az I. Alkalmazott Nyelvészeti Doktoranduszkonferencia előadásaiból* (pp. 117–129). Budapest: MTA Nyelvtudományi Intézet.
- Sass, B. (2022). Principles of corpus querying: A discussion note. *Acta Linguistica Academica*, 69(4), 599–614. <https://doi.org/10.1556/2062.2022.00581>
- Tóth E. (szerk.). (2017). *Magyar helyesírási szótár*. Budapest: Akadémiai Kiadó.