

**Comenius Egyetem Bölcsészettudományi Karának  
Magyar Nyelv és Irodalom Tanszéke**

**Filozofická fakulta Univerzity Komenského  
Katedra maďarského jazyka a literatúry**



**Lanstyák István**

# **Magyar nyelvi váltakozások vizsgálata a Magyar nemzeti szövegtár segítségével**

Szenci Molnár Albert Egyesület  
Pozsony  
2023

**Magyar nyelvi váltakozások vizsgálata**  
a Magyar nemzeti szövegtár segítségével

**Výskum variantnosti v maďarskom jazyku**  
pomocou Maďarského národného korpusu

*Lektorálta:*

Horváth Péter Iván (Pázmány Péter Katolikus Egyetem, Budapest)

M. Pintér Tibor (Károli Gáspár Református Egyetem, Budapest)

*Recenzenti:*

Doc. Tibor M. Pintér, PhD. (Reformovaná univerzita Gašpara Károliho,  
Budapešť)

Mgr. Péter Iván Horváth, PhD. (Katolícka univerzita Petra Pazmáňa, Budapešť)

Kiadja a Szenci Molnár Albert Egyesület, Pozsony, 2023.

Vydalo Združenie Alberta Szenci Molnára, Bratislava, 2023.

ISBN 978-80-69002-09-8

EAN 9788069002098

© Lanstyák István

# Tartalom

ELŐSZÓ .....	5
ÁLTALÁNOS TUDNIVALÓK .....	7
1. Korpusznyelvészet .....	7
2. Magyar nemzeti szövegtár .....	10
KERESÉSI BEÁLLÍTÁSOK.....	14
1. A lekérdezés típusának beállítása.....	15
A) Egyszerű keresés .....	15
B) Szótőre való keresés .....	17
C) Szóalakra való keresés .....	18
D) Karaktorsorozatra való keresés.....	22
E) CQL-keresés .....	23
F) Részletes keresés .....	25
2. Egyéb keresési beállítások.....	26
MEGJELENÍTÉSI BEÁLLÍTÁSOK .....	28
1. Mentés.....	28
2. KWIC/mondat .....	32
3. Megjelenítés .....	36
4. Rendezés.....	38
5. Véletlen minta .....	43
6. Szűrés.....	43
7. Gyakoriságok .....	44
8. Kollokációk.....	45
A KONTEXTUS MEGADÁSA .....	48
AZ ALKORPUSZOK JELLEMZŐI .....	55
1. Régió szerinti alkorpuszok .....	58
2. Stílusréteg szerinti alkorpuszok.....	60
A) A „szépirodalmi” alkorpusz sajátosságai .....	61
B) A „sajtónyelvi” alkorpusz sajátosságai .....	63
C) A „tudományos” alkorpusz sajátosságai .....	64
D) A „hivatalos” alkorpusz sajátosságai.....	66
E) A „beszélt nyelvi” alkorpusz sajátosságai .....	67
F) A két „személyes” alkorpusz sajátosságai .....	69
a) A „személyes-fórum” alkorpusz sajátosságai .....	69
b) A „személyes-közösségi” alkorpusz sajátosságai.....	71
RÉGIÓ SZERINTI MENNYISÉGI VIZSGÁLAT .....	72
1. Általános kérdések.....	72
2. A <i>nyílt nap</i> vs. <i>nyitott nap</i> példája .....	73

STÍLUSRÉTEG SZERINTI MINŐSÉGI VIZSGÁLAT .....	82
1. Általános kérdések.....	82
2. A <i>szok</i> + főnévi igenév példája .....	82
A) A „szépirodalmi” alkorpusz találatainak vizsgálata .....	85
B) A „sajtónyelvi” alkorpusz találatainak vizsgálata .....	90
C) A „tudományos” alkorpusz találatainak vizsgálata .....	91
D) A „hivatalos” alkorpusz találatainak vizsgálata .....	93
E) A „beszélt nyelvi” alkorpusz találatainak vizsgálata .....	93
F) A „személyes” alkorpusz találatainak vizsgálata .....	96
G) Összefoglalás .....	100
 STÍLUSRÉTEG SZERINTI MENNYISÉGI VIZSGÁLAT .....	 102
1. Általános kérdések.....	102
2. A <i>kaszárnya</i> és a <i>laktanya</i> példája.....	103
A) Az előfordulás relatív gyakorisága az egyes alkorpuszokban.....	104
B) A két szó gyakorisági aránya egy-egy alkorpuszon belül.....	107
 SZAKIRODALOM .....	 110
 A SOROZAT KÖTETEI .....	 123

## & ELŐSZÓ

Ez a jegyzet kiegészítésként szolgál egy többrészes jegyzetsorozathoz, amely a Comenius Egyetem Bölcsészettudományi Kara Magyar Nyelv és Irodalom Tanszékének kiadványszerkesztő szakos hallgatói számára készül. A sorozat legtöbb kötete a *Nyelvalakítás* (FiF.KMJL/A-boED-119/15 – *Jazykový manažment*) című tantárgy elméleti tudnivalóit tartalmazza. Jelen kötet célja, hogy a hallgatók megtanulják alapszinten használni a Magyar nemzeti szövegtárat, annak v2.0.5. sz. verzióját, vagyis képesek legyenek kinyerni, elemezni és értelmezni a nyelvi változók egyes változataira vonatkozó szövegtári adatokat, s az adatok ismeretében ki tudják választani a rendelkezésükre álló változatok közül az adott szövegbe legjobban beleillőt, ill. kommunikációs céljaiknak leginkább megfelelőt.

A jegyzetsorozatnak ez a darabja a *Nyelvalakítás* című tantárgyon kívül több más kurzushoz is felhasználható. Ilyen mindenképp a *Magyar nyelvváltozatok* (FiF.KMJL/A-buMA-211/20 – *Jazykové variety v maďarčine*), a *Szociolingvisztika* (FiF.KMJL/A-boED-220/15 – *Sociolingvistika*), a *Fejezetek az alkalmazott nyelvészetből* (FiF.KMJL/A-boED-223/15 – *Vybrané kapitoly z aplikovanej jazykovedy*), a *Beszélt nyelv* (FiF.KMJL/A-boED-135/15 – *Hovorený jazyk*), valamint a *Szókészlet- és szójelentéstan* (FiF.KMJL/A-boED-021/15 – *Lexikológia a sémantika*).

A kötet érdemi része nyolc fejezetből áll. Az első általános tudnivalókat tartalmaz. A következő három részletes gyakorlati útmutatás a szövegtár alapszintű használatához. Az ezt követő fejezetek az alkorpuszok szerinti vizsgálatokhoz szükséges legfontosabb tudnivalókat tartalmazzák, továbbá konkrét példákat hoznak az alkorpuszok szerinti vizsgálatra. A jegyzetet irodalomjegyzék zárja, amely a korpusznyelvészetről, ill. általánosabban a nyelvtechnológiáról szóló munkák mellett olyan tanulmányokat tartalmaz, melyeknek szerzői a kutatásaikhoz felhasználták a Magyar nemzeti szövegtárat vagy más korpuszokat.

Mivel a jegyzetsorozat megjelentetésére – időhiány és más okok miatt – nem adtunk be pályázatot, nincs aki fedezze a sorozat megjelentetésének költségeit. Ezért a kötetet magam tördeltem be, idő és tördelési gyakorlat híján a lehető legegyszerűbb módon, a szövegben való tájékozódást megkönnyítő fejléc és navigáció alkalmazása nélkül. A navigációt részben pótolja a fejezetcímek előtt álló & jel, valamint az alfejezetcímek előtt található ÷ jel: ezekre rákeresve fejezetről fejezetre, ill. alfejezetről alfejezetre tudunk ugrani a szövegben.

Köszönettel tartozom azoknak, akik a sorozat megjelentetését lehetővé tették: lektoraimnak, M. Pintér Tibornak és Horváth Péter Ivánnak, továbbá a Magyar Tudományos Akadémiának, amely Domus ösztöndíjprogramja keretében támogatta a témával kapcsolatos kutatásaimat, részben közvetlenül, részben a somorjai Fórum Kisebbségkutató Intézetben működő Gramma Nyelvi Iroda közvetítésével. A nyelvvalakítás némely általánosabb vonatkozásával a Comenius Egyetem Magyar Nyelv és Irodalom Tanszékén foglalkoztam a Szlovák Kutatás- és Fejlesztéstámogató Ügynökség APVV-17-0254 számú projektjének keretében. Ezenkívül munkámban felhasználtam annak a kutatásnak némely, a váltakozással kapcsolatos eredményét, melyet az 1/0106/21 számú VEGA-projekt keretében folytattam.

Köszönet illeti a munkahelyemet is, a Comenius Egyetem Bölcsészettudományi Karának mindenkori vezetését, valamint a Magyar Nyelv és Irodalom Tanszék mindenkori tanszékvezető-

jét, akik évtizedekig nyugodt körülményeket biztosítottak nekem a kutatási (és persze az oktatási) munkámhoz. Nem ők tehetnek róla, hogy a minisztérium abnormális követelményeket támasztó minősítési rendszere az elmúlt években egyre inkább lehetetlenné tette az elmélyült kutatást és azt, hogy igazán kiérlelt munkákat adjak ki a kezemből.

# & ÁLTALÁNOS TUDNIVALÓK

## ÷ 1. Korpusznyelvészet

A korpusznyelvészet a nyelvi jelenségek tanulmányozásának az a módja, amely **nagy mennyiségű, előre meghatározott szempontok** (pl. keletkezési idő, dialektus, regiszter, médium) **alapján összeválogatott szöveg számítógépes elemzésén** alapul. Némelyek a korpusznyelvészetet önálló tudományágnak vagy elméletnek tekintik, nem csupán az alkalmazott nyelvészet egyik területének.

Bár a tudatosan összeválogatott szövegek statisztikai feldolgozására már a múlt század 60-as éveitől is vannak példáink, a (szöveg)korpusz, ill. a korpusznyelvészet fogalma csak a 80-as években jelent meg.

A korpusznyelvészet kibontakozása mind a nyelv-, mind a fordítástudományban megerősítette az **empirikus megközelítést**, a **használatközpontúságot**, amelyet a nyelvtudományban korábban legmarkánsabban a **dialektológia**, majd a **szociolingvisztika** képviselt. A szövegek korpuszok ugyanis javarészt a tényleges nyelvhasználatot rögzítő szövegeket tartalmaznak (sőt a *javarészt* határozószó el is hagyható, ha „tényleges nyelvhasználatnak” tekintjük például a játékból tudatosan hibás vagy másokat karikírozó nyelvi megnyilvánulásokat is).

A korpusznyelvészeken kívül, akiknek a korpusznyelvészet a fő érdeklődési területük, gyakorlatilag **mindenfajta nyelvész** számára hasznos lehet olykor korpuszadatokhoz fordulni. A korpuszalapú kutatások legnyilvánvalóbb és leglátványosabb haszna a **szótárírás** és általában a **nyelvreírás**, valamint a **fordítás** számára van, legalábbis ha azokra a területekre gondolunk, melyekkel a laikus beszélők is találkoznak.

Mivel a szövegek korpuszok nem csekély hányada **bárki számára elérhető** az interneten, ezek a nyelvészeken kívül gyakorlati segítséget is nyújthatnak a fogalmazásban vagy az idegennyelvtanulásban, -tanításban a **nyelv hivatásos használóinak** (pl. az íróknak, szerkesztőknek, fordítóknak, magyar- és idegennyelvtanároknak), de a **hétköznapi beszélőknek** is

## Szövegek korpusz

A **KORPUSZ** vagy **SZÖVEGKORPUSZ** a szó modern értelmében *elektronikusan* tárolt, egy adott nyelvre nézve többé-kevésbé *reprezentatív*nak tekinthető írott vagy lejegyzett beszélt nyelvi *annotált és gépileg elemezhető szövegek*, általában nagyméretű gyűjteménye.

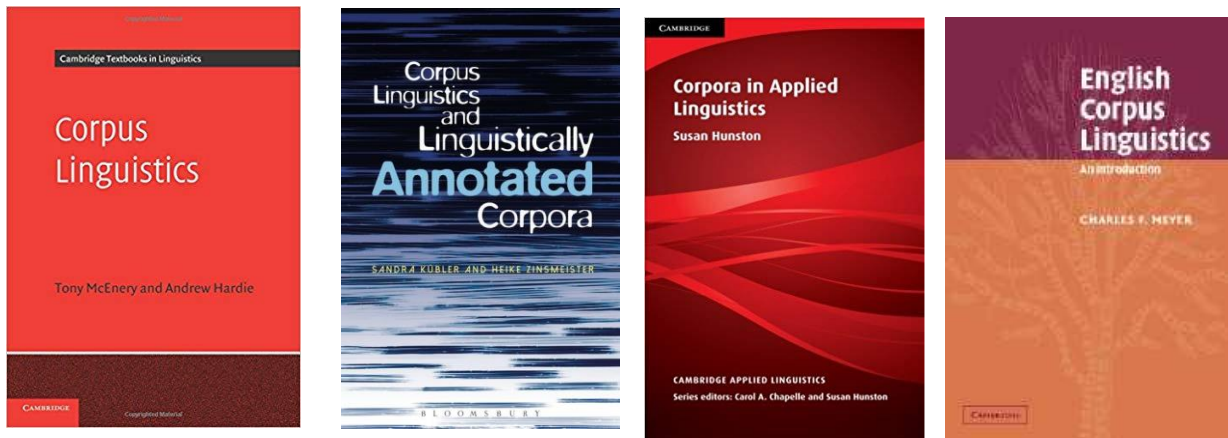
A nagyméretűség mindenekelőtt a nyilvánosan elérhető, nagy szövegek korpuszokra jellemző; speciális kutatási célokra kisebb terjedelmű korpuszokat is építenek.

A szövegek **annotáltsága** azt jelenti, hogy el vannak látva bibliográfiai, szövegszerkezeti és egyéb adatokkal.

A szövegek **gépi elemezhetősége** azt jelenti, hogy az egyes szavak és nagyobb szerkezeti egységek mellett is kódok találhatóak, melyek lehetővé teszik, hogy a program a szavakat morfémákra tudja bontani.

A legújabb szövegtörzsek **sok százmillió, sőt több milliárd szövegszót** is tartalmazhatnak.

Ha a szövegtörzsek fogalmát tudottnak tekintjük, a törzsenyelvészet fogalmát egyszerűen úgy határozzuk meg, hogy az nem más, mint **a nyelvi adatok szövegtörzsekben való tudományos vizsgálata**.



A modern szövegtörzseknek ezek a legfontosabb ismérvei, bővebben kifejtve:

- 1) A szövegtörzsek a „**való életből**” vett írott vagy (lejegyzett) **beszélte nyelvi** szövegek gyűjteménye.
- 2) A különböző forrásokból származó, különböző témájú szövegek **meghatározott szempontok** szerint vannak **összeválogatva** és **összerendezve**, s az összeállítók szándéka szerint valamilyen mértékben **reprezentálják a nyelv egészét**.

Speciális célú törzsek esetében a szövegek az adott kutatási témához kapcsolódó szövegek körét fedik le.

- 3) A szövegtörzsek célja, hogy **nagy mennyiségű empirikus adatot** szolgáltatson a **nyelvészeti kutatásokhoz**.

A nagy mennyiség, amint említettük, manapság már több milliárd is lehet. A legnagyobb magyar szövegtörzsek, a Magyar nemzeti szövegtár, melyről ez a jegyzet szól, jelenleg (2023-ban) mintegy egymilliárd szövegszót tartalmaz.

- 4) A szövegek a szövegtörzsekben **elektronikus adatbázisként** vannak tárolva, gépileg olvasható és automatikusan vagy félautomatikusan elemezhető (tehát nem képi) formában.

- 5) A szövegek ún. **KONKORDANCIAPROGRAM** segítségével válnak elemezhetővé. A program képes a szavakat alaktanilag elemezni (**lemmatizálni**), azaz megkülönböztetni a **szótöveget** és a különféle **toldalékokat**.



**6)** A korpuszban a szövegek, valamint azok mikroszintű egységei, a szavak számos **információval** vannak ellátva, amelyek megkönnyítik a szövegek nyelvészeti feldolgozását.

A **bibliográfiai adatokon** kívül kódolva vannak a szöveg **szerkezeti egységei** (mondatok, bekezdések), az egyes szavak **szófaji értéke**.

Mindazokat az információkat, amelyeket a korpuszba bekerült szöveg eredetileg nem tartalmazott, **KORPUSZANNOTÁCIÓNAK** nevezzük.

Erről Sass Bálint (2016: 799) így ír:

A szövegek nyelvi adatolást – ún. annotációt – kell, hogy tartalmazzanak. Az annotáció azt jelenti, hogy a szöveg egyes egységeihez különféle adatok vannak rendelve. Például a dokumentumokhoz a szerző vagy a szavakhoz a szófaj, de a bekezdésekhez, mondatokhoz is társíthatók adatok, mint például az adott egység nyelve vagy mérete. Az annotációk általában valamilyen egységes kódrendszer szerint szerepelnek a korpuszokban.

Példa: A körülültük szóalak morfológiai annotációja, azaz ami a szó alakjának, a benne lévő elemeket írja le a következő lehet: IK.IGE.TMt1. Ebből kiderül, hogy ez a szóalak tehát egy igekötős (IK) ige (IGE), mely határozott ragozású (T) – vö: körülültük az asztalt és nem egy asztalt – , múlt idejű (M), és többes szám első személyű (t1).

Az annotációk kézi, gépi (automatikus) vagy félautomatikus (gépi annotálás + kézi ellenőrzés) úton kerülnek bele a korpuszokba. Számos nyelvre számos automatikus szótövező, morfológiai elemző, szintaktikai elemző stb. eszköz létezik. Nagyobb szövegeknél és megbízható gépi eszközök esetén kap teret az automatikus gépi annotálás. A korpusz fogalmát a fentiek alapján tovább szűkíthetjük az annotált szövegekre.

**7)** A szövegtörzsek a felhasználó számára a **korpuszlekérdező felület** vagy **szövegkereső felület** segítségével érhetők el. Sass Bálint (2016: 798) így határozza meg a „korpuszlekérdező” fogalmát:

Olyan számítógépes eszköz, gyakran online felület, melynek segítségével szövegekben, általában nagy méretű szöveggyűjteményekben kereshetünk adott tulajdonságú szavakat, szókapcsolatokat, és a keresési feltételeket különféle nyelvészeti szempontok, a szavak nyelvészeti tulajdonságai alapján adhatjuk meg.

A közismert internetes keresők (pl. a Google) és a korpuszlekérdezők közti különbségre jól világít a következő részlet, melyben Sass Bálint (2016: 798) Adam Kilgarriff tanulmányának egy részét ismerteti:

Az internetes keresőkben (1) nincs nyelvészeti adatolás, mint például az egyes szavakhoz a szótő vagy a szófaj; (2) a keresőkifejezést csak nagyon behatárolt módon lehet megadni; (3) korlátozva van a naponta futtatható lekérdezések száma (ez akkor probléma, ha automatikusan szeretnénk futtatni sok, akár több ezer lekérdezést) és nem kapjuk meg az összes találatot; valamint (4) egy találat egy dokumentumot (internetes oldalt) jelent, nem pedig egy szóelőfordulást.

Az internetes keresők tehát nem nyelvészeti szövegkeresők, a nyelvészeti adatolás hiánya, illetve az egyszerű lekérdezési formátum miatt nem tudunk nyelvészetileg releváns lekérdezéseket megfogalmazni. Az internetes keresők ún. információ-visszakereső (information retrieval – IR) rendszerek, feladatuk az, hogy azt a dokumentumot adják eredményül, mely a lekérdezésnek megfelelő releváns információt tartalmazza. A nyelvészeti szövegkeresők feladata ezzel szemben az, hogy adott, precízen körülírt nyelvi jelenség összes előfordulását szolgáltatassák, így

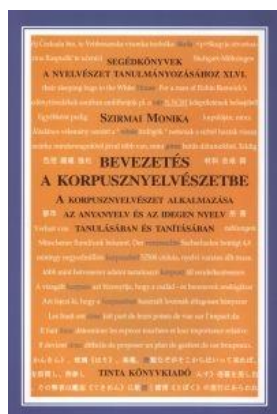
nemcsak példákat kapunk a jelenségre, hanem lehetővé válik a jelenség statisztikai vizsgálata is.

Amint az előzőekből is kiderül, a nyelvészeti szövegkeresőkben – szemben a Google-féle internetes keresőkkel – a keresési alapegység nem a dokumentum, hanem **leggyakrabban a szó**, de sok korpuszban **szónál nagyobb vagy kisebb egységre** is kereshetünk, pl. **mondatra, tagmondatra** egyfelől, **fonémára** másfelől.

Szónak az elemzőprogramok az egymástól szóközzel elválasztott vagy írásjellel lezárt betűsorokat tekintik.

A keresés tehát valamilyen **nyelvészeti szempontból meghatározható egységre** irányul. Ismét Sass Bálintot idézzük:

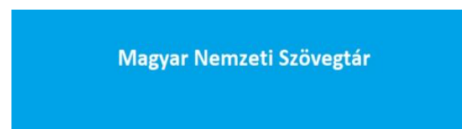
A filológiával, irodalomtudománnyal szemben, mely talán nagyobb jelentőséget tulajdonít a dokumentumokhoz rendelt, dokumentumszintű adatoknak (szerző, forrás, megjelenés ideje stb.), a nyelvészeti szövegkeresőkben a szó a központi elem. Egyrészt általában a szavak kapják a legtöbb fajta és legrészletesebb annotációt, másrészt a találatok sem dokumentumok, hanem szavak. Másképp fogalmazva: egy nyelvészeti szövegkereső esetében alapvető követelmény, hogy a keresett szó minden egyes előfordulását külön találatként jelenítse meg.



## ÷ 2. Magyar nemzeti szövegtár

A Magyar Nemzeti Szövegtár bővített és átalakított változata, az MNSZ2 – éppúgy, mint az eredetibb, kisebb terjedelmű és könnyebben kezelhető, gondosabban összeállított MNSZ1 – a **Nyelvtudományi Kutatóközpont** honlapján található (<http://www.nytud.hu/>).

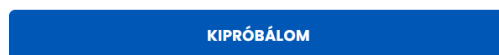
A Nyelvtudományi Kutatóközpont a 2021-ben megszüntetett Nyelvtudományi Intézet jogutódja. A Nyelvtudományi Intézet 1949-ben jött létre, 1951-től a Magyar Tudományos Akadémia egyik szervezeti egységeként működött egészen 2019-ig, amikor egy újonnan



SZÖVEGKORPUSZ

Magyar Nemzeti Szövegtár

BŐVEBBEN →



létrehozott intézményhez, az Eötvös Loránd Kutatási Hálózathoz kapcsolták; ennek neve 2023-ban Magyar Kutatási Hálózat lett; rövidítésként az angol név alapján keletkezett HUN-REN (Hungarian Research Network) betűszót használják.

A Nyelvtudományi Kutatóközpont honlapjáról a Magyar nemzeti szövegtárat – pontosabban ennek korábbi változatát – úgy érzük el, hogy az oldal fejlécében található menüsor „**Szolgáltatások**” menüpontját választva az „**Adatforrások**” linkjére (<https://nytud.hu/adatforrasok>) kattintunk, melynek felületén – sok más hasznos adatbázis, ill. adatforrás mellett – megtaláljuk a Magyar nemzeti szövegtárat is.

A **világoskék** felületre kattintva **tájékoztatót** olvashatunk a Magyar nemzeti szövegtárról, a **sötétkék** felületre kattintva pedig magának a **szövegtárnak** – pontosabban az MNSZ1-nek – az oldalát érzük el.

Az MNSZ1 oldalán (<http://corpus.nytud.hu/mnsz/>), annak elején, a bekeretezett részben olvashatunk az MNSZ2-ről, onnan a „**Kattintson**” szóra kattintva jutunk el az MNSZ2 oldalára, ahonnan (regisztrálás után) a szövegkiemelt „**Keresőfelület**” szóra kattintva lehetővé válik a keresés.

Hasonlóképpen érhető el az MNSZ2 a jobboldali menüsor „Keresés” menüpontjára kattintással, ahonnan egy másik oldalon keresztül érzük el az MNSZ2-t.

A Nyelvtudományi Kutatóközpont honlapjának főntebb említett „**Adatforrások**” oldaláról (<https://nytud.hu/adatforrasok>) nyitható meg az ún. **Nemzeti korpuszportál** is, melyről elérhetők a legfontosabb magyar szövegkorpuszok.

A Nemzeti korpuszportál oldalának felső része így néz ki:

Szeretnénk, ha minél több magyar nyelvű, szóalapú online keresővel rendelkező korpusz összegyűlne itt a köz javára.

**Az Ön által karbantartott korpusz is!**  
**Kérjük írjon nekünk, hogy felvehessük a portálra.**

A csatlakozás feltételei:

- magyar nyelvű szöveganyag
- szóalapú online kereső/lekérdező
- kapcsolat = egy emailcím, akihez lehet fordulni

Ezenkívül az a kérésünk, hogy tegyen **egy linket** a korpusz oldalára, mely az NKP-re mutat, például így:

A [...] a [Nemzeti Korpuszportál tagja](#).

Ezzel hozzájárul ahhoz, hogy a korpuszhasználók megismerhessék az NKP tagjait! Köszönjük.

## Korpuszok

2016	„új” Magyar történeti szövegtár	<a href="#">kereső</a>	<a href="#">kapcsolat</a>
2015	Mikes-szótár	<a href="#">főoldal</a> <a href="#">kereső</a>	<a href="#">kapcsolat</a>
2014	MNSZ2	<a href="#">főoldal</a> <a href="#">kereső</a>	<a href="#">regisztráció</a> <a href="#">kapcsolat</a>
2014	Történeti Magánéleti Korpusz	<a href="#">főoldal</a> <a href="#">kereső</a>	<a href="#">kapcsolat: nyelvi tech</a>
2013	Ómagyar korpusz	<a href="#">főoldal</a> <a href="#">kereső</a>	<a href="#">kapcsolat</a>
2012	BUSZI	<a href="#">főoldal</a> <a href="#">kereső</a>	<a href="#">regisztráció</a> <a href="#">kapcsolat</a>
2010	Hunglish Korpusz	<a href="#">főoldal</a> <a href="#">kereső</a>	<a href="#">kapcsolat</a>
2009	Mazsola	<a href="#">főoldal</a> <a href="#">kereső</a>	<a href="#">regisztráció</a> <a href="#">kapcsolat</a>
2006	Webkorpusz	<a href="#">főoldal</a> <a href="#">kereső</a>	<a href="#">kapcsolat</a>
2005	MNSZ1	<a href="#">főoldal</a> <a href="#">kereső</a>	<a href="#">regisztráció</a> <a href="#">kapcsolat</a>
1999	Orwell: 1984	<a href="#">kereső</a>	<a href="#">kapcsolat</a>

A „**Nemzeti korpuszportál**” oldalról az MNSZ2-nek három oldalát érhetjük el:

- a) a „**főoldal**” link használatával eljuthatunk a főoldalra;
- b) a „**kereső**” linkre kattintva elérhetjük az MNSZ2 lekérdezőfelületét;
- c) a „**regisztráció**” linkre kattintva elvégezhetjük a regisztrálást, amely feltétele mind az MNSZ2, mind az MNSZ1 használatának (egyszeri regisztrálással válik elérhetővé mindkettő, sőt a regisztráció lehetőséget ad a Mazsola igeivonzat-kereső használatára is).

A „**kapcsolat**” linkre kattintva levelet írhatunk a szövegtár működtetőinek.

Először nézzünk szét a **főoldalon** (<http://clara.nytud.hu/mnsz2-dev/>)! Ezt megnyitva ez a kép fogad bennünket:

## MNSZ2

Ezen az oldalon a Magyar Nemzeti Szövegtár több mint egymilliárd szavas, felújított változatát találja.

>> 2018. június 8. óta érhető el a legfrissebb [v2.0.5](#)-ös verzió. <<

[Hírek](#). [Keresőfelület](#). [Súgó](#). [Alkorpuszok mérete](#). [Morfoszintaktikai kódok rendszere](#).

Teljes egészében tartalmazza a [rég](#)i, [első változat](#) anyagát, de ahhoz képest számos új tulajdonsággal, előnnyel bír.

- **Nagyobb.** A szövegtár mérete jelentősen megnőtt. A nyelvi anyag frissebb, új mintavétel a mai magyar nyelvből.
- **Felújított nyelvi elemzés.** Jó minőségű nyelvi elemzést tartalmaz, mely információt nyújt többek között az összetett szavakról, a fonológiai jegyekről, a képzőkről is.
- **Új funkciók.** A találatokat [elmenthetjük](#), [szűrhetjük](#), különféle [gyakorisági listákat](#) készíthetünk, sőt [kollokációs](#) vizsgálatot is végezhetünk.
- **Szofisztikált keresés.** A [részletes keresés](#) felület és a [CQL](#) lekérdezőnyelv segítségével.
- **Összes találat.** A korábbi maximum 500-zal szemben lehetőség van az összes találat megtekintésére.
- **Gyorsabb.** A keresőfelület válaszideje lerövidült.

Ismerkedjen meg az új szövegtárral és az új [keresőfelülettel](#), mely ingyenes [regisztráció](#) után érhető el. A korábbi regisztrációk az új szövegtárra is érvényesek. A használatról részletes [súgó](#) ad tájékoztatást. Elérhetők az [alkorpuszok méretére](#) vonatkozó adatok. Az MNSZ annotációja és felülete időről időre változhat, jelentős változás esetén a [korábbi verziók](#) is elérhetők maradnak. Ha hibát vagy nem megfelelően működő funkciót talál, kérjük, [írjon nekünk](#).

Kérjük, hogy a Magyar Nemzeti Szövegtár felhasználásával létrejött műről [értesítsen bennünket](#), illetve hivatkozzon az alábbi cikkekre:

Oravecz Csaba, Váradi Tamás, Sass Bálint: [The Hungarian Gigaword Corpus](#). In: *Proceedings of LREC 2014*, 2014.

Próbálja ki az igék és argumentumaik közvetlen vizsgálatára szolgáló [Mazsola](#) eszközünket is!

*Magyar Nemzeti Szövegtár - v2.0.5*

*Ha észrevétele van, kérjük, [tudassa velünk](#).*

© *MTA Nyelvtudományi Intézet, Nyelvtechnológiai Kutatócsoport, 1998-2020.*

Az érdemi szöveg második sorában öt linket találunk, logikátlan sorrendben (a linkek sem fontossági sorrendet nem követnek, sem betűrendbe nincsenek állítva).

Az első a „**Hírek**” (<http://clara.nytud.hu/mnsz2-dev/hirek.html>); ebből a korábbi fejlesztésekről kaphatunk áttekintést. Ha most ismerkedünk az MNSZ2-vel, számunkra ezeknek az információknak nincs különösebb jelentősége.

A második, kiemelt opció, a „**Keresőfelület**” ([http://clara.nytud.hu/mnsz2-dev/bonito/run.cgi/first\\_form](http://clara.nytud.hu/mnsz2-dev/bonito/run.cgi/first_form)). Ez gyakorlati szempontból a legfontosabb, mert erre kattintva jutunk el a keresőfelületre, s válik így lehetővé a lekérdezés (amennyiben már regisztráltunk).

A harmadik választási lehetőség a „**Súgó**” (<http://clara.nytud.hu/mnsz2-dev/sugo.html>). Ez olykor ténylegesen eligazít, ha az MNSZ2 használata közben problémába ütközünk. Sajnos azonban nagyon sokszor nem kapunk belőle a kérdésünkre választ, mivel fájdalmasan kevés információt tartalmaz.

A súgó az évek során semmit sem változott, pedig a visszajelzések alapján a fejlesztőknek tudniuk kell azokról a problémákról, amelyekkel a felhasználók küszködnek, vagy legalábbis egy részükről.

Mivel az MNSZ2-n belül több alkorpuszban van lehetőségünk keresni, fontos az „**Alkorpuszok mérete**” opció is (<http://clara.nytud.hu/mnsz2-dev/stat.html>), különösen olyankor, amikor statisztikai vizsgálatot végzünk annak megállapítására, hogy egy-egy nyelvi forma használata a nyelvhasználat különféle színtereire mennyire jellemző, s ebből mire lehet következtetni a vizsgált nyelvi forma stílusértékére nézve.

A méretadatok természetesen a szóelőfordulásokra vonatkoznak; ezeket nevezik *szövegszónak*.

A táblázat a „személyes-fórum” és a „személyes-közösségi” alkorpusz méretadatait összevontan tartalmazza, „személyes” címszó alatt. Ennek az az oka, hogy ezeket az alkorpuszokat később hozták létre a korábban egységes „személyes” alkorpuszból. A két alkorpuszra vonatkozó – becsült – méretadatok a táblázat alatt olvashatók.

A sorban utolsó, „**Morfoszintaktikai kódok rendszere**” elnevezésű opcióra (<http://clara.nytud.hu/mnsz2-dev/msd.html>) akkor van szükségünk, ha például nyelvtani formákra, nem pedig konkrét szavakra szeretnénk rákeresni.

Nem tudjuk, hogy az összes lehetséges kódot tartalmazza-e, vagy csak a legfontosabbakat: neve alapján (a határozott névelő hiányából) azt várnánk, hogy csak válogatást tartalmaz, de az is lehet, hogy a határozott névelő csak tévesztésképpen maradt el a szerkezet elejéről, s valójában „A morfoszintaktikai kódok rendszeréről” van szó.

A következő sorból az MNSZ1-nek az általunk már fentebb megismert oldala (<http://corpus.nytud.hu/mnsz/>) is elérhető (ez nyílik meg, amikor a Magyar nemzeti szövegtárat a Nyelvtudományi Kutatóközpont „adatforrások” oldaláról nyitjuk meg). Az MNSZ1 összehasonlíthatatlanul felhasználóbarátabb volt, mint amilyen az MNSZ2, viszont az a mai szóanyagnak csupán mintegy egyötöd részét tartalmazza, s új funkciói is vannak, ezért az MNSZ1-et már nem érdemes használni.

További előnye az MNSZ2-nek, hogy a program (corpus manager) és a felhasználói felület (graphic user interface) nem egyedi, hanem külföldi fejlesztésű, más nemzeti korpuszokban is találkozhatni vele. Így például a Szlovák nemzeti korpusz is ezt használja.

Ezután az MNSZ2 rövid jellemzését találjuk, melyből megismerhetjük az MNSZ2 előnyeit az MNSZ1-hez képest. Az MNSZ2 hátrányaival a tényleges használat során szembesülünk majd.

A legújabb, v2.0.5-ös verzió sajátosságairól a „**Hírek**” rovatból tájékozódhatunk; fentebb már jeleztük, honnan érhető el ez a doku.

## & KERESÉSI BEÁLLÍTÁSOK

Első feladatunk az lesz, hogy elvégezzük a **beállításokat**. Ehhez a fentebb jelzett „Keresőfelület” linkjére kell kattintani ([http://clara.nytud.hu/mnsz2-dev/bonito/run.cgi/first\\_form](http://clara.nytud.hu/mnsz2-dev/bonito/run.cgi/first_form)). Ugyanezt elérhetjük, amint jeleztük, a Nemzeti korpuszportál oldaláról is (<http://corpus.nytud.hu/nkp/>).

Ha a keresőfelület angol nyelvű változata jelenne meg, a **jobb alsó sarokban** át tudunk váltani a magyar nyelvre:

Ha a „**korpusz**” mezőre kattintunk, lehetőségünk nyílik az MNSZ2 korábbi verzióinak a választására:

Egy korábbi verzió használatának csak az olyan felhasználók esetében van értelme, akik e régebbi verziók valamelyikével kezdték el a kutatást, s ugyanazzal szeretnék be is fejezni.

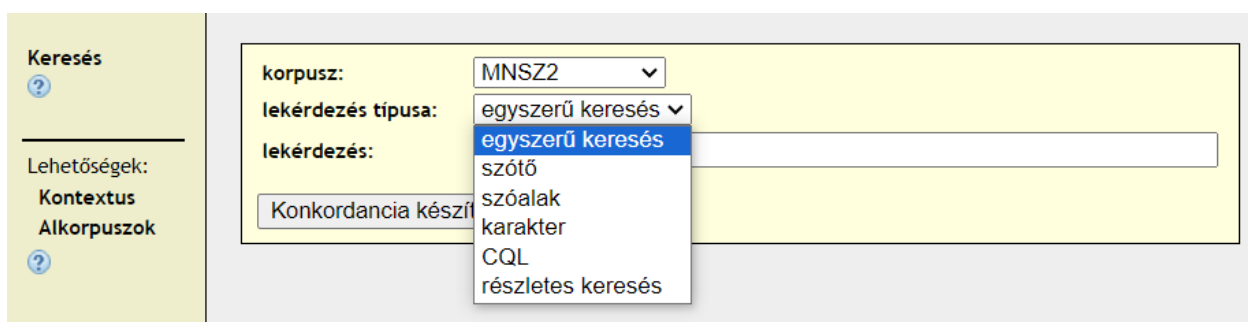
Mi a legújabb, legnagyobb terjedelmű változatot fogjuk használni, annak is a legújabb v2.0.5.-ös verzióját, az **alapbeállítás** szerint.

## ÷ 1. A lekérdezés típusának beállítása

A következő sorban a „**lekérdezés típusát**” lehet beállítani. Ezek közül a lehetőségek közül választhatunk:

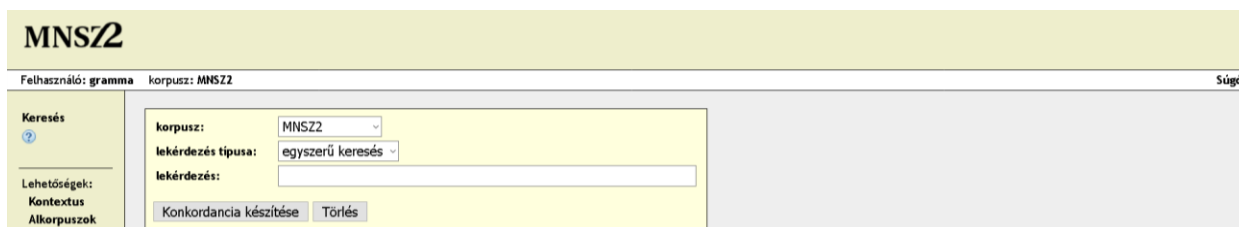
1. **egyszerű** keresés,
2. **szótőre** való keresés,
3. **szóalakra** való keresés,
4. **karakterre** való keresés,
5. a **CQL-nyelv** segítségével történő keresés,
6. **részletes** keresés.

Az, hogy melyiket választjuk, alapvetően a feladat jellegétől függ, de a választásban olykor személyes preferenciánk is szerepet játszhat (magyarán ugyanazt az eredményt többféle módon is elérhetjük, s tőlünk függ, melyiket választjuk).



Nézzük meg egyenként ezeket a lehetőségeket!

## ÷ A) Egyszerű keresés



Az „**egyszerű keresés**” opcióval a célforma **egész paradigmájára** rákereshetünk, vagyis olyankor élünk ezzel a lehetőséggel, amikor nemcsak a beírt szótári alakra mint lexémára vagyunk kíváncsiak, hanem a lexéma különféle todalékos alakjaira is.

Például ha a *posta* szót írjuk be, nemcsak a *posta* szóalakot kapjuk meg, hanem az olyanokat is, mint *postát*, *posták*, *postán*, *postánkon*, *postára*, *postáról*, *postákról*, *postákat*, *postámat*

stb., viszont nem kapunk találatot a *postás*, a *postai* vagy a *postáz* alakra, mivel az *-s*, az *-i* és a *-z* nem ragok vagy jelek, hanem képzők, és így új lexémát (új szótári egységet) alkotnak.

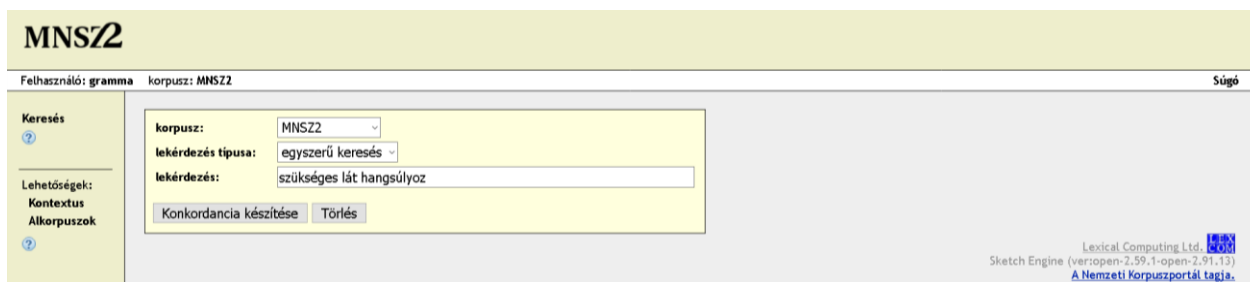
Kivételt képeznek az igék esetében az igenévképzős alakok, ezeket megkapjuk, ha az igére keresünk rá, pl. az *énekel* beírásával kijön az *éneklek*, *énekel*, *énekeltek*, *énekelik*, *énekelte*, *énekelj* stb. mellett az *énekelni*, *énekelniük*, *éneklő*, *éneklendő*, *énekelve* stb. is.

Az ilyen típusú keresést a **szótőre** való keresésnek fogjuk nevezni.

Olyankor van értelme ezzel a lehetőséggel élni, amikor a szónak nincs más szófajú homonimája (mert a szófaji értéket itt nem tudjuk beállítani), így tehát egyaránt megjelennek a *sírok* 'én sírok' és a *sírok* 'halottak nyugvóhelyei' típusú találatok.

Nagy előnye ennek az opciónak, hogy **több szóra** is rákereshetünk vele. Ha tehát két, három vagy több szóból álló **szókapcsolatok** meglétére vagyunk kíváncsiak, és nincs kedvünk részletesebb beállításokkal bibelődni, ezt az opciót érdemes választani.

A következő mezőben, a **„lekérdezésben”** tüntetjük fel a célformát, azaz a találati szót vagy szavakat. Írjuk be ide mondjuk a *szükséges*, a *lát* és a *hangsúlyoz* szavakat:



MNSZ2

Felhasználó: **gramma** korpusz: **MNSZ2** Súgó


Keresés

Lehetőségek:  
Kontextus  
Alkorpuszok

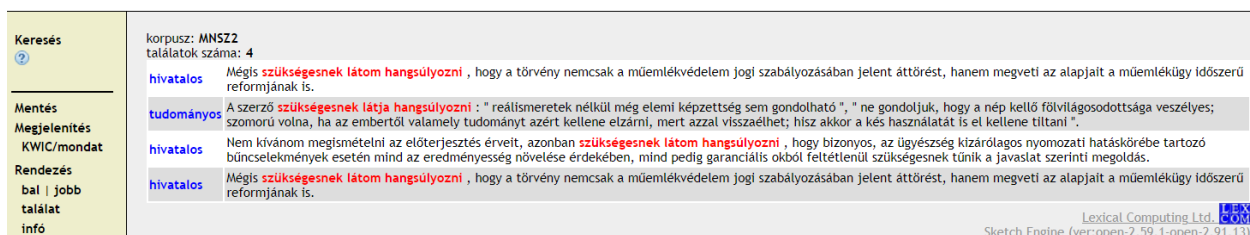
korpusz:

lekérdezés típusa:

lekérdezés:

Lexical Computing Ltd.   
Sketch Engine (ver:open-2.59.1-open-2.91.13)  
A Nemzeti Korpuszportál tagja.

A **„Konkordancia készítése”** gombra kattintva négy találatot kapunk a *szükségesnek lát hangsúlyozni* kifejezésre:



Keresés

Mentés  
Megjelenítés  
KWIC/mondat  
Rendezés  
bal | jobb  
találat  
infó


korpusz: **MNSZ2**  
találatok száma: 4

**hivatalos** Mégis **szükségesnek látom hangsúlyozni**, hogy a törvény nemcsak a műemlékvédelem jogi szabályozásában jelent áttörést, hanem megveti az alapjait a műemlékügy időszerű reformjának is.

**tudományos** A szerző **szükségesnek látja hangsúlyozni**: "reálismeretek nélkül még elemi képzettség sem gondolható", "ne gondoljuk, hogy a nép kellő fölvilágosodottsága veszélyes; szomorú volna, ha az embertől valamely tudományt azért kellene elzárni, mert azzal visszaélhet; hisz akkor a kés használatát is el kellene tiltani".

**hivatalos** Nem kívánom megismételni az előterjesztés érveit, azonban **szükségesnek látom hangsúlyozni**, hogy bizonyos, az ügyészség kizárólagos nyomozati hatáskörébe tartozó bűncselekmények esetén mind az eredményesség növelése érdekében, mind pedig garanciális okból feltétlenül szükségesnek tűnik a javaslat szerinti megoldás.

**hivatalos** Mégis **szükségesnek látom hangsúlyozni**, hogy a törvény nemcsak a műemlékvédelem jogi szabályozásában jelent áttörést, hanem megveti az alapjait a műemlékügy időszerű reformjának is.

Lexical Computing Ltd.   
Sketch Engine (ver:open-2.59.1-open-2.91.13)

Mi ugyan a szótári alakot tüntettük fel a **„lekérdezés”** mezőben, mégis látjuk, hogy todalékos alakokat kaptunk eredményként, mivel ez a három szó együttesen a megjelenített kifejezésben fordul elő.

Természetesen ez a todalék zéró morféma is lehetett volna, csak történetesen e találatok közt ilyen nincs.



## ÷ B) Szótőre való keresés

A következő lehetőség a **szótőre** való keresés: ha ide írjuk be a célformát, itt is megkapjuk a lexéma **minden toldalékos alakját**.

The screenshot shows the MNSZ2 search interface. The search criteria are: korpusz: MNSZ2, lekérdezés típusa: szótő, and szótő: (empty). There are buttons for 'Konkordancia készítése' and 'Törítés'.

A különbség az előzőhöz képest az, hogy míg az egyszerű keresésnél több szó beírásával szó szerkezeteket is kereshetünk, a szótő esetében csak **egyetlen lexéma** keresése engedélyezett.

Keressünk rá például a *kandeláber* szótőre:

The screenshot shows the MNSZ2 search interface with the search criteria: korpusz: MNSZ2, lekérdezés típusa: szótő, and szótő: kandeláber. There are buttons for 'Konkordancia készítése' and 'Törítés'.

Ezt az eredményt kapjuk:

The screenshot shows the MNSZ2 search results page for the word 'kandeláber'. The search criteria are: korpusz: MNSZ2, találatok száma: 918. The results are listed in a table with columns for 'szépirodalom' and the text of the results. The results are:

szépirodalom	szépirodalom
	Novemberben különösen vigyázni kell, mert a nyálkás hidegben a galambszar is jobban csúszik, láttam már elegáns, fekete kabátos férfiakat, akiknek olasz cipője nem bírt el a madárkakival, ott hempergöztek a koraesti szürkéségben, a sárga fényű <b>kandeláberek</b> alatt, korhely dísznők módjára, a kovácsoltvas biciklilitartóba kapaszkodva támaszkodtak volna talpra, hogy bejussanak végre a patika ajtaján.
	Szürke, egyöntetű romhalmaz volt a város, minden összeomlott, csak a gúnárnyakú <b>kandeláberek</b> álltak.
	Ő csak feküdt ott fölpattanva szája s mondta a pap kezében bibliája A halottasház ajtaja az egyik rosszul betámasztván zörögve megnyílt Volt aki ingujjban jött más megint ki feketében nyakkendővel vonult ki A gyászhuszárok szinte félcívilben: kék-szürke nadrágban s kék-szürke ingben S a kondoleálók sorjában egyre hol csak kezét szorítva hol ölelte A fejnél két vas <b>kandeláber</b> ormán spirítuszlámg égett izgága-formán S ő csak feküdt horgassá vált az orra s nyakát mintha kissé behúzta volna szokattanul hiszen tartása állva s jártában is olyan volt mint a szálfá Ki bagatellizálta a világot most ime ő is bagatellizált volt kit nem soká a föld alá bedugnak étkévé sok jóévtvágyú kukacnak de így is most is még csodák csodája...
	<b>Kandeláberek</b> lángjától?
	A szobában hőség volt, hatalmas <b>kandeláberek</b> lobogtak.
	Hábetler egy lángoló <b>kandeláber</b> mellett állt.
	Az öreg Zentay a fiával rózsá- és ciklámenkoszorút hozott, a ravatalozó emberek alig tudták elhelyezni a <b>kandeláberek</b> között illatozó virághegyen.
	A szobában hőség volt, hatalmas <b>kandeláberek</b> lobogtak.
	Én meg pontosan a lángoló <b>kandeláber</b> mellett álltam, ömlött rólam a veríték.
	Az volt a szerencsénk, hogy pórázával a komondor egy felfordult <b>kandeláber</b> rozsdás tömegébe gabalyodott, és most azt ugatta.
	A komondor akkorát rántott pórázán, hogy a <b>kandeláber</b> nagyot kalimpált.
	Mert a komondor sem ismert; és nyálát vadul csurrantva - a vállalt rabszolgaság és a szeretet jogán - visszakövetelte meleg vérünket, szagunkat, tenyerünket, bálvány mivoltunkat: a hóban maga után vonszolta a <b>kandelábert</b> .
	A kutya pórása lesiklott a <b>kandeláber</b> szarváról; közelhúzta karikáját, egészen ki a huzal végéig, s most ott ugrált, veszekedett a gránát vágta falnyílás alatt, ahol Anibel eltűnt.
	A bronzarany <b>kandelábereket</b> !
	Marioláne tétlenül üldögélt a falnál, az üvegfürtös <b>kandeláberek</b> alatt, és ölébe ejtett kézzel, révetegen bámult maga elé.
	- Illúzióinak - mondta mereven Morelli, és fölnézett a piszkossárga fényvel kigyulladó <b>kandeláberre</b> .
	Valami világos gesztus kéne: egyértelmű jelzés, tiszta fény, nem gyertyában villamosáram vagy <b>kandeláber</b> , spirítuszszagú.
	A <b>kandeláberek</b> spirítusz-szaga émelyítő; Villó az esküvői kalapját viseli, sűrű fátlyallal, feketére festve.
	Alvadtvér-színű árnyéksáv szivarog a lomb közt s cseppenként alápereg, kialusznak a gesztenyevirágok, e nappal égő <b>kandeláberek</b> .
	Waltzerek a rozgatonvirágok. S nappal égő <b>kandeláberek</b> .

Látjuk, hogy a *kandeláber* mellett itt is szerepelnek a konkordanciában a különféle toldalékos alakok is (*kandeláberek, kandelábereket, kandelábert, kandeláberre*). A találatok persze folytatódnak, hiszen 918 van belőlük, ez a kép csak az első néhányat tartalmazza.

### ÷ C) Szóalakra való keresés

A következő lehetőség a **szóalakra** való keresés. Ezt akkor érdemes használni, amikor például csak egy-egy igealakra vagyunk kíváncsiak, mondjuk az *éreződjék* felszólító módú ikes formára.

korpusz:

lekérdezés típusa:

szóalak: 
 kisbetű/nagybetű számít:

Ezt az opciót használva az *éreződik* igenek csak erre az egyetlen alakjára kapunk találatokat. A következő kép a találatok egy részét mutatja, további rendezés nélkül:

Keresés	korpusz: MNSZ2 találatok száma: 38
Mentés	<b>szépirodalom</b> A könyv bármiről szól is, <b>éreződjék</b> benne egyfajta könnyedség, mely arra emlékeztet, hogy a mű sohasem természetzerűleg adott valami, hanem igény és adomány.
Megjelenítés	<b>szépirodalom</b> Vigyáz, a hangján ne <b>éreződjék</b> szorongása.
KWIC/mondat	<b>szépirodalom</b> Az író igyekszik nyugodtnak látszani, vigyáz, hogy hangján ne <b>éreződjék</b> a szorongás.
Rendezés	<b>szépirodalom</b> A tisztelendő úr pedig azon igyekezve, véletlen se <b>éreződjék</b> hangjából a rökönnyödés.
bal   jobb	<b>szépirodalom</b> A halál tétje <b>éreződjék</b> rajta, igen, hogy az életben meg is halhatunk.
találat	<b>szépirodalom</b> De ne <b>éreződjék</b> a rögtönzés rajta, hebegés-habogás és oly biztonsággal, mintha tekercedd hosszúra húzva olvasnád!
infó	<b>szépirodalom</b> (Amerre a képzelet épp kanyarodik, de úgy, hogy <b>éreződjék</b> : szabad asszociáció ez, " imagination creatrice " ! Nő-rím szabályosan négy soronként.)
keverés	<b>hivatalos</b> S azzal is tisztában vagyok, hogy időnként az én kijelentéseim sem aratnak éppen osztatlan sikert, de nagyon kérem képviselőtársaimat, hogy ez az ellenszenv ne <b>éreződjék</b> a kárpótlás ügyén.
Véletlen minta	<b>sajtó</b> De mindig ott vagyok - mondhatnám a pálya szélén -, hogy az élmény friss hatása <b>éreződjék</b> az alkotáson.
Szűrés	<b>sajtó</b> Riporter :- Tehát a csiksomlyói helyszellem <b>éreződjék</b> ?
Gyakoriságok	<b>sajtó</b> Arra van szükség, hogy ez a valóságban is, a választási eredményeken <b>éreződjék</b> .
szóalakok	<b>sajtó</b> Demszkyvel azt akarjuk, hogy minden megnyilatkozásunkon, döntésünkön <b>éreződjék</b> ez az értékrend.
szótövek	<b>sajtó</b> Személyesebb, nyugodtabb, zajongást kerülő periódust akarok az életemben, s az is célom, hogy mindez <b>éreződjék</b> a műveimben is. Illés Endre - akinek Borsos Miklós szavaival az " izlése a legigényesebb " - művészete évtizedek óta vonz.
elemzések	<b>sajtó</b> A gyerekeket (15-16 éves csitrik) hamar megkedvelte, de azért vigyázott a hangjára, percig se lágyuljon annyira, hogy ne <b>éreződjék</b> ki belőle a tulaj parancsoló utasítása.
dokumentumok	<b>szépirodalom</b> Ennyi idő elegendő lehetett volna ahhoz, hogy legalább a bűneink többsége új legyen, szellemes, lefegyverző, <b>éreződjék</b> meg rajtuk az értelem kalandja, de a nagy részük, sajnós, unalmas, kisszerű, ismerős, bejáratott.
Kollokációk	<b>szépirodalom</b> Valami mégis akadályozza, hogy úgy kifejecesedjen, kikristályosodjon egy mag, hogy ne <b>éreződjék</b> rajta Pécs vagy Debrecen, meg Szeged vagy Budapest, szóval oda-vissza ne legyenek rossz szálak, már hangulati értelemben.
	<b>szépirodalom</b> Az első köszöntők nyikoroghatnak kissé; egyáltalán: <b>éreződjék</b> feszültség.
	<b>szépirodalom</b> Lágyan, természetesen; <b>éreződjék</b> még a tehének lehelete is.
	<b>szépirodalom</b> (Igyekszik úgy fogalmazni, hogy ne <b>éreződjék</b> a megaláztatása.)

Fontos tudni, hogy ennél az opciónál lehetőségünk van a **kis és nagy kezdőbetű megkülönböztetésére** is, ami többek között akkor hasznos, ha a célszóval azonos vezetéknev vagy nicknév nagy számban ad irreleváns találatot.

Például a *kandeláber* szó vizsgálata során kiderült, hogy ki kell szűrni a *Kandeláber* nicknevet.

Ha azt akarjuk, hogy csak kis kezdőbetűs találatokat kapjunk, kipipáljuk a „kisbetű/nagybetű számít” jelölőnégyzetet:

korpusz: 
  
 lekérdezés típusa: 
  
 szóalak: 
 kisbetű/nagybetű számít:

Írjuk most be a keresőmezőbe a *kandeláber* szót, s hagyjuk kipipálva a „kisbetű/nagybetű számít” opciót:

korpusz: 
  
 lekérdezés típusa: 
  
 szóalak: 
 kisbetű/nagybetű számít:

A „Konkordancia készítése” gombra kattintva ezek a találatok jelennek meg:

**MNSZ2**

Felhasználó: **gramma** korpusz: MNSZ2 >> v2.0.5 - friss verzió 2018. június 6-tól << [Hírek](#) [Súgó](#)

<p><b>Keresés</b></p> <p><b>Mentés</b></p> <p><b>Megjelenítés</b></p> <p><b>KWIC/mondat</b></p> <p><b>Rendezés</b></p> <p>bal   jobb</p> <p><b>találat</b></p> <p>infó</p> <p>keverés</p> <p><b>Véletlen minta</b></p> <p>Szűrés</p> <p><b>Gyakoriságok</b></p> <p>szóalakok</p> <p>szótövek</p> <p>elemzések</p> <p>dokumentumok</p> <p><b>Kollokációk</b></p> <p>Visszajelzés</p> <p>Próbálja ki a <b>Mazsola</b> eszközt ígék és argumentumaik közvetlen vizsgálatához!</p>	<p>korpusz: MNSZ2 találatok száma: 165</p> <p><b>szépirodalom</b></p> <p>Ő csak feküdt ott fölpattanva szája s mondta a pap kezében bibliája A halottasház ajtaja az egyik rosszul betámasztván zörögve megnyílt Volt aki ingujjban jött más megint ki feketében nyakkendővel vonult ki A gyászuszárok szinte félcivilben: kék-szürke nadrágban s kék-szürke ingben S a kondoleálók sorjában egyre hol csak kezét szorítva hol ölelte A fejnél két vas <b>kandeláber</b> ormán spirituszláng égett izgága-formán S ő csak feküdt horgassá vált az orra s nyakát mintha kissé behúzta volna szokatlanul hiszen tartása álva s jártában is olyan volt mint a szálfa Ki bagatellizálta a világot most íme ő is bagatellizált volt kit nem soká a föld alá bedugnak étkévé sok jóévtvágyú kukacnak de így is most is még csodák csodája...</p> <p><b>szépirodalom</b></p> <p>Hábetler egy lángoló <b>kandeláber</b> mellett állt.</p> <p><b>szépirodalom</b></p> <p>Én meg pontosan a lángoló <b>kandeláber</b> mellett álltam, ömlött rólam a veríték.</p> <p><b>szépirodalom</b></p> <p>Az volt a szerencsénk, hogy porázával a komondor egy felfordult <b>kandeláber</b> rozsdás tömegébe gabalyodott, és most azt ugatta.</p> <p><b>szépirodalom</b></p> <p>A komondor akkorát rántott porázán, hogy a <b>kandeláber</b> nagyot kalimpált.</p> <p><b>szépirodalom</b></p> <p>A kutya poráza lesiklott a <b>kandeláber</b> szarváról; közelhúzta karikáját, egészen ki a huzal végéig, s most ott ugrált, veszekedett a gránát vágta falnyílás alatt, ahol Anibel eltűnt.</p> <p><b>szépirodalom</b></p> <p>Valami világos gesztus kéne: egyértelmű jelzés, tiszta fény, nem gyertyában villamosáram vagy <b>kandeláber</b>, spirituszszagú. mezsgyeoszlopától, ettől a mennyezetes ágytól, amelynek viaszrózsa baldachinja alatt valamikor én is helyet foglalok, csak tudatlanságom szédületével jöhetnek el, és hogy csak egy kevéssel többet tudjak adottságaimról, mint a villanypásztor drótkeretében legelő ökör, amelynek nemsokára egy különleges teherautó magasított oldaldeszkájához nyolcadmagával hozzákötözik a szarvát az utolsó utazásra, amelynek jelentőségét a vágóhíd állatorvosa szerint az ökrök nem érzik meg, s hogy hátratevő választásainknak tíz ujjal szorítható zsinórmértéke legyen, legalább egy-egy pillanatra tudnom kellene, mivé leszek e faggatózás elvillanó többlete nélkül, látogatóimra és érdemeim fölserorolására várva itt a tömjénzaggal föllángoló két <b>kandeláber</b> között, s mivé lett apám, véletlen és testvéri előképem, ez a puha halotti maszk, ez a domború és elkeskenyedett, de mégsem egészen természetű arcmása múlt-heti önmagának, ez a szürkületi figyelem, ez a ráncatlan végérvényesség, ez a görcstelen befeléfordulás, ez az elfüggönyözött vetítővászon, ez a sérthetetlen itt-nem-lét, az énnék mondott bioelektronikus varázscső káprázatainak elvonulása erről az égő álarcról, ez a széttépett függetlenségi nyilatkozat, egy páratlan szolgálati szabályzat valamennyi pontjának érvénytelenítése, ez a tökéletes egyezés gallérjával, az óralánccal, amelyről leszerelték a zsebórát, ez a térelem</p> <p><b>szépirodalom</b></p> <p>A társaság önmagával elfoglalva, néhány <b>kandeláber</b> és távoli ágakra függesztett lampion világánál szétszóródott az éjszakába süppedt tájban, csak az jelezte, hogy itt is, ott is mulatók csatangolnak, hogy amott zajosan üdvözölték az éppen felkelő holdat, emitt egy tenorhang, Petrarca szonettjeinek kis kötetét tartva kezében, hamisan dalolta, hogy " Ha nem szeretmet, akkor hát mit érzek?... "</p> <p><b>szépirodalom</b></p> <p>Az ebédőnkben ekkor már nem volt bútor, csak ember és ravatal meg két nagy <b>kandeláber</b>.</p> <p><b>szépirodalom</b></p> <p>Napszemüvegben megcsodálja színpompás városunk (bástyaormzat <b>kandeláber</b>) jobbra a gomb az esti séta balra a gomb a végzetes nő középuitt a férfivécé</p> <p><b>szépirodalom</b></p> <p>Egyszer tágas, komor házban járkáltam, egyik néptelen teremből a másikba haladva: sehol semmi mozgás, csak a falakon függő tükrökben, a padló márványlapjainak négyesögeiben s a nehéz, faragott bútorok fényezett fájában tükröződött a falakra erősített számtalan <b>kandeláber</b> gyertyáinak lángja.</p> <p><b>szépirodalom</b></p> <p>S ahogy ott álltam és ámulva néztem, odakinn megint fölüzött a szél, egy szélroham betört a nyitott ablakon, s a fehér függönyt befelé fújta a terembe, s ahogy ferdén elnyúlva lobogott a könnyű anyag, egyszerre csak belékapott az egyik <b>kandeláber</b> gyertyáinak lángja, s az ide-oda verdeső függönyszárnyról a tűz átterjedt a ravatalon fekvő alak hajára és ruhájára.</p> <p><b>szépirodalom</b></p> <p>Kisvártatva nyughatatlanul fölpattant a székről, és én már azt hittem, ismét magamra hagy, de csak a réz gyertyatartóba szűrt gyertyát gyújtotta meg, majd annak lángjáról a falra erősített <b>kandeláber</b> gyertyáit, azután újra odaült mellém, s egy darabig némán várakoztunk tovább.</p>
--	--

Látjuk, hogy a találatot száma radikálisan csökkent (918-ról 165-re), ám nem csupán azért, mert eltűntek a nagy kezdőbetűs formák, hanem azért is, mert a toldalékos formák is eltűntek, hiszen itt szóalakra való keresésről van szó.

Természetesen ezt az opciót választva kizárjuk azokat a találatokat is, melyekben a célszó nem tulajdonnév, hanem köznévi, ha azok valamilyen oknál fogva nagy kezdőbetűsek: többnyire azért, mert mondatkezdő helyzetben állnak.

Emiatt a gyakorlati munka során célszerűbbnek látszik egyenként átnézni a találatokat, és kihagyni azokat, amelyekben a célszó tulajdonnév (amennyiben csak a közneveket vizsgáljuk). Ha a találatok száma túl nagy ehhez, véletlen mintával dolgozhatunk (erről alább lesz szó).

A kis és nagy kezdőbetű megkülönböztetésének a **mondattani kutatásokban** is hasznát vehetjük, ha például azt vizsgáljuk, milyen kötőszók szoktak mondatkezdő helyzetben állni, s ezek mennyire gyakoriak.

Nézzük meg példaként a *de* kötőszót mondatkezdő helyzetben! Ha ezt akarjuk vizsgálni, a szót nagy kezdőbetűvel írjuk be a keresőbe, s persze kipipáljuk a „kis/nagy kezdőbetű számít” jelölőnégyzetet:

A találatok rendezéséhez a „keverés” opciót választjuk, hogy ne csupa szépirodalmi találat jelenjen meg az első találati oldalon (a rendezésnek erről a módjáról alább lesz szó). Annak érdekében, hogy szembetűnőbb legyen a mondatkezdő helyzet, nemcsak egy mondatot jelenítettünk meg, hanem mondattól függetlenül kontextusba helyeztük a találati szót (ennek beállítási módjáról alább lesz szó). A keresés eredményeként a találati felület első oldalán ez a kép fogad bennünket:



Amint a fenti képről leolvashatjuk, az MNSZ2 a *de* kötőszóra több mint négy és fél millió találatot ad. Könnyen kiszámíthatjuk, hogy a nagy kezdőbetűs *De* az összes előfordulás 18,4%-át teszi ki. Vagyis a *de* kötőszónak minden ötödik-hatodik előfordulása nagy kezdőbetűs. Ennek az aránynak nagyjából meg is felel, hogy mindjárt az első találati oldalnak a képernyőképen látható részén 7 nagy kezdőbetűs *De* található, melyből 6 releváns.

### ÷ D) Karaktersorozatra való keresés

A „**karakter**” opció abban különbözik a többitől, hogy míg az előzőben nyelvi elemekre kerestünk rá (melyeket a program úgy is kezel, tehát nyelvi elemekként, és megkülönböztette, ha kellett, a szótőt a toldalékoktól), most karaktersorozatra (azaz karakterek, például betűk, írásjelek egymásutánjára) keresünk rá: ezt a program természetesen nem is kezeli nyelvi elemként, nem elemzi.

The screenshot shows the MNSZ2 search interface. At the top, the logo 'MNSZ2' is displayed. Below it, the user is identified as 'Felhasználó: gramma' and the corpus as 'korpusz: MNSZ2'. On the left side, there is a navigation menu with 'Keresés' (Search) selected, and other options like 'Lehetőségek: Kontextus Alkorpuszok'. The main search area contains a form with the following fields: 'korpusz:' with a dropdown menu set to 'MNSZ2', 'lekérdezés típusa:' with a dropdown menu set to 'karakter', and 'karakter:' with an empty text input field. Below the input fields are two buttons: 'Konkordancia készítése' and 'Törlés'.

Ez az opció akkor nagyon hasznos, ha egy szórészre vagyunk kíváncsiak. Írjuk be pl. az *alom* karaktersorozatot!

This screenshot is similar to the previous one, but the 'karakter:' input field now contains the text 'alom'. The rest of the interface, including the 'MNSZ2' logo, user information, and search options, remains the same.

A találatokat egy-egy mondatnyi kontextusban jelenítjük meg.

találatok száma: 1297072  
 oldalszám: 1 / 865 Ugrás következő | utolsó

**szépirodalom** Vagy csipetnyi puha vatta Meszes falak között szunyadt Jó sokáig, az ebadta Csukott ablak óvta álmát Hosszantartó éjszakáját Jó melegben, csend ölében Mint a tészta közepében Ha tehetné, ki se bújna Összegömbölyödik szépen De egy napon arra ébred A **fájdalom** elborítja Szikrázó fény elvakítja Körül minden tisztán ragyog Megszületett, nincs más dolga Selyemfészek tenyerében Kicsiny szárnyan sok-sok pihe Pityang bóbítája volna?

**szépirodalom** Merőleges járatot véstek a leomlott falba, apró csőrükkel fáradhatatlanul hordták a cseppnyi homokot, nem is törődtek vele, hogy a laza homok néha beomlik, ők minden **alkalommal** újra kezdték a járat építését, s végül diadalmaskodtak, a másfél méter hosszú folyosó elkészült, hozzáfoghattak a bentli fészek kialakításához.

**szépirodalom** Ilyenkor **izgalom** támadt a gyerekek között, az anyukák elérhetetlen magányba húzódtak, a könyörgés, elutasítás, magyarázkodás hangjai töltötték be az udvarokat.

**szépirodalom** **Malomkerekek** .

**szépirodalom** Én egy vödörrel jártam az utcákat, leginkább a vasútállomás felé kóricáltam, a **teherforgalom** jó részét még muralovakkal vontatott szekerek bonyolították le.

**szépirodalom** Minden **alkalommal** kitaláltunk valamiféle új játékot.

**szépirodalom** S akármennyi pofont is kaptam később a szerb nyelv hiányos tudása miatt, már nem voltak képesek leszoktatni róla, hogy **önszorgalomból** tanuljam, szeressem, kutassam az ezen a nyelven föllelhető szépségeket.

**szépirodalom** Cili néni azonban bátran szedett a fákról is gombákat, mondta is nekem minden **alkalommal** , hogy ez vörös tinóru, ez meg özláb, ez meg sárga rókaomba.

**szépirodalom** Vannak némely nők, akik minden valódi, öröklött **hatalom** nélkül boszorkánynak tetszenek, ámbár mindössze utazatai ezeknek.

**szépirodalom** Institoris mesterrel hajdanában, majd később, egészen napjainkig, számos boszorkány halálánál voltunk jelen, s nem egyet magunk idéztünk elő - a boszorkányok halálát azonban megszokni nem lehet, s minden egyes **alkalommal** újra megérintik a lelket, nyomot hagyva benne.

**szépirodalom** Ennek megtörténte után a boszorkány teljesen ártalmatlanná válik az ilyen férfi számára, és sokkal kevésbé lesz veszélyes másokra is - ugyanis, mint leiratott, a boszorkányok élető eleme a **hatalom** és a szenvedély, s e kettő nélkül üressé válnak, és értelmetlenné létezésük.

**szépirodalom** Úgy múlt el, akár a vihar, mely fákát dönt ki, vizeket áraszt és falvakat söpör el, majd egy pillanat múlva néma, embertelen csend és **nyugalom** marad helyén az elpusztított helyen.

**szépirodalom** A boszorkányok halálának pillanatában az el- és felszabadult démoni **hatalom** , az érzelmetek és érzeteket lebilincselő erő még egyszer, utoljára átsöpör a földön, s egy kis időre újra hatalmába kerít mindenkit, akihez a boszorkánynak valaha köze volt.

**szépirodalom** Éppen ellenkezőleg: Institoris intelligens volt, művelt, s minden idegensége dacára kitűnően eligazodott a legváratlanabb helyzetben is; ráadásul valami hűvös **nyugalom** lebegte körül, ami csak a nagyon kiegyensúlyozott emberek sajátja.

**szépirodalom** Enyhe, ibolyaillatú parfümöt árasztott, s ebből a közelségből csak még nyilvánvalóbbá vált rendkívüli szépsége - szinte fizikai **fájdalomba** került uralkodnom magamon, hogy ne érintsem meg, s még tekintetemet levenni róla is nehéz volt.

**szépirodalom** Mivelhogy pedig, mint mondtuk, a testiség mellett a **hatalom** a boszorkányok legfőbb stimulusa, megesis néha, hogy az utóbbi okából erőnyesebbek az orsolyás nővéreknél, ha érdekük úgy kívánja, s ezzel jobban lenyűgözhetik a férfit.

**szépirodalom** Harmadjára pedig, s nem utolsó sorban, a gonoszágban erő rejtezik, és sötét **hatalom** , míg a rosszindulat csupán a gyáva emberek fegyvere - márpedig, amint azt korábban megírtuk, nincs a **hatalomnál** nagyobb delejesség és vonzerő a boszorkányok számára.

**szépirodalom** Harmadjára pedig, s nem utolsó sorban, a gonoszágban erő rejtezik, és sötét **hatalom** , míg a rosszindulat csupán a gyáva emberek fegyvere - márpedig, amint azt korábban megírtuk, nincs a **hatalomnál** nagyobb delejesség és vonzerő a boszorkányok számára.

**szépirodalom** Az **izgalom** , a megfoghatatlan varázs elült - ez azonban vészjósló csend volt, s benne a poeta szava léleklarang.

**szépirodalom** Az Anyaszentegyháznak pedig nincsen már meg a világi cselekvéshez szükséges ereje és hatalma, így ' 72-ben, mikor Őszentsége kinyilatkoztatta: a Sátán elszabadult, és démonai szerte garázdálkodnak a földön, a rettenetes tényből röpké életű sajtócsemege, kuriózum váltott; s ebben nem kevés része volt maguknak a boszorkányoknak is, akik mindenkor jelen voltak az irástudók és betűvetők mellett, és azokat befolyásolva, ellenőrizve kétszeres **hatalomra** tehettek szert.

Látjuk, hogy nagyon eltérő szavak jöttek elő, melyekben csupán az a közös, hogy az *alom* szórészt tartalmazzák: *fájdalom, alkalom, izgalom, Malomkerekek, teherforgalom, önszorgalom, hatalom, nyugalom*.

Ezekben az *alom* szórészt eltérő funkciójú lehet: lehet egy tőszónak egy darabja (pl. *Malomkerekek*), lehet egy képzőnek egy darabja (pl. *fájdalom*), lehet képző (pl. *izgalom*).

### ÷ E) CQL-keresés

A következő opció, a **CQL**, lehetővé teszi, hogy egy korpuszlekérdező nyelven (Corpus Query Language) fogalmazzuk meg a keresést.

Ennek az opciónak a használatához el kell sajátítani ezt a lekérdező nyelvet (<https://www.sketchengine.eu/documentation/corpus-querying/>). Olyankor érdemes ezt a funkciót használni, amikor bonyolultabb keresési feltételeket akarunk megadni, olyanokat, melyeknek a megadása az alább bemutatott részletes kereséssel nem lehetséges. A CQL-keresés azt is lehetővé teszi, hogy szótőre is rá tudjunk keresni a kis és a nagy kezdőbetű megkülönböztetésével.

Végezzünk most el az egyszerűség kedvéért egy olyan keresést, amit amúgy meg tudunk oldani más módon is.

Írjuk be a „karakter” mezőbe azt, hogy [lemma = „kandeláber”].

**MNSZ2**

Felhasználó: **gramma** korpusz: **MNSZ2** >> v2.0.5 - friss verzió 2018. június 6-tól << [Hírek](#) [Súgó](#)

Keresés [?](#)

Lehetőségek:  
 Kontextus  
 Alkorpuszok [?](#)

korpusz:   
 lekérdezés típusa:   
 CQL:  alapértelmezett attribútum:

Lexical Computing Ltd. [LCC](#)  
 Sketch Engine (ver:open-2.59.1-open-2.91.13)  
 A Nemzeti Korpuszportál tagja.  
 MNSZ2 - v2.0.5 - 2018-06-08  
 Nyelvtípus: [English](#) | [magyar](#)

Az eredmény az lesz, hogy a program a *kandeláber* szótőre keres rá:

Felhasználó: **gramma** korpusz: **MNSZ2** >> v2.0.5 - friss verzió 2018. június 6-tól << [Hírek](#) [Súgó](#)

Keresés [?](#)

Mentés  
 Megjelenítés  
 KWIC/mondat  
 Rendezés  
 bal | jobb  
 találat  
 infó  
 keverés  
 Véletlen minta  
 Szűrés  
 Gyakoriságok  
 szóalakok  
 szótövek  
 elemzések  
 dokumentumok  
 Kollokációk [?](#)

Visszajelzés  
 Próbálja ki a **Mazsola** eszközt igék és argumentumaik közvetlen vizsgálatához!

korpusz: **MNSZ2**  
 találatok száma: **918**

**szépirodalom** Novemberben különösen vigyázni kell, mert a nyálkás hidegben a galambszar is jobban csúszik, láttam már elegáns, fekete kabátos férfiakat, akiknek olasz cipője nem bírt el a madárkakival, ott hempergöztek a koraesti szürkeségben, a sárga fényű **kandeláberek** alatt, korhely disznók módjára, a kovácsoltvas biciklitartóba kapaszkodva támaszkodtak volna talpra, hogy bejussanak végre a patika ajtaján.

**szépirodalom** Szürke, egyöntetű romhalmoz volt a város, minden összeomlott, csak a gúnárnyakú **kandeláberek** álltak.

**szépirodalom** Ő csak feküdt ott fölpattanva szája s mondta a pap kezében bibliája A halottasház ajtaja az egyik rosszul betámasztván zörögve megnyílt Volt aki ingujjban jött más megint ki feketében nyakkendővel vonult ki A gyászhuszarak szinte félcvilben: kék-szürke nadrágban s kék-szürke ingben S a kondoleálók sorjában egyre hol csak kezét szorítva hol ötleve A fejnél két vas **kandeláber** ormán spiritszláng égett izgága-formán S ő csak feküdt horgassá vált az orra s nyakát mintha kissé behúzta volna szokattanul hiszen tartása állva s jártában is olyan volt mint a szálfá Ki bagatellizálta a világot most íme ő is bagatellizált volt kit nem soká a föld alá bedugnak étkévé sok jóétvágyú kukacnak de így is most is még csodák csodája...

**szépirodalom** **Kandeláberek** lángjától?

**szépirodalom** A szobában hőség volt, hatalmas **kandeláberek** lobogtak.

**szépirodalom** Hábétel egy lángoló **kandeláber** mellett állt.

**szépirodalom** Az öreg Zentay a fiával róza- és ciklámenkoszorút hozott, a ravatalozó emberek alig tudták elhelyezni a **kandeláberek** között illatozó virághegyen.

**szépirodalom** A szobában hőség volt, hatalmas **kandeláberek** lobogtak.

**szépirodalom** Én meg pontosan a lángoló **kandeláber** mellett álltam, ömlött rólam a veríték.

**szépirodalom** Az volt a szerencsénk, hogy porázával a komondor egy felfordult **kandeláber** rozsdás tömegébe gabalyodott, és most azt ugatta.

**szépirodalom** A komondor akkorát rántott porázán, hogy a **kandeláber** nagyot kalimpált.

**szépirodalom** Mert a komondor sem ismert; és nyálát vadul csurrantva - a vállalt rabszolgaság és a szeretet jogán - visszakövetelte meleg vérünket, szagunkat, tenyerünket, bálvány mivoltunkat: a hóban maga után vonszolta a **kandelábert** .

**szépirodalom** A kutya poráza lesiklott a **kandeláber** szarváról; közelhúzta karikáját, egészen ki a huzal végéig, s most ott ugrált, veszekedett a gránát vágta falnyílás alatt, ahol Anibel eltűnt.

**szépirodalom** A bronzarany **kandelábereket** !

**szépirodalom** Marioláne tétlenül üldögélt a falnál, az üvegfürtös **kandeláberek** alatt, és ölébe ejtett kézzel, révetegen bámult maga elé.

**szépirodalom** - Illúzióknak - mondta mereven Morelli, és fölneézett a piszkossárga fénnel kigyulladó **kandeláberre** .

**szépirodalom** Valami világos gesztus kéne: egyértelmű jelzés, tiszta fény, nem gyertyában villamosáram vagy **kandeláber** , spiritszagú.

**szépirodalom** A **kandeláberek** spiritszag-élményítő; VIII az esküvői kalapját viseli, sűrű fátyollal, feketére festve.

**szépirodalom** Alvadt-vér-színű árnyéksáv szivárog a lomb közt s cseppenként alapereg, kialusznak a gesztenyevirágok, e nappal égő **kandeláberek** .

**szépirodalom** Kialusznak a gesztenyevirágok, e nappal égő **kandeláberek** .

**szépirodalom** Alvadt-vér-színű árnyéksáv szivárog a lomb közt, s cseppenként alapereg, kialusznak a gesztenyevirágok, e nappal égő **kandeláberek** .

**szépirodalom** Majd a hadak is megjelentek s nyomulni kezdtek északról, keletről, nyugatról - délnak. A Tejút ívlámpái sorra elsötétedtek, a Göncölök, mint a kisiklott villamosok, oldalra dőltek. Aztán sötét lett teljesen. A felhők megszállták az eget, s kihirdették az ostromállapotot. De nem volt már kinek. Kihirdették a szabad rablást. Erre az addig fegyelmezett és zárt csapatok örült fosztogatásba kezdtek: törtek-zúztak, föltépték a vízvezetéseket, bevették az ablakokat, kidobálták az égő **kandelábereket** . A tehetetlen düh, hogy senkit se találtak, az dolgozott bennük. Bár egy bujkáló, egy ellenálló, egy nő vagy legalább egy csecsemő, akit falhoz kenhetek volna, de senki. Fölhasogatták a párnákat, a dunyhákat, s aztán napok múltán, mert nem tehettek mást, kihirdették a békét.

Pontosan ugyanazt kapjuk tehát, mint ha a „szótő” opciót választottuk volna, s a keresőbe beírtuk volna a *kandeláber* szót (l. fentebb).



## ÷ F) Részletes keresés

A legfontosabb opció számunkra a **részletes keresés**. Itt is megadhatjuk, hogy **szóalakra** vagy **szótőre** akarunk-e keresni.

Az alábbi képen a „szóalak” opció szerepel.

MNSZ2

Felhasználó: [gramma](#) korpusz: [MNSZ2](#) >> v2.0.5 - friss verzió 2018. június 6-tól << [Hírek](#) [Súgó](#)

Keresés ?

Lehetőségek:  
Kontextus  
Alkorpuszok ?

korpusz:

lekérdezés típusa:

részletes keresés:   szófaj:

összetett szó | morf:  morféma:

Lexical Computing Ltd. [LCC](#)  
Sketch Engine (ver:open-2.59.1-open-2.91.13)  
[A Nemzeti Korpuszportál tagja.](#)  
MNSZ2 - v2.0.5 - 2018-06-08  
Nyelvváltás: [English](#) | [magyar](#)

Megjegyezzük, hogy itt még a „szóalak” opciónál sincs lehetőségünk a kis és nagy kezdőbetű megkülönböztetésére, ezért ha szükségünk van a kis és nagy kezdőbetű megkülönböztetésére, sajnos részletes keresést nem végezhetünk; erre csak – amint fentebb jeleztük – a CQL nyelv használata esetén van lehetőség.

Amint az alfejezet első képéről kiderül, a „részletes keresés” opció választásával lehetőségünk nyílik a célszó **szófaji értékének** megadására, ami többek között akkor hasznos, ha a **homónímiát** akarjuk kikerülni, pl. a *zavar* igét vizsgáljuk, s így a *zavar* főnévre vonatkozó adatokat szeretnénk kiküszöbölni.

Nem biztos, hogy százszázalékosan sikerül, a két szófaj azonos toldalékai gondot okozhatnak; pl. a *zavar* főnévre rákeresve a találatok közt a *zavarom* igealakra vonatkozó találatok egy része is megjelent, nyilván azért, mert alakilag nem különbözik a *zavarom* névszóalaktól. Ezek csak a találatok kézi tisztításával küszöbölhető ki.

Fontos tudni, hogy úgy is kereshetünk, hogy **csak a szófaji értéket állítjuk be**, s a keresőmezőt üresen hagyjuk. Az alábbi képen például a „főnévi igenév” szófaji értéket állítottuk be:

MNSZ2

Felhasználó: [gramma](#) korpusz: [MNSZ2](#) >> v2.0.5 - friss verzió 2018. június 6-tól << [Hírek](#) [Súgó](#)

Keresés ?

Lehetőségek:  
Kontextus  
Alkorpuszok ?

korpusz:

lekérdezés típusa:

részletes keresés:   szófaj:

igekötő:  alapak/szám:

összetett szó | morf:  morféma:

Lexical Computing Ltd. [LCC](#)  
Sketch Engine (ver:open-2.59.1-open-2.91.13)  
[A Nemzeti Korpuszportál tagja.](#)  
MNSZ2 - v2.0.5 - 2018-06-08  
Nyelvváltás: [English](#) | [magyar](#)

Ebben az esetben a keresés indítása után tetszőleges igék főnévi igeneve fog megjelenni:

<p>Keresés ?</p> <p>Mentés</p> <p>Megjelenítés</p> <p>KWIC/mondat</p> <p>Rendezés</p> <p>bal   jobb találat</p> <p>infó</p> <p>keverés</p> <p>Véletlen minta</p> <p>Szűrés</p> <p>Gyakoriságok</p> <p>szóalakok</p> <p>szótövek</p> <p>elemzések</p> <p>dokumentumok</p> <p>Kollokációk ?</p> <p>Visszajelzés</p> <p>Próbálja ki a Mazsola eszközt igék és argumentumaik</p>	<p>korpusz: MNSZ2 találatok száma: 10934260</p> <p>oldalszám: 1 / 7290 Ugrás <a href="#">következő</a>   <a href="#">utolsó</a></p> <p><b>szépirodalom</b> mászni kezdtem én, de lehuppantam a ló túlfelén, másztam, leestem, másztam újra s megint, látom: a boltos erre csak nevet és legyint, végül megsajnál és feltuszkol a lóra, a fából való, kantáros fakóra.</p> <p><b>szépirodalom</b> S láttam lábam előtt <b>heverni</b> , a világot!</p> <p><b>szépirodalom</b> Egycsapásra el akartam <b>érni</b> a tengert, a távoli vizeket, a bőszi vihart, tajtékos óceánt, hogy megmutassam, ki ám a legény!</p> <p><b>szépirodalom</b> ültünk, nagyapám Csikos csergébe bugyolálta testünk Két bivaly állt a házunk udvarán Sötét bőrük fehér páráit lestük Az éj füstje fölöttük lobogott Beleborzongtunk hűvös harmatába Kapu nyílt és a bivalyszekeret <b>Ballagni</b> hagyta egy távoli világba Vasárnapon, egy szekér s rajta mi Kölykök, kik titkon nem alusznak éjjel Egy hegyi faluból mentünk virrasztani Egy vásárba, mely álmainkban vész el.</p> <p><b>szépirodalom</b> ültünk, nagyapám Csikos csergébe bugyolálta testünk Két bivaly állt a házunk udvarán Sötét bőrük fehér páráit lestük Az éj füstje fölöttük lobogott Beleborzongtunk hűvös harmatába Kapu nyílt és a bivalyszekeret Ballagni hagyta egy távoli világba Vasárnapon, egy szekér s rajta mi Kölykök, kik titkon nem alusznak éjjel Egy hegyi faluból mentünk <b>virrasztani</b> Egy vásárba, mely álmainkban vész el.</p> <p><b>szépirodalom</b> szomszéd kislány egyre babázott, autót <b>vezetni</b> sehogy se vágyott.</p> <p><b>szépirodalom</b> Bodri bezzeg örül neki, ami lehull, eltünteti, és mert neki <b>enni</b> adok, nem bánja, hogy ember vagyok, barátságunk örök kincsem, pedig ma már Bodri - nincsen.</p> <p><b>szépirodalom</b> hogy majom vagyok, pedig én csak grimaszkodok, s nem tudok a fára <b>mászni</b> , az ablakon ki-bejárni.</p> <p><b>szépirodalom</b> hogy majom vagyok, pedig én csak grimaszkodok, s nem tudok a fára mászni, az ablakon <b>ki-bejárni</b> .</p> <p><b>szépirodalom</b> bicikli sírva kért: ne hagyjam <b>veszni</b> , felfalja a rozsda!</p> <p><b>szépirodalom</b> Lassan, dőcögve vészorgunk tova, itt egy pedál hull el, ott egy kerék, már csak a kormányt szorongatja kezem, s haza kellene <b>menni</b> még...</p> <p><b>szépirodalom</b> <b>Hallani</b> , hogy zümmög a légy.</p> <p><b>szépirodalom</b> Sokat fogunk így <b>várni</b> még?</p> <p><b>szépirodalom</b> 30A hasraesést még <b>gyakorolni</b> kell</p> <p><b>szépirodalom</b> nem lehet bárhol <b>fölbukni</b> az utcán.</p> <p><b>szépirodalom</b> Kinő újra Észrevétlenül, tolvajmódra Nyelvemhez simult, tiéd vagyok De téled azért most búcsúzó Nő helyembe más, itt az óra Ne siess, kérlektem, el ne hagyj Még csak egy éjszakát, egy napot Tejfogam vagy s a tiéd vagyok Hogy is válthatnánk el, még maradj Nem lehet, súgta, életemnek Szabálya: <b>menni</b> kell, hát megyek Kihullott.</p> <p><b>szépirodalom</b> <b>Aludni</b> kell, ébren vagyok, és azt hiszem, hogy álmodok: ezernyi fürge pukkanás, megannyi szikravillanás.</p> <p><b>szépirodalom</b> Sok ágat pusztítottam, mert szerettem fűz fasípot <b>gyártani</b> .</p> <p><b>szépirodalom</b> Néha eszembe jut, hogy anyám nélkül nem voltam képes <b>elaludni</b> .</p>
--	---

Látjuk, hogy a legkülönbözőbb igékből képzett főnévi igenevek jelennek meg, a konkordancia egészében csaknem 11 millió találatot kaptunk.

Ennek a lehetőségnek a választása olyankor hasznos például, amikor a főnévi igenévnek ket-tős alany részeként való használatát vizsgáljuk. Ehhez be kell állítani, hogy a főnévi igenév együtt forduljon elő a *látszik* igével. Erről alább, a kontextus megadásánál lesz szó.

## ÷ 2. Egyéb keresési beállítások

Miután megismerkedtünk a lekérdezés típusaival, vessünk még egy pillantást a bal oldali menü sorra, ahol a „Keresés” menüpontot látjuk legfelül. Ezzel indíthatunk új és üres keresést: rákattintva a keresőmezőből eltűnik az előző keresésre utaló karaktersor.

A találatokat mutató felületen is megtaláljuk a „Keresés” menüpontot, ezzel kezdhethetünk a találatok megtekintése után új keresést, visszajutva a fő keresőablakba.

MNSZ2

Felhasználó: **gramma**

Keresés  
?

Lehetőségek:  
Kontextus  
Alkorpuszok  
?

Bizonyos esetekben ilyenkor törlődik a keresőmezőből a célszó és néhány más beállítás, így a célszót újra be kell írunk, ill. az eltűnt beállításokat vissza kell állítanunk.

Ezt kétféleképpen kerülhetjük el. Az egyik lehetőség az, hogy a böngésző visszafelé mutató nyílával lépünk vissza; ez olyankor járható út, ha nem „mozogtunk” sokat a konkordanciában, és egy-két visszalépés után már a keresési felületen is vagyunk. Ha azonban a konkordancia felü-

letére érkezve már több oldalt nyitottunk meg, az összesen végig kell lépnünk visszafelé, ami kényelmetlen.


Ilyen szempontból jobb a másik módszer, amelyet néhány böngésző (például a Chrome) lehetővé tesz, más (például a Firefox) nem: miközben a „Konkordancia keresése” gombra kattintunk, vagy miközben a billentyűzet „Enter” gombjának segítségével indítjuk el a keresést, lenyomva tartjuk a Ctrl-billentyűt. Ilyenkor a konkordancia egy új ablakban nyílik meg, így ablakváltással vissza tudunk térni a keresőfelülethez (a konkordancia az MNSZ1-ben eleve új ablakban jelenik meg, így ott ilyen problémánk nincs).

A „Keresés” menüpont alatt egy kérdőjelet találunk: erre kattintva megnyílik a súgó. Egy további lehetőség a súgó megnyitására a felület jobb oldalán adódik, a „Súgó” menüpont választásával. Amint fentebb már említettük, az MNSZ súgója sajnos nagyon kevés eligazítást nyújt az MNSZ2 használatához.

## MNSZ2

### Súgó

#### 1. Bevezető

Ezen az oldalon a Magyar Nemzeti Szövegtár felületét, a keresési és feldolgozási lehetőségeket mutatjuk be. Az oldal tanulmányozása közben érdemes külön ablakban megnyitni a [keresőfelületet](#), hogy az alább szereplő  ikonnal jelölt példákat, gyakorlatokat rögtön kipróbálhassuk. A keresőfelülethez [regisztráció](#) után lehet hozzáférni. A [>>](#)-re kattintva pedig rögtön megtekintheti a szóban forgó lekérdezés eredményét. Az MNSZ kibővített változatának felülete a szabadon elérhető [NoSketch Engine](#) rendszer felhasználásával készült el. Jelen dokumentum merít e rendszer [angol nyelvű súgójából](#). Az alábbiakban a *szövegtár* és a *korpusz* kifejezéseket szinonimaként fogjuk használni.

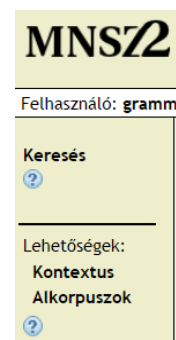
A felület segítségével hozzáférünk a szövegtár teljes anyagához, és az összes hozzáadott információhoz, annotációhoz.

A nyitólapra mindig a bal felső sarokban lévő MNSZ2 logóra kattintva térhetünk vissza.

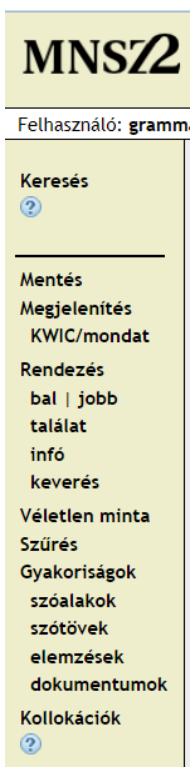
A bal oldali menüsorban látható további lehetőségekre a „Kontextus” és az „Alkorpuszok” menüpontok utalnak.

Ezekkel lehet a keresést kétféleképpen is leszűkíteni: bizonyos kontextusra vagy bizonyos alkorpuszra (régióra vagy stílusrétegre).

A két lehetőséget alább külön-külön fejezetekben tekintjük majd át, előtte azonban még megismerkedünk a megjelenítési beállításokkal.



## & MEGJELENÍTÉSI BEÁLLÍTÁSOK



Mielőtt megkezdénénk a munkát, el kell végezni a megjelenítési beállításokat. Mivel ezeket **csak a találatokat mutató felületről lehet elérni**, keressünk rá bármilyen szóra bármilyen módon. A találati eredményektől balra új menüsort látunk.

### ÷ 1. Mentés

Az első pont a **„Mentés”**, mellyel el tudjuk menteni az éppen elkészült konkordanciát (a találatokat tartalmazó szövegrészleteket), alapbeállítás szerint egy egyszerű **.txt fájlba**. Például a „hitelbe” szóalakra való rákeresés eredménye, a baloldali szekvencia szerint rendezve ez:

```
# Corpus: MNSZ2
# Hits: 580
# Query word,[word=„hitelbe”] 580
# Sort word in left context 580
```

sajtó sztrádaépítés, mert a cégek jószerivel sajtó költségükön, „ < hitelbe > „ építik az utakat. A kormányzat pedig csak utólag fizet. A tudományos pedig az „ oktatóiskola „ elvont logikája szerint - úgymond: „ < hitelbe > „ - azért, hogy „ okosabbak legyünk „). Valójában a „ dolog személyes-fórum fordítani akarják az IMHO nem más mint, mint a választások előtt „ < hitelbe > megvett szavazatok „ ellenértékének kifizetése. Mert a hitelek személyes-közösségi elfogyasztunk a bolygó éves erőforráskészletét, vagyis inentől „ < hitelbe > „ fogyasztunk. Az Earth Overshoot Day-ról bővebben az alábbi szépirodalom Éppen elég lehetett nekik, hogy a Szovjetunióknak továbbra is „ < hitelbe > „ kellett szállítaniok.) Összegezve kiderült, hogy Magyarország sajtó pénzekből. Nem kizárt azonban, hogy a szakértői kamara a jövőben „ < hitelbe > „ nem dolgozik majd a rendőrségnek, azaz más megállapodási személyes-fórum gazdálkodó szervezetek által a költségvetésnek „ rövidlejáratú „ < hitelbe > adott pénz. Az önkormányzatok finanszírozásában szintén hallani sajtó mellett a környezet állapota. Az idén már szeptember 27-étől „ < hitelbe > fogyasztja „ a Föld erőforrásait az emberiség, vagyis már feléltük személyes-fórum „ megteremtését „ ügyesen „ későbbre halasztották, vagyis „ < hitelbe > „ dolgoztatták az érintetteket. speaker 02/27/99 13:00:15 Az személyes-fórum jelentkezni. Ha gondolod, az én jegyemet is kifizethetné ( < hitelbe > ), mert valszeg egyszerűbb

odamenni az INDEXesekhez és ott befizetni  
személyes-fórum beszélgetés, piálás. Ez utóbbi rendkívül olcsó, önköltségi ár alatt, < hitelbe > is. Politizálás szinte semmi.  
Most még. Majd szép fokozatosan  
szépirodalom Úgy vesztegette ezt a sok holmit potya pénzen, magára alkudva, < hitelbe > is adta, kínálta, doboltatta a falunak. Legisinkább persze  
beszélnyelvi Hobbyméhészek. A fiataloknak nem elég ez. Nincs pénz befektetni, < hitelbe > ne szabad menni, úgy hogy ez egy kicsit nehezebb indulunk,  
sajtó forint) értékre elkövetett váltóhamisítással, banki csalással, < hitelbe > vett áruval való üzérkedéssel és biztosítási bűncselekményekkel  
személyes-fórum benne minden pénzzel meg irattal. A benzinem persze elfogyott, < hitelbe > tankoltam és kétszer kidumáltam magam igazoltatásnál -- fokozott  
tudományos tárgyaljanak. Szállást, bőséges vacsorát kaptak a fogadóstól, < hitelbe > még szerszámokat, élelmet (egy hétre) és két felszerelt öszvért  
sajtó postásnő. - Nem, írja csak: „Etel a Hotelbe, kapitánya a fotelbe, < hitelbe > hatol be. Textiláru föl, selyemáru le, szőrmeáru összevissza  
sajtó kénytelen beérni ilyesfajta ajánlatokkal, sőt terven fölül, < hitelbe > szállítani gázt - csak az idei év utolsó negyedében 5 milliárd  
beszélnyelvi tehát a hírekben. De hogy ez mit jelent tulajdonképpen hitelből, < hitelbe >, de miért? A kérdésre Duronelli Péter a Budapest Alapkezelő  
sajtó a kertben, melyek a kedvenc cicák emlékét őrizték. Hozomra, < hitelbe >, ingyen dolgoztak a budakeszi iparosok, ha valami földanivaló  
szépirodalom mortgage, magyarul? NŐ Nem tudom - VENDÉG Maguk is így élnek itt, < hitelbe > - az egész ország hitelbe él - hát Amerikában is - csak ott  
sajtó gatyában, ingben, mezitláb a bolt pultjánál, és pálinkát ittak, < hitelbe >. Öten voltak, öt napszámos, és úgy döntöttek, elmennek Amerikába  
sajtó meg. Ezért, ha totálkáros a gépkocsi és a kis kezdőrészlettel, < hitelbe > vásárolták, akkor a kártérítés az első néhány évben nem fedezi  
sajtó megborítani a családi költségvetést és szükségtelen kiadásokba, < hitelbe >, miegymásba hajszolni magukat és szeretteiket. Áprilisi teljes  
személyes-fórum reggel ott talalod, ahol hagyta, - bemesz a kedvenc kocsmadba, < hitelbe > ihatsz, - kocsival 5-10 perc alatt kint vagy a közeli erdőben  
szépirodalom szilárdságától tellett - eligyekezett. S most, útonálló, ki így lopsz, < hitelbe >, most ezüstőlő prémjére esett a pillantásod, mert mint a fókán  
személyes-közösségi bank kifizetni a betéteseket - az a pénz ki van adva másoknak, < hitelbe >... De fölösleges is magyarázni, lövésetek sincs az egészről  
személyes-fórum . A revesz 5 aranyat ker a munkaert, de a nonek nincs penze, < hitelbe > pedig nem dolgozik. Elmegy a szeretojehez es penzt akar kerni  
tudományos használt ruhát kap. A szomszéd sűrűn vág disznót, és ha nincs pénz, < hitelbe > is ad húst. Napszámba is járnak. Ha hárman elmennek, megvan  
személyes-fórum mindörökké. Ámen „Megvette az az egyházakat (a mi pénzünkön, < hitelbe >) és fehérvári (szül.h.szerint is megyei) illetőségű, valamit  
hivatalos adót, míg a befektetések összege nem csökken. Ha úgy tetszik, < hitelbe > kapja. A befektetési kedvezmények érvényesek a brókerek által  
hivatalos adót, míg a befektetések összege nem csökken. Ha úgy tetszik, < hitelbe > kapja. A befektetési kedvezmények érvényesek a brókerek által  
szépirodalom európai államok versenyeztek, hogy melyik szállítson többet, < hitelbe > is. - Akkor nem értem. Miért kellett ezt megszüntetni? A hatalom  
szépirodalom leadni a bejáróné rágalmait s kérni a vágott hízó vesevelőjét, < hitelbe > a többihez? Vagy Amáreszéknél kopogjon, tüntetni megbotránkozásával  
szépirodalom zsidó ember. Kezesek, váltók erejével autóbust vásároltak, < hitelbe >. A reménység nagy sárga autója ott állt a hosszú utca közepe  
sajtó miniszterelnöksége alatt? bőszyoknyás házaló kereskedőtől vásárolt zugaranyat, < hitelbe >, ott, a rendőrpalotában. Akkor valami vidéki tyúk tojta az  
szépirodalom egy este cserépkancsóikkal: „tessen mán egy kis bort adni „ - < hitelbe >, ha lehet. Arcuk gyászos, hangjuk kunyeráló - de a cserépkancsóba  
szépirodalom majd nem ülhet fel lovára... volt hogy cédákért esengett: - < hitelbe > hozzál kurvát ha meglóptál te szolga! - sziszegte máskor meg  
szépirodalom a rúdjára, mondott egy nagy „ csekk! „ -et, és elkezdődött - < hitelbe >! - a nap. Ha keveset, ha eleget, ha túl sokat meséltem eddig  
sajtó summát - a Hivatásos Labdarúgó Alsövetség közreműködésével - < hitelbe > kapta az FTC. Szeiler József, a futball kft. ügyvezető igazgatója  
szépirodalom , az eladósodást meg a féltékenységet és már nem kell sok... < hitelbe > haljak meg? talán majd azt is Herkules szegezi le rajtam, a

Látjuk, hogy a program által létrehozott fájlban a találati felület adatai jelennek meg. A találati szó **csúcsos zárójelben** van. Az olvasást némiképp megnehezíti, a szövegmegjelenítést pedig rontja, hogy a nyitó zárójel után, ill. a záró zárójel előtt **szóköz** van. Erre azért van szükség, mert előfordulhat, hogy a találati szó az eredeti szövegben is csúcsos zárójelben van. A találati szó előtti és utáni szóköz tehát azt jelzi, hogy a csúcsos zárójel nem az eredeti szövegből származik, hanem a program iktatta be.

Természetesen nem az összes találatot idézzük.

Ha a kimenetet .xml kiterjesztésű fájlban kérjük, ezt kapjuk:

```
<?xml version="1.0" encoding="iso-8859-2"?>

<concordance>

  <heading>
    <corpus>MNSZ2</corpus>
    <hits>580</hits>
    <query>word,[word="hitelbe"] 580 word in left context 580 </query>
  </heading>
  <lines>
    <line>
      <ref>sajtó</ref><left_context> sztrádaépítés, mert a cégek jószereivel saját költségükön, „ </left_context><kwic> hitelbe
</kwic><right_context> „ építik az utakat. A kormányzat pedig csak utólag fizet. A </right_context>
    </line>
    <line>
      <ref>tudományos</ref><left_context>pedig az „ oktatóiskola „ elvont logikája szerint - úgymond: „
</left_context><kwic> hitelbe </kwic><right_context> „ - azért, hogy „ okosabbak legyünk „). Valójában a „ dolog
</right_context>
    </line>
    <line>
      <ref>személyes-fórum</ref><left_context>fordítani akarják az IMHO nem más mint, mint a választások előtt „
</left_context><kwic> hitelbe </kwic><right_context> megvett szavazatok „ ellenértékének kifizetése. Mert a hiteleket
</right_context>
    </line>
    <line>
      <ref>személyes-közösségi</ref><left_context>elfogyasztanunk a bolygó éves erőforráskészletét, vagyis innentől „
</left_context><kwic> hitelbe </kwic><right_context> „, fogyasztunk. Az Earth Overshoot Day-ről bővebben az alábbi
</right_context>
    </line>
    <line>
      <ref>szépirodalom</ref><left_context>Éppen elég lehetett nekik, hogy a Szovjetunióknak továbbra is „
</left_context><kwic> hitelbe </kwic><right_context> „, kellett szállítaniok.) Összegezve kiderült, hogy Magyarország
</right_context>
    </line>
    <line>
      <ref>sajtó</ref><left_context>pénzekből. Nem kizárt azonban, hogy a szakértői kamara a jövőben „
</left_context><kwic> hitelbe </kwic><right_context> „, nem dolgozik majd a rendőrségnek, azaz más megállapodási
</right_context>
    </line>
    <line>
      <ref>személyes-fórum</ref><left_context>gazdálkodó szervezetek által a költségvetésnek „ rövidlejáratú „
</left_context><kwic> hitelbe </kwic><right_context> adott pénz. Az önkormányzatok finanszírozásában szintén hallani
</right_context>
    </line>
    <line>
      <ref>sajtó</ref><left_context>mellett a környezet állapota. Az idén már szeptember 27-étől „ </left_context><kwic>
hitelbe </kwic><right_context> fogyasztja „, a Föld erőforrásait az emberiség, vagyis már feléltük </right_context>
    </line>
    <line>
      <ref>személyes-fórum</ref><left_context> „ megteremtését „, ügyesen „, későbbre halasztották, vagyis „
```

</left\_context><kwic> hitelbe </kwic><right\_context> „ dolgoztatták az érintetteket. speaker 02/27/99 13:00:15 Az  
</right\_context>  
</line>  
<line>  
<ref>személyes-fórum</ref><left\_context> jelentkezni. Ha gondolod, az én jegyemet is kifizethetnéd (  
</left\_context><kwic> hitelbe </kwic><right\_context>), mert valszeg egyszerűbb odamenni az INDEXesekhez és ott befizetni </right\_context>  
</line>  
<line>  
<ref>személyes-fórum</ref><left\_context>beszélgetés, piálás. Ez utóbbi rendkívül olcsó, önköltségi ár alatt,  
</left\_context><kwic> hitelbe </kwic><right\_context> is. Politizálás szinte semmi. Most még. Majd szép fokozatosan  
</right\_context>  
</line>  
<line>  
<ref>szépirodalom</ref><left\_context>Úgy vesztegette ezt a sok holmit potya pénzen, magára alkudva,  
</left\_context><kwic> hitelbe </kwic><right\_context> is adta, kínálta, doboltatta a falunak. Legisinkább persze  
</right\_context>  
</line>  
<line>  
<ref>beszéltnyelvi</ref><left\_context>Hobbyméhészek. A fiataloknak nem elég ez. Nincs pénz befektetni,  
</left\_context><kwic> hitelbe </kwic><right\_context> ne szabad menni, úgy hogy ez egy kicsit nehezebb indulunk,  
</right\_context>  
</line>  
<line>  
<ref>sajtó</ref><left\_context>forint) értékre elkövetett váltóhamisítással, banki csalással, </left\_context><kwic> hitelbe  
</kwic><right\_context> vett áruval való üzérkedéssel és biztosítási bűncselekményekkel </right\_context>  
</line>  
<line>  
<ref>személyes-fórum</ref><left\_context>benne minden pénzzel meg irattal. A benzinem persze elfogyott,  
</left\_context><kwic> hitelbe </kwic><right\_context> tankoltam és kétszer kidumáltam magam igazoltatásnál -- fokozott  
</right\_context>  
</line>  
<line>  
<ref>tudományos</ref><left\_context>tárgyaljanak. Szállást, bőséges vacsorát kaptak a fogadóstól,  
</left\_context><kwic> hitelbe </kwic><right\_context> még szerszámokat, élelmet (egy hétre) és két felszerelt öszvért  
</right\_context>  
</line>  
<line>  
<ref>sajtó</ref><left\_context>postásnő. - Nem, írja csak: „ Etel a Hotelbe, kapitány a fotelbe, </left\_context><kwic>  
hitelbe </kwic><right\_context> hatol be. Textiláru föl, selyemáru le, szőrmeáru összevisza </right\_context>  
</line>  
<line>  
<ref>sajtó</ref><left\_context> kénytelen beérni ilyesfajta ajánlatokkal, sőt terven fölül, </left\_context><kwic> hitelbe  
</kwic><right\_context> szállítani gázt - csak az idei év utolsó negyedében 5 milliárd </right\_context>  
</line>  
<line>  
<ref>beszéltnyelvi</ref><left\_context>tehát a hírekben. De hogy ez mit jelent tulajdonképpen hitelből,  
</left\_context><kwic> hitelbe </kwic><right\_context>, de miért? A kérdésre Duronelli Péter a Budapest Alapkezelő  
</right\_context>  
</line>  
<line>  
<ref>sajtó</ref><left\_context> a kertben, melyek a kedvenc cicák emlékét őrizték. Hozomra, </left\_context><kwic>  
hitelbe </kwic><right\_context>, ingyen dolgoztak a budakeszi iparosok, ha valami földanivaló </right\_context>  
</line>  
<line>  
<ref>szépirodalom</ref><left\_context>mortgage, magyarul? NŐ Nem tudom - VENDÉG Maguk is így élnek itt,  
</left\_context><kwic> hitelbe </kwic><right\_context> - az egész ország hitelbe él - hát Amerikában is - csak ott  
</right\_context>  
</line>  
<line>  
<ref>sajtó</ref><left\_context>gatyában, ingben, mezítláb a bolt pultjánál, és pálinkát ittak, </left\_context><kwic> hitel-  
be </kwic><right\_context>. Öten voltak, öt napszamos, és úgy döntöttek, elmennek Amerikába </right\_context>  
</line>  
<line>  
<ref>sajtó</ref><left\_context>meg. Ezért, ha totálkáros a gépkocsi és a kis kezdőrészlettel, </left\_context><kwic> hitel-

```

be </kwic><right_context> vásárolták, akkor a kártérítés az első néhány évben nem fedezi </right_context>
</line>
<line>
<ref>sajtó</ref><left_context>megborítani a családi költségvetést és szükségtelen kiadásokba, </left_context><kwic>

```

## ÷ 2. KWIC/mondat

A mentési lehetőségeket előhívó opció alatt találjuk meg a megjelenítési beállításokat. A „KWIC/mondat” opcióra kattintva lehetővé válik két megjelenítés közti választás:

Megjelenítés  
KWIC/mondat

1. a **KWIC** annyi mint „**key word in context**” típusú megjelenítés; ha ezt választjuk, akkor megszabadhatjuk, hogy mekkora kontextussal kívánjuk a célformát megjeleníttetni (de maximum 1000–1200 karakteres kontextus fog megjeleneni);
2. a másik a „**mondatkontextus**”: ha ezt a formát választjuk, szebb, rendezettebb lesz ugyan az összkép, viszont így csak egyetlen mondat fog megjeleni kontextusként, akkor is, ha a mondat nagyon rövid.

Nézzük meg a kétféle beállítás közötti különbséget!

## ÷ A) KWIC

Ha a *láthatóak* szóalakra keresve a „KWIC” megjelenítést választjuk, ez a kép tárul elénk:

The screenshot shows a search interface with the following elements:

- Search bar:** "keresés" with a search icon and a link to "találatok száma: 5615".
- Filters:** "oldalszám: 1 / 4" with buttons for "Ugrás", "Következő", and "Utolsó".
- Left sidebar:** A list of filters including "Mentés", "Megjelenítés", "KWIC/mondat", "Rendezés", "bal | jobb találat", "Info", "Keverés", "Véletlen minta", "Szűrés", "Gyakoriságok", "szóalakok", "szótövek", "elemzések", "dokumentumok", "Kollokációk", "Visszaállítás", "Próbálja ki a Mazzola eszközt ígék és argumentumaim közöttlen vizsgálathoz!".
- Main content:** Search results for "láthatóak". The first result is from "BALOGH BERTALAN Fény és árnyék... Két pólus. Szobrász barátomtól értem meg, mennyire nélkülözhetetlen az árnyék jelenléte. Egyszerűen nem volnának láthatóak a térbeli formák, ha minde kisseri halál...". The text is displayed in a KWIC format, with the search term "láthatóak" highlighted in red. The text is truncated on both sides to fit the KWIC format.

Itt most viszonylag szerény kontextust állítottunk be, amely a találati felületen megjelenítő-dik (150 karaktert), de még így is használnunk kell a csúszkát ahhoz, hogy az egész sort el tudjuk olvasni, legalábbis ha szabad szemmel jól olvasható betűnagyságot akarunk (a betűméret csökkentésével több fér rá a képernyőre). A KWIC-beállításnál ugyanis nincs sortörés, annak érdekében, hogy a célformák pontosan egymás alatt lehessenek. Ez azonban a kontextus tanulmányozásához nem túl kényelmes megoldás.



Fontos tudni, hogy nem kell megelégednünk az így megjelenő kontextussal, hanem a **pirossal jelölt célformára rákattintva** alul megjelenik egy **új ablak**, melyben a bal felső és alsó sarkokban található + jelekre kattintva a **kontextus növelhető**. (A + jel alatt található – jel nem használható semmire.)

Ez az ablak a – jelle kattintva eltüntethető. Sajnos utána **semmi sem jelzi, melyik sorban voltunk**, így türelmet próbáló feladat újra megtalálni az éppen vizsgált sort a találati felületen. Ez egy a számos kényelmetlenség közül, amivel az MNSZ2 használata során szembesülünk.

The screenshot shows the MNSZ2 search interface. At the top, there's a search bar with 'gramma' entered and 'MNSZ2' as the corpus. Below this is a list of search results. Each result has a title, a snippet, and a 'Keresés' button. The first result is 'BALOGH BERTALAN Fény és árnyék...'. The search results are displayed in a table-like format with various filters on the left and a search bar at the top.

A képen a legelső találat ablaka van megnyitva, nagyobb kontextusban (miután rákattintottunk a fönti és a lenti pluszjelre).

Ha a kékkel szedett **forrásadatra** (a képernyőképeken látható beállítás az alkorszak tünteti fel) kattintunk rá, megtaláljuk a forrásközlés részleteit (pl. a szépirodalmi alkorszak esetében a szerző és a mű címét, amelyből az idézet való). Ez az ablak a – jelle kattintva itt is eltüntethető.

This screenshot is similar to the first one, but the context menu for the first search result is expanded. It shows details about the source, including the author 'BALOGH BERTALAN', the title 'Fény és árnyék...', and other metadata like 'doc.region' and 'doc.type'. The context menu is a small window that appears over the search results, providing more information about the document being viewed.

Amint a „doc.file” sorában olvashatókból látszik, a megoldás itt sem éppenséggel olvasóbarát: a szokásos bibliográfiai adatok helyett csupán kódokat kapunk.

Ennél súlyosabb probléma az információk megbízhatósága.

Erre nézve nincsenek tapasztalataink, mert nem szoktuk a besorolások hitelességét megkérdőjelezni, az épp bemutatott felületen azonban felettébb feltűnő, hogy az első két találat azonos: szerzője is azonos (Balogh Bertalan), mindkettő az *Oszlopszent* című munkából való, de ezek más-más dokumentumazonosítóval vannak ellátva az adatbázisban.

Ennél nagyobb probléma, hogy a két találatnak eltérő a régió szerinti besorolása: a bemutatott esetben erdélyi magyarként van feltüntetve, a másik esetben szlovákiai magyarként.

Ha utánanézzünk az interneten, az *Oszlopszent* című művet író Balogh Bertalan önéletrajzából kiderül, hogy Balogh Bertalan Magyarországon (!!!), Újpesten született 1935-ben, majd a második világháború után Olaszországba disszidált, aztán 14 évig Svédországban élt, később pedig az Egyesült Államokban telepedett le, ahol a mai napig él. Nem tudunk meg semmit valamiféle erdélyi vagy felvidéki kötődésről.

Rejtély, hogyan kerülhetett bele az MNSZ-be a magyarországi–svédországi–amerikai Balogh Bertalan *Oszlopszent* című műve erdélyiként és szlovákiaiként.

Az ilyen hibák miatt nagyon fontos tüzetesen vizsgálni a találatokat, különösen az alacsony elemszámú vizsgálatok során.

Ennyit a „KWIC” beállításról, most nézzük meg a „Mondat” típusú beállítást!

## ÷ B) Mondat

A „Mondat” típusú beállításnál a konkordancia **minden találatához csak azt a mondatot rendel hozzá, amelyben a célforma előfordult**. Ebből következően az egyes találatok kontextusa meglehetősen **változó**, egy esetleges tényezőtől, a mondat hosszától függő.

MNSZ2	
Felhasználó: <b>gramma</b>	korpusz: MNSZ2
>> v2.0.5 - friss verzió 2018. június 6-tól << <a href="#">Hírek</a> <a href="#">Súgó</a>	
<b>Keresés</b>	korpusz: MNSZ2 találatok száma: <b>5615</b> oldalszám: <input type="text" value="1"/> / <input type="text" value="4"/> <a href="#">Ugrás</a> <a href="#">következő</a>   <a href="#">utolsó</a>
<b>Mentés</b>	<b>szépirodalom</b> Egyszerűen nem volnának <b>láthatóak</b> a térbeli formák, ha mindent beborító egyenletes fény világítana be a szobrot.
<b>Megjelenítés</b>	<b>szépirodalom</b> Egyszerűen nem volnának <b>láthatóak</b> a térbeli formák, ha mindent beborító egyenletes fény világítana be a szobrot.
<b>KWIC/mondat</b>	<b>szépirodalom</b> Az uszodában, magyarkénta tovább hírlko, teljes valójukban szinte egyáltalán nem <b>láthatóak</b> , ugyanis a kőörvben különböző színes hárttyákat növesztenek.
<b>Rendezés</b>	<b>szépirodalom</b> Kisebbség, nagyobbak, <b>láthatóak</b> , láthatatlanok.
<b>bal   jobb találat</b>	<b>szépirodalom</b> amilyenek egy éjszakai kertnek lennie kell az utolsó pillanatban, nagy, üres terek váltakoztak szép, sűrű bokrokkal, sűrűt a hold, s a kert gyepeire lékes, sugárvó fényvet bocsátott, csupán az árnyékok voltak fésztesen hosszúak és <b>láthatóak</b> , de talán, gondolta, az éjszaka títka kedvez más títoknak is, Drachos mosolygott, ellenkezett, tevékenykedett, dadogott, mommogott, habogott, gügyögött, babusgott, ayonbeszélte őt, ő boldog meg hagyta, hadd müljék az idő, föl kellett volna nyalabolnia Drachost és basta, úgy érte, és kész, de hát addig-addig, míg egyik pillanatról a másikra, mint egy rossz fuvalat, kivilágosodott, a tő felől jött a tömény, csirizes, világos massa, melyben mozdulni is nehéz, nem vitás, véget ért, véget
<b>info</b>	<b>szépirodalom</b> A megjelenés pillanatában <b>láthatóak</b> .
<b>keverés</b>	<b>szépirodalom</b> Mindenütt ott <b>láthatóak</b> , munkánál, menetnél vagy appelnél a sorban, amint ütemesen előre-hátra himbálózva mormolják magukban kifoghatatlanul, mint valami soha le nem törleszhető adósságot: az inadságukat.
<b>Véletlen minta</b>	<b>szépirodalom</b> Aztán megint az ónszínű Starnbergen reggelizünk, az Alpok nem <b>láthatóak</b> , a tő, a levegő szürke, fehér hattyúkkal pettyezve.
<b>Szűrés</b>	<b>szépirodalom</b> Az észcert már megmutatták egymásnak, a testi formátlanságok is <b>láthatóak</b> voltak, de a kékrókak és a nercok még nem.
<b>Gyakoriságok</b>	<b>szépirodalom</b> Meglepett, hogy a könyvek aránybetűl mégis oly tisztán <b>láthatóak</b> .
<b>szóalakok</b>	<b>szépirodalom</b> A fekete istennék, a Fürták vijjogva követtek, hogy rám húzassák kényszerzubbonyukat, tépőfoguk elől ide-oda cikázva mitoszokat kevertem a Bibliával, és mindkettőt a hétszónapjaimmal - elegendő életmű a foglalkozásszerű túléle, örökös rögtönzés, szemmel <b>láthatóak</b> a varratok, mint koponyámon a két műtét hege: költészetnek földszagú, de legalább minden hire szó szerint igaz.
<b>szótövek</b>	<b>szépirodalom</b> Barátágunk előzmény nélkül, egyszerre tört ki, mint valami betegség, és forrasztotta össze figuráinkat oly határig, hogy - miután mindig együtt voltunk <b>láthatóak</b> - akik nem ismertek bennünket, azt gondolták, testvérek vagyunk.
<b>dokumentumok</b>	<b>szépirodalom</b> Ezen kívül az életüket is veszélyeztetik, hiszen nem eléggé <b>láthatóak</b> a közúti forgalomban.
<b>Kollokációk</b>	<b>hivatalos</b> Az állampolgárok első alkalommal 1994-ben választhattak kisebbségi képviselőket, és az így létrehozott kisebbségi önkormányzatok működésének mintegy három éves tapasztalatai alapján már <b>láthatóak</b> e sajátos öngazdáltsági rendszer előnyei és hiányosságai is.
<b>Visszaállítás</b>	<b>hivatalos</b> Ennek eredményeképpen az oszloplagramokat tekintve, melyeken a számadatok is <b>láthatóak</b> , megállapítható, hogy csekély számban nőtt a budapesti, és csökkent a vidéki lnditványozók száma.
<b>Problálja ki a Mazzola eszközt igek és argumentumaiak közvetlen vizsgálatához!</b>	<b>hivatalos</b> AB-határozat), vagyis "a jog szabályai világosak, egyértelműek, hatásukat tekintve kiszámíthatóak és a címzettjei számára is előre <b>láthatóak</b> legyenek." (1/1995. AB-határozat).
	<b>hivatalos</b> Ennek jelei az épületek külső falain jól <b>láthatóak</b> , illetve még az utcán - az intézmény előtti járdán - is érezhetőek voltak.
	<b>hivatalos</b> Az országgyűlési biztos ajánlásában felhívta Budapest Főváros IV. Ker. Újpest önkormányzat polgármesterének figyelmét, hogy az önkormányzati rendelet előkészítése során, különösen a támogatási feltételek módosításakor az állampolgároknak a jogbiztonsághoz való jogát - hogy a jog egésze, egyes részterületei és egyes szabályai is világosak, egyértelműek, hatásukat tekintve kiszámíthatóak és a norma címzettjei számára előre <b>láthatóak</b> legyenek - minden esetben érvényesítse.
	<b>hivatalos</b> A jogbiztonság az állam s Alkotmánybíróság meghatározása szerint elsősorban a jogalkotó, de a megfogalmazásból eredően a jogalkalmazó köteletségévé is teszi annak biztosítását, hogy a jog egésze, egyes részterületei és az egyes jogszabályok is világosak, egyértelműek, működésüket tekintve kiszámíthatóak és előre <b>láthatóak</b> legyenek a norma címzettjei számára.
	<b>hivatalos</b> Ha önk. kinyitják az előterjesztés 2. oldalát, ott <b>láthatóak</b> azok az észrevételek, amelyeket a köztársasági megbízott tett még a tavalyi évben és amelyeket a jogi bizottság - az előző jogi bizottság munkáját folytatva - megtárgyalta és ideterjesztett a képviselő-testület elé.
	<b>hivatalos</b> A magas tartalékok képzése mögött egyelőre nem <b>láthatóak</b> a valós szándékok, azonban a tartalék-fejezetnél kiderül: a főváros a most tescsőg beruházási hullám után újabb hullámmra készülődik - következnek a Városháza 5-6 milliárd forintos szépítése.

A bemutatott képen jól látható a kontextusnak ez a különbözősége: van a találatok közt olyan is, amelyben a mondat mindössze négy szóból áll, de olyan is, amelyben a mondat négy sort tesz ki.

Nyilvánvaló, hogy ez a beállítás nem alkalmas a kontextusvizsgálatra, ám az anyag gyors áttekintését sikeresen elvégezhetjük ennél a beállításnál.

Bár a „Mondat” beállításnál egy találathoz mindig csak egy-egy mondat jelenik meg, tudni kell, hogy a pirossal jelölt célformára kattintva **ugyanúgy lehetővé válik a kontextus növelése, mint a „KWIC” beállításnál**, az ott leírtak erre a beállításra is érvényesek, amint ezt az alábbi kép is mutatja:

The screenshot shows the MNSZ2 search interface. The search term is 'láthatóak' and the results are displayed in a table with columns for 'szépirodalom' and the text snippet. The results include various literary and official documents, such as 'Egyszerűen nem volnának láthatóak a térbeli formák...' and 'Ezen kívül az életüket is veszélyeztetik, hiszen nem elege láthatóak a közúti forgalomban.'

A kép az első találat bővebb kontextusát mutatja a lent megnyílt szövegdobozban.

### ÷ 3. Megjelenítés

Eddig a „Megjelenítés” alatti „KWIC/mondat” opcióval foglalkoztunk. Most a különféle megjelenítési lehetőségeket nézzük meg.

Megjelenítés  
KWIC/mondat

Ha rákattintunk a „Megjelenítés” opcióra, eltűnik a konkordancia, és egy ilyen új ablak jelenik meg. Mivel a „struktúrák” menüpont alatt több opció van, mint ami egyszerre megjeleníthető, két képen mutatjuk be a lehetőségeket:

Felhasználó: **gramma** korpusz: MNSZ2

Keresés ?

Megjelenítés ?

attribútumok	struktúrák	infó
<input checked="" type="checkbox"/> word	<input checked="" type="checkbox"/> <doc>	<input checked="" type="checkbox"/> szó sorszáma
<input type="checkbox"/> lemma	<input checked="" type="checkbox"/> <div>	<input checked="" type="checkbox"/> dokumentumazonosító
<input type="checkbox"/> msd	<input checked="" type="checkbox"/> <p>	<input checked="" type="checkbox"/> doc.file
<input type="checkbox"/> ctag	<input checked="" type="checkbox"/> <s>	<input checked="" type="checkbox"/> doc.region
<input type="checkbox"/> ana	<input checked="" type="checkbox"/> <i>	<input checked="" type="checkbox"/> doc.style
<input type="checkbox"/> word_cv	<input checked="" type="checkbox"/> <lg>	<input checked="" type="checkbox"/> div.column
<input type="checkbox"/> word_syll	<input checked="" type="checkbox"/> <poem>	<input checked="" type="checkbox"/> div.type
<input type="checkbox"/> lemma_cv	<input checked="" type="checkbox"/> <head>	
<input type="checkbox"/> lemma_syll	<input checked="" type="checkbox"/> <title>	
<input type="checkbox"/> word_phon	<input checked="" type="checkbox"/> <opener>	
<input type="checkbox"/> lemma_phon	<input checked="" type="checkbox"/> <byline>	
	<input checked="" type="checkbox"/> <closer>	
	<input checked="" type="checkbox"/> <dateline>	

Attribútumok megjelenítése

minden szón

csak a találati szavakon

Oldalanként: 1500 sor

Kontextus mérete: 60 karakter

Megjelenítés megváltoztatása Beállítások mentése

Felhasználó: **gramma** korpusz: MNSZ2

Keresés ?

Megjelenítés ?

attribútumok	struktúrák	infó
<input checked="" type="checkbox"/> word	<input checked="" type="checkbox"/> <head>	<input checked="" type="checkbox"/> szó sorszáma
<input type="checkbox"/> lemma	<input checked="" type="checkbox"/> <title>	<input checked="" type="checkbox"/> dokumentumazonosító
<input type="checkbox"/> msd	<input checked="" type="checkbox"/> <opener>	<input checked="" type="checkbox"/> doc.file
<input type="checkbox"/> ctag	<input checked="" type="checkbox"/> <byline>	<input checked="" type="checkbox"/> doc.region
<input type="checkbox"/> ana	<input checked="" type="checkbox"/> <closer>	<input checked="" type="checkbox"/> doc.style
<input type="checkbox"/> word_cv	<input checked="" type="checkbox"/> <dateline>	<input checked="" type="checkbox"/> div.column
<input type="checkbox"/> word_syll	<input checked="" type="checkbox"/> <date>	<input checked="" type="checkbox"/> div.type
<input type="checkbox"/> lemma_cv	<input checked="" type="checkbox"/> <time>	
<input type="checkbox"/> lemma_syll	<input checked="" type="checkbox"/> <docAuthor>	
<input type="checkbox"/> word_phon	<input checked="" type="checkbox"/> <sp>	
<input type="checkbox"/> lemma_phon	<input checked="" type="checkbox"/> <speaker>	
	<input checked="" type="checkbox"/> <stage>	
	<input checked="" type="checkbox"/> <g>	

Attribútumok megjelenítése

minden szón

csak a találati szavakon

Oldalanként: 1500 sor

Kontextus mérete: 60 karakter

Megjelenítés megváltoztatása Beállítások mentése

Az eddig bemutatott találati képek a most bemásolt képeken látható beállításokat tükrözik. Más lehetőségeket is ki lehet próbálni; alább azokat mutatjuk be, amik a munkánkhoz a leghasznosabbnak tűnnek.

Az „**attribútumok**” címszó alatt található lehetőségek közül érdemes változatlanul hagyni az alapbeállítottat, a „**wordöt**”, azaz a szót. Azzal, hogy az attribútumok minden szón vagy csak a találati szavakon (célformákon) jelenjenek meg, nem kell foglalkozni, maradjon az alapbeállítás („csak a találati szavakon”).

A „**struktúrák**” menüpont alatt különféle méretű szerkezeti egységeket tehetünk láthatóvá. A bekezdések eleje (<p>) és vége (</p>) az alapértelmezett, de beállíthatjuk például a mondat- vagy a dokumentumhatárok megjelenítését is. Mivel nekünk ezekre a kódokra ritkán van szükségünk, legjobb, ha egy nagy egységet jelenítünk meg, a **dokumentumot**, vagyis a **<doc>** opciót választjuk.

A „**struktúrák**” alatt állíthatjuk be azt is, hogy az **írásjelek előtt ne legyenek szóközök**. Ehhez az utolsó opcióra, a **<g>**-re kell rákattintani („g” mint *glue* ’ragasztó’, azaz tapadnak az írásjelek). **A szokásos beállítás tehát az, hogy a <doc> és a <g> opciók legyenek kiválasztva**. Ahhoz, hogy kettőt lehessen kiválasztani, a Ctrl-billentyűt megnyomva kell tartani.

Az „**infó**” menüpont segítségével meghatározhatjuk, hogy a konkordanciasorok elején milyen információ jelenjen meg. Az eddigi találati képeken csak a **stílusréteg (doc.style)** volt kiválasztva, ez jelent meg kék színnel a találati szövegrészletek előtt, azoktól balra. Némely

szónál jelentős eltérés van a magyarországi és a határon túli találatok relatív gyakorisága közt; ilyenkor hasznos a régió (doc.region) kiválasztása is. Ezenkívül sok esetben célszerű a **dokumentumazonosítót** is megjeleníteni, mivel ebből azonnal látjuk, hogy a következő sorban lévő adat még ugyanabból a dokumentumból származik-e, vagy már egy másikból. (Az egynél több opció kiválasztása itt is a Ctrl-billentyű megnyomásával történik.)

A dokumentumazonosító megjelenítése nagyon jó szolgálatot tesz a példák kiválogatásában, mert az az egyik fontos elvünk, hogy ha van elég adatunk, egy szerzőtől csak egyetlen példát idézünk (lehetőleg egyetlen művéből egyetlen példát, de ha kevés példánk van, vagy valamilyen okból több példa is nagyon szemléletes, figyelmen kívül hagyhatjuk ezt az elvet).

Az alábbi képernyőkép azt az esetet mutatja, amikor az „infó” menüpontnál a föntebb jelzett három lehetőség van beállítva (1. dokumentumazonosító, 2. régió, 3. stílusrétteg):

Felhasználó: **gramma** korpusz: MNSZ2 >> v2.0.5 - friss verzió 2018. június 6-tól << Hírek Sűgő

<b>Keresés</b>	korpusz: MNSZ2 találatok száma: 253	
<b>Mentés</b>	<a href="#">doc#2011,magyarországi,szépirodalom</a>	S olyan volt ez a bosszús, ősi reggel, mint egy bozontos, <b>morgó</b> kutya, amelyik csatát veszített az éjjel, s most saját mérgétől sötétül a szőre.
<b>Megjelenítés</b>	<a href="#">doc#1015,magyarországi,személyes-fórum</a>	Hanem mint az ordas, ki farkára lépett vágnál Te is felém nagy ártatlan képet s nem vonyit, nyüzszöge jön felém a farka, jelezvén szemével - szívével is falna Várják parancsokat dehogy acsarognak mellőlem - soha már - el nem csavarognak És ilyen mindegyik! öreg vagy fiatal, jóképű vagy randa - nem tudja mit akar Mikor jószándékkal egyszer rájuk nézek rettegnek hogy velem mély verembe lépnek S akármilyen kedves lehetek veletek, automatizmus lesz abból, hogy szeretek <b>morgó</b> futószalag, csikorgó gépezet, s nem látják mögöttem a ténykedő kezet Pedig amit adott, vissza tudja fogni s félév múlva elkezd mindegyik nyafogni Hanem ami elkelt vissza én nem veszem, Úgy tudok feledni, hogy észre sem veszem Hogyha egyik elmegy, jön helyette majd más Ha tudsz akkor tanulj, Te is ebből pajtás Nem fogom mi eltört újból összerakni későn érkezéneked nincs már mit aratni.
<b>Rendezés</b>	<a href="#">doc#312,magyarországi,szépirodalom</a>	Királyfi áttülekszik az egyre hevesebben <b>morgó</b> ügyfeleken az első íróasztalhoz.
<b>bal   jobb találat</b>	<a href="#">doc#1192,magyarországi,tudományos</a>	Coboly nyikorgó hangjára alszom, medve <b>morgó</b> hangjára ébrednek, szarvas köhhögő hangjára ébrednek.
<b>infó</b>	<a href="#">doc#2761,magyarországi,személyes-közösségi</a>	Vagy legyél olyan mint, Kasperov a sakkozó ne legyél <b>morgó</b> a hét törpé#x00e8;ből Hajnik&#x00e0;m!!!
<b>keverés</b>	<a href="#">doc#2321,magyarországi,sajto</a>	A németországi Hamburgban férfigérgőző nyílt, ahová a nők betehetik <b>morgó</b> partnerüket és nyugodtan elmehetnek vásárolni.
<b>Véletlen minta</b>	<a href="#">doc#355,magyarországi,szépirodalom</a>	Házuk mélyében egy <b>morgó</b> zongora fogta fogva.
<b>Szűrés</b>	<a href="#">doc#1942,magyarországi,szépirodalom</a>	Olyan voltál, mint egy emberré varázsoltt, jó medve, csúnya, tagbaszakadt, <b>morgó</b> , a lelked már emberé, a riadalmaid meg az ösztöneid hajdani medve voltodból maradtak.
<b>Gyakoriságok</b>	<a href="#">doc#2470,magyarországi,sajto</a>	A huszonnégy kutyából, amelyeknek kutya képet vetítettek, hűz jól azonosította be a <b>morgó</b> kutya méretét, és rögtön a helyes képre nézett, és tovább is figyelte azt.
<b>szóalakok</b>	<a href="#">doc#1866,magyarországi,szépirodalom</a>	Most a Szentatya ellép a <b>morgó</b> fegyverek közt, torz áldást osztogatnak méltatlan kezei.
<b>szótövek</b>		
<b>elemzések</b>		
<b>dokumentumok</b>		
<b>kollokációk</b>		
<b>Fredei konk</b>		

A következő beállítási lehetőség az, hogy hány sor jelenjen meg **oldalanként**. Ha nincs kedvünk sokat lapozgatni, akkor nagyobb mennyiséget állítunk be, akár **1200–1500 sort** is. Ez még nem lassítja meg annyira a találatok betöltését, hogy érdemes lenne kevesebbet választani, és aztán lapozgatni.

Az utolsó megjelenítési beállítás segítségével megszabhatjuk, hogy a konkordanciában milyen legyen a „**kontextus mérete**”. Bármilyen szám feltüntethető, de ténylegesen csak 1000–1200 karakter fog megjelenni, akkor is, ha 5000-et írunk be. Ha 1200-at írunk be, a maximális kontextust kapjuk ugyan, de ezt a felületen nagyon kényelmetlen olvasni, mert nincs sortörés, a képernyőn pedig annyi szöveg nem fér el. Ezért **az egyes találatokat a célformára való rákattintás után lent megnyíló ablakban érdemes olvasni**. A másik lehetőség, hogy a **konkordanciát elmentjük**, megnyitjuk Wordben vagy belemásoljuk egy Word-dokuba, átalakítjuk nekünk tetsző formátumba, és ott olvassuk.

A találatok egész kontextusának elolvasására a **célforma stílusértékének megállapításához** van szükség. Ha a célforma környezetében például inkább informális szavak vannak, akkor esély van rá, hogy a célforma informális vagy közömbös stílusértékű. Ha inkább formális szavak vannak a célforma környezetében, akkor a célforma is formális vagy közömbös stílusértékű lehet. Ha régies a szöveg, akkor elképzelhető, hogy a célforma is régies vagy elavult.

Ha csak a változtatásoknak a konkordancia megjelenítésére gyakorolt hatását akarjuk megnézni, akkor elég a kilépés előtt a „**Megjelenítés megváltoztatása**” gombra kattintani. Ha a

beállításokat véglegesnek tekintjük, kattintsunk rá a „**Beállítások mentése**” gombra, majd lépünk ki a bal oldali menüsorban található „**Konkordancia**” menüre való kattintással, vagy pedig a „**Megjelenítés megváltoztatása**” gombra kattintással.

## ÷ 4. Rendezés

A bal oldali menüsor következő pontja a „**Rendezés**”. A legfontosabb rendezési szempontok akkor is érvényesíthetők, ha erre nem kattintunk rá, mert ezek a „Rendezés” opció alatt külön-külön is fel vannak tüntetve:

- 1.1. a találati szó előtti (bal) szó szerint,
- 1.2. a találati szó utáni (jobb) szó szerint,
2. a találati szó szerint,
3. a sor elején látható infó szerint,
4. rendezetlen megjelenítés (keverés).

**Rendezés**  
bal | jobb  
találat  
infó  
keverés

Ezek a lehetőségek arra valók, hogy meghatározzuk, **mi szerint szeretnénk betűrendbe tenni** (ha szeretnénk) a megjelenő konkordanciasorokat (szövegrészleteket).

Ha a lexéma különböző toldalékokkal ellátott formáiból szeretnénk példákat venni (és általában erre van szükségünk), akkor érdemes a **találati szó szerinti rendezés** lehetőségét választani (2. sz. lehetőség). Akkor először a toldaléktalan formák jelennek meg, majd az egyes toldalékok – elvileg – betűrendben.

korpusz: MNSZ2 találatok száma: 32489 - ebből vett 30 elemű véletlen minta		
tudományos	1715-ben Felsőzsemberben 4 kúria és 32 háztartás, Alsózsemberben 15	<b>háztartás</b>
beszélnyelvi	Természetesen, kevés a pénz. Szerintem ez úgy működik, mint egy	<b>háztartás</b>
személyes-közösségi	Környezetvédelem Energia megtakarítás Egészségvédelem Vegyszermentes	<b>Háztartás</b>
tudományos	napig " Janyicsárok" hívják. 1715-ben szőlőskertje és 27	<b>háztartása</b>
tudományos	század közepéig a Perneszi család bírta. 1848-ban 109 úrbéres	<b>háztartása</b>
beszélnyelvi	eljutott oda, hogy két-három, négy kukába gyűjt ugye minden	<b>háztartása</b>
személyes-közösségi	2013-09-26 UPC Magyarország Kedves Renától Szeretnénk minél több	<b>háztartásba</b>
sajtó	gyermeknapon teljesíteni akarja csemetéje kívánságát. Félmillió	<b>háztartásban</b>
tudományos	a felsőoktatásban is, a nők egy része pedig otthon maradt a	<b>háztartásban</b>
személyes-közösségi	háztartásfővel élők körében, valamint azok között, akik olyan	<b>háztartásban</b>
személyes-közösségi	tényleg van olyan ahol van 3 vagy akár 4 TV készülék is egy	<b>háztartásban</b>
személyes-közösségi	amennyit szívesen elvegne és örömet okoz. Hisz vannak azért a	<b>háztartásban</b>
személyes-közösségi	amennyit szívesen elvegne és örömet okoz. Hisz vannak azért a	<b>háztartásban</b>
sajtó	központjában. A januári 24, 9 százalékos áramáremelést a honi	<b>háztartások</b>
sajtó	hasonló időszakhoz képest. Valamilyen biztosítást az összes	<b>háztartások</b>
hivatalos	terméktípusok képesek -e nagyobb piaci részesedést nyerni [14]. 6.7 A	<b>háztartások</b>
hivatalos	foglalkoztatás megszűnésével mindenkélt azok a gyermeket nevelő	<b>háztartások</b>
sajtó	a Magyar Nemzeti Bank pénteken közzétett statisztikájából. A	<b>háztartások</b>
sajtó	különbség a megtakarítások időhorizontjában ugyan, de a külföldi	<b>háztartások</b>
sajtó	Európa-szerte több mint 188 millió 10 évnél idősebb készülék van a	<b>háztartások</b>
tudományos	népszámlálási adatok szerint Aberdeén lakónépessége 1603 fő, a	<b>háztartások</b>
sajtó	szepetembere és novembere között regisztrációs űrlapokat küldenek ki a	<b>háztartásokba</b>
sajtó	marad a kutya, a macska, a tengerimalac, a kanári, és egyes	<b>háztartásokban</b>
szépirodalom	női szerepet, ő egy tréfával ütötte el: - Ez úgy van, mint a	<b>háztartásokban</b>
személyes-közösségi	forintra tartották a létminimumot Magyarországon az egyszemélyes	<b>háztartásokban</b>
sajtó	évente elszámol az adófizetőkkel (a település honlapján vagy a	<b>háztartásoknak</b>
tudományos	közzé tette, hogy több mint azonos nemű párok által alkotott	<b>háztartást</b>
személyes-közösségi	gyermekemmel is foglalkozni, ha leteszem a kicsit aludni, de a	<b>háztartást</b>
sajtó	venni. Gyest igényelhet a szülő, a nevelőszülő, a gyám a saját	<b>háztartásában</b>
sajtó	is). Sir Elton Johnnak is lesz miből fenntartania költséző	<b>háztartását</b>
sajtó		
	volt. 1828-ban Felsőzsemberben 100 házbán 614 lakos, Alsózsemberben	
	, hogy annyit kell költeni, amennyi van. Azzal kell számolni	
	Mezőgazdaság: Bio Termelés Állattartás: elhullás csökke...	
	volt. 1787-ben 120 házbán 920 lakos élt. Lakói mezőgazdasággal	
	volt, ekkor Somogy nagyobb települései közé tartozott. 1930-ban	
	a hulladékok, a hasznosítható hulladékok, de nyilván ez a jövő	
	elérhetővé tenni a szolgáltatásink, viszont jelenleg nincs	
	a szülők megengedhetik maguknak az ajándékozást, 160-200 ezer	
	, vagy segit a családi vállalkozásban. Ők a régi szocialista	
	élnék, ahol a háztartásfő munkanélküli vagy inaktív. A parlamenti	
	. És ott akkor bizony korlátozza azt a jogot, hogy a 3. és 4	
	is olyan munkák, amik örömet okoznak, ha másért nem, mert a	
	is olyan munkák, amik örömet okoznak, ha másért nem, mert a	
	negyede nem tudja megfizetni. A baj orvoslására Karl Imre energetika	
	közel kétharmada kötött, és arányuk alig változik - tarták	
	és a kifizasztók az energia mintegy 30 %-át igénylik. Az	
	kerültek dráma helyzetbe, ahol a háztartásfő munkanélküli	
	összesített hitelállománya ugyanakkor közel 200 milliárd forinttal	
	portfoliójában nagyobb arányt képviselnek a részvények, kötvények	
	használatában, melyeket a legújabb technológiát alkalmazó gépekre	
	száma 689 és 436 család él itt a faluban. A népsűrűség 437,7	
	, ezek visszaküldésével lehet felkerülni a névjegyzékbe. Az	
	a pithon. (Ahol még nem nyelte le a gazdit.) A kutya viselne	
	szokott lenni, hol ő mosogat és én törölgetek - hol fordítva	
	. 2013-09-23 30 perc alatt 20 szegedi gól ötödik meccs is	
	elküldött kis információs füzet útján). A 30.000 korona fölötti	
	detektált az Egyesült Államok területén. Azonban, egyéb felmérések	
	is el kell látni, és főzőm is kell. A fiam panaszkodott, hogy	
	nevelt gyermek 3. életévének betöltéséig, ikergyermekek esetén	
	- az Elton John's Greatest Hits: 1970-2002 arany-, illetve	

A képen a *háztartás* szótőre adott találatokból létrehozott 30 elemű véletlen minta látható.

Az egyes toldalékos alakok jól áttekinthetőek, bár a betűrendet megzavarja, hogy a toldalékos betűket a program a toldaléktalanok után sorolja (ezért a *háztartása* után nem a *háztartásában* következik, mint ahogy kellene, hanem a *háztartásba*, s a *háztartásában* csak a *háztartást* után következik).

Amikor dolgozatot írunk, abban az idézett példák ne betűrendben legyenek, hanem a szokásos sorrendben: először egyes számú formák, aztán többes számú formák, az igéknél az egyes számún belül 1., 2., 3. személy, ugyanúgy a többes számún belül; először a kijelentő mód, aztán a feltételes, majd a felszólító stb.

Az anyag tisztításakor a találatokat a találati szó előtti **bal** vagy **jobb oldali szó** szerint kell rendeznünk (1.1. sz. lehetőség), mert akkor egymás alá kerülnek az azonos találatok.

személyes-közösségi	labdarúgó-bajnokság zárófordulójának Atletico Paranaense-Vasco da Gama	találkozójáról	, amely a drukkeriek összecsapása miatt több mint egy órát állt
személyes-közösségi	labdarúgó-bajnokság zárófordulójának Atletico Paranaense-Vasco da Gama	találkozójáról	, amely a drukkeriek összecsapása miatt több mint egy órát állt
személyes-közösségi	szombaton megrendezték a második magyarországi Volksvagen Golf	találkozót	a Hungaroringen, amely idén a Golf... " dc: creator= " Super
személyes-közösségi	5 gól a ponthoz Rendkívül izgalmas és többségében gólzaporos	találkozót	hozott a férfi futball NB I 13. fordulójá hétfőn, hiszen az
személyes-közösségi	rendezték Angliában, a mindig kiélezett és parázs hangulatú	találkozón	ezúttal már az 5. percben megszerezte a vezetést a forduló
személyes-közösségi	legfontosabb eseményével, az európa kupával fogunk játszani. Ma hat	találkozó	fog kezdődni este nyolc órakor. Típpeljétek meg hogy ezek közül
személyes-közösségi	ellen, hogy több szektort is le kell zárnia a klubnak a hazai	találkozón	. " A Ferencváros az első olyan magyar klub, amely kitilt szurkolókat
személyes-közösségi	A nagy meccseken, mint az ETO, vagy a Videoton elleni hazai	találkozón	sem tudunk igazán jó teljesítményt nyújtani. Támadni fogunk
személyes-közösségi	petrencerúddal tverbar; A húsz csapat részvételével megrendezett hentesi	találkozón	kétszázal is többen vettek részt. 2013-08-19 Jól rajtoltak
személyes-közösségi	Sox otthonában játszzák az első két összecsapást, majd három	találkozóra	kerül sor St. Louisban (már amennyiben mindhárom szükséges)
személyes-közösségi	útfelújításokkal kapcsolatban Gazdafórum - Gyomaendrődön tartották meg idei	találkozójukat	a gazdák, ahol kiértelmezhetők az elmúlt időszak eredményeit
személyes-közösségi	Egerszalókon tverbar; Szerdán kezdődött a 31. Egerszalóki Ifjúsági	Találkozó	és LelkiGYakorlat, amely több évtizede magyar katolikus fiatalok
személyes-közösségi	Egerszalókon tverbar; Szerdán kezdődött a 31. Egerszalóki Ifjúsági	Találkozó	és LelkiGYakorlat, amely több évtizede magyar katolikus fiatalok
személyes-közösségi	primás, esztergom-budapesti érsek a Hungaró nemzeti Ifjúsági	Találkozó	záróeseményén. A Rio de Janeiro-i Katolikus Ifjúsági Világtalálkozóval
személyes-közösségi	Augusztus 16-a és 19-e között rendezték meg a Zenei Ifjúsági	Találkozó	és Nemzetközi Fúvószenekari Versenyt. 2013-08-20 Balesetek
személyes-közösségi	olvastam - most azonban direktben is kérem. A Közösség többször is	találkozót	tartott, melyben csak úgy fürdöttünk a nagy találkozásokban
személyes-közösségi	elszámazottakat tverbar; Az önkormányzat augusztus 31-ére ismét	találkozóra	hívja a Kenderesről és Bánhalomról elszámazottakat, tudtuk
személyes-közösségi	ötzör lépett pályára, gólt nem szerzett, a Serie B-ben kilenc	találkozón	játszott, itt sem volt eredményes. A támadó tavasszal a harmadosztályú
személyes-közösségi	Magyarországról és más országokból, hogy tanúi lehessenek a királyok	találkozójának	és a lovagi tornának. Mint azt a heroldok már jó előre kihirdették
személyes-közösségi	fordulójában az osztrákok voltak az ellenfelek. A kiélezett	találkozó	a hajrában dönt el, amikor a mieink számára felcsillant a győzelem
személyes-közösségi	fejlesztésekről és a fenntarthatóságról A Kelet-Magyarország TOP 100 Klub	találkozójának	a Dunapack Kft. nyíregyházi központja adott otthont. 2013-11-10
személyes-közösségi	fejlesztésekről és a fenntarthatóságról A Kelet-Magyarország TOP 100 Klub	találkozójának	a Dunapack Kft. nyíregyházi központja adott otthont. 2013-11-10
személyes-közösségi	Brüsszelbe, az európai unió tagországok állam és kormányfői	találkozója	a Dunapack Kft. nyíregyházi központja adott otthont. 2013-11-23
személyes-közösségi	mérkőzés legjobbjá. A védekezésünk ezen a meccsen, és a korábbi	találkozón	. Ezeknek a javaslatoknak az lett volna a céljuk, hogy a lehető
személyes-közösségi	egy csapat rendre hitehagyottan lép pályára a kulcsfontosságú	találkozón	fantasztikus volt, Mathieu Flamini érkezésével még tovább erősödöttünk
személyes-közösségi	hasát, mint a kutya 2013-10-20 plaa. Meg volt a mai kutya	találkozó	, az mindenképpen edzői hiba. Vagy azért, mert a szakvezető
személyes-közösségi	történelem meg a csoportban. A harmadik alkalommal, azaz két	találkozó	is. És teljesen véletlenül Veie is összefutottam: D 2013-10-20
személyes-közösségi	helyettesítő Keith Tandy nincs a helyzet magaslatán. A legutóbbi két	találkozón	után sorra arról számoltak be a nők, hogy változott az életszemléletük
személyes-közösségi	Navracsics Tibor elmondta, jelezte az ír kollégájának, hogy kétoldalú	találkozót	, amikor vele kezdett a Buacs, összesen 58 pontot nyalt be a
személyes-közösségi	Navracsics Tibor elmondta, jelezte az ír kollégájának, hogy kétoldalú	találkozót	tart szükségesnek az ügyben az október 7-én tartandó bel- és
személyes-közösségi	kedves családiás vásár volt! Köszönjük a szervezőknek, a kézműves	találkozót	tart szükségesnek az ügyben az október 7-én tartandó bel- és
személyes-közösségi	programszorozat és közösségének legfontosabb fóruma a nyár végi közéleti	találkozó	megjelenteket a sok szép kínálatot a vásárlók érdeklődését
személyes-közösségi	egyetértettünk, hogy a Karma törvényei nem ismernek tréfát. A következő	találkozó	. Idén augusztus 21. és 25. között rendezik meg. A közeléti
személyes-közösségi	öztündíjasok, kutatók találkozásának az elősegítése. A következő	találkozónk	Augustus 22 -re 17h -ra tervezjük, A Szoboszlói Bárdos Lajos
személyes-közösségi	tehát, hajrá, ez a program pont állandó elem lesz a következő	találkozónk	október 26-án, szombaton 20 órakor kezdődik a Rómái Magyar
személyes-közösségi	kultúra frontján is megállja a helyét. A Szepezdí Köztisztviselő	találkozónk	apró változtatásokkal. Az est utolsó harmadában pedig ismét
személyes-közösségi	kultúra frontján is megállja a helyét. A Szepezdí Köztisztviselő	találkozóra	Arany János: A walesi bárdok című balladáját állítják színpadra
személyes-közösségi	kultúra frontján is megállja a helyét. A Szepezdí Köztisztviselő	Találkozóra	Arany János: A walesi bárdok című balladáját állítják színpadra

A képen a *találkozó* szótőnek a „személyes-közösségi” alkorpusz találataiból vett 250 elemű véletlen minta egy részét látjuk; a találatokat a bal oldali szó szerint rendeztünk.

Felületes szemlélet során is feltűnik, hogy csak ebben a kis részletben két helyen három-három, további három helyen két-két azonos találatot látunk.

Statistika készítésekor szükség lehet rá, hogy az ismétlődő találatokat kiküszöböljük, különösen olyankor, amikor kevés találatot kapunk. Ezt nevezzük a találatok tisztításának; a jegyzet későbbi fejezeteiben látunk majd példákat az adatok tisztítására.

Ha azt akarjuk, hogy a **lehetséges kollokációk**, ill. **állandósult szókapcsolatok** egymás alá kerüljenek, ugyancsak érdemes a **találati szó előtti bal vagy jobb oldali szó** szerint rendezni a találatokat.

Az alábbi kép a *búcsút* szóalak találataiból vett 50 elemű minta találatainak egy részét tartalmazza, a célszó utáni jobb oldali szó szerint rendezve: látjuk, hogy ránézésre kiemelkednek a *búcsút int*, *búcsút mond* és *búcsút vesz* szókapcsolatok:

szépirodalom  
személyes-közösségi  
sajtó  
sajtó  
sajtó  
sajtó  
sajtó  
tudományos  
szépirodalom  
tudományos  
sajtó  
szépirodalom  
beszélnyelvi  
sajtó  
sajtó  
sajtó  
sajtó  
szépirodalom  
szépirodalom  
sajtó  
tudományos  
sajtó  
sajtó  
sajtó  
szépirodalom  
szépirodalom  
sajtó  
személyes-közösségi  
sajtó  
sajtó  
személyes-közösségi  
szépirodalom  
szépirodalom  
sajtó  
személyes-közösségi

ősz, csepergett, köd szállt le és bár még enyhéség uralodik, **búcsút** intek a vénasszonyok nyarának. Sötét, nem egészen jó minőségű ébredés halál volt. Veled fent voltam a napfényen, Most már **búcsút** intenek a rögök és a fák, a hajnali álommal eltűnök örökre, Bár, ráadásul most Szalóki Ági is jelen lesz, aki azóta már **búcsút** intett a kollégáknak. A hosszú pihenő előtt még négy koncertet lapnak adott - interjúban kilátásba helyezte, hogy a nyáron **búcsút** inthet az angol élvonalnak. Sporthircentrum Ha a sportvilág 1998-ban Kohl elvesztette a választásokat, kijelölt utódja így **búcsút** inthetett ennek az esélynek. Schäubele a CDU-elnökséget vizsgálódhatott készítették időtálló lemezeit, most, 70 év fölött hirtelen **búcsút** inthetnek a jogdíjakból befolyó összegeknek. Az EU-biztos hangsúlyozta szögekben és magasságokban állítható monitorálványának köszönhetően **búcsút** inthetünk a nyakban tomboló fájdalomnak. Arra az esetre pedig kedvező, kilátás pazar. Este hazautazás. Nehéz szívvel vettem **búcsút** kellett vennie tőle; a kaszárnyában aztán az a kellemetlen egyedül élt és még nem volt ötvenéves; másnap azonban mégis **búcsút** kellett vennie tőle; a kaszárnyában aztán az a kellemetlen a szelek, csak a sóhaj rezgetteti, egész messziről, mint egy **búcsút** lebegtető patyolatsebkeendő... Az a csodálatos, feshetetlen erősítette a Kalapácsosokat, azaz 12 évnyi szolgálat után vett **búcsút** Londontól. Hurstöt 75 ezer fontért szerezték meg a piros-fehérek Katséré karjában. A csillagasszony meglehetősen volna, hogy ő is **búcsút** mond a földnek, de nem tudott elszakadni a kunyhótól, ahol Több mint valószínű, hogy az egyik magyar érdekltségű klub **búcsút** mond a Premier League-nek, hogy ez ne a Bolton legyen, azért miniszteri biztatásra) Hegyi mégis belátta, jobban teszi, ha **búcsút** mond az Andrassy úti palotának. Egy-két elfecsérelt, tehetséges nemcsak e miatt forog a pénz körül a média. A miatt is, hogy **búcsút** mondanak az egy és ötköpejás érméknek. A legutóbbi közvélemény-kutatás bevételük tették lehetővé. Ha most elköltönek másra, örökre **búcsút** mondhatna a fávóras a 4-es metróra. Moskovits Péter, a pénzügyi szakos új esztendő jó kívánások elhangzottak, hát akkor **búcsút** mondott a háziának s nemkülönben a két legénynek is, akik társadalomból származik. Igaz, hogy még nagyon kora ifjúságában **búcsút** mondott annak. A komaságot azonban sok mindenben tartja vele Amolyan ezredvégi requiem lesz, melynek során újra átéljük, majd **búcsút** mondunk a huszadik századnak. - Milyennek tartja ezt az időszakot kiadott brévében, melyek szerint aki hittel hordja a Medált, **búcsút** nyer, ha a megszentelő kegyelem állapotában van (ez a katolikus Szébia-Montenegrótól is kikaptak (24-29), az Izland elleni 35-31-gyel csupán a **búcsút** sikerült széppé tenni. Az Európa-bajnok végül Franciaország április elején láthatja a szabadban a hófehér bocot. Ezek vettek **búcsút** Sir Edmund Hillarytól Ezek vettek búcsút Új-Zélandon Sir Edmund kolostor- és börtönudvara, ahol aznap a pálos rendi szerzetesek **búcsút** tartanak. Ehhez a szakrális eseményhez csatlakoznak a performansz-előadók szupermodellek fémjelzte Fashion Cafénak éppen Claudia Schiffer mondott **búcsút** tavaly szeptemberben. Nagy Gábor HVG 99/02. szám 1999. január 16. nézték, azok sajnos hamarosan kénytelenek lesznek tőle könnyes **búcsút** venni. Az ORIGO információ szerint a T-Kábel az éjszakai csatorna viszont ha már így alakult, szerdán mindenképpen győzelemmel kell **búcsút** vennünk egymástól. Bánki József [origo] lkszelt az MTK Pécsset kifizető. A közrádióban az terjedt el, hogy a dolgozók feletől **búcsút** vesznek, akik viszont megússzák a kirügást, több fizetést fognak " 25 alatt! " -ígt; oké, Milán az egyik tombolától könnyes **búcsút** vett... " 10 alatt!!!! " Milán kiesett a játékból, ott maradt szabott csónakházi szobalabirintusban. Lefekvés előtt Clifford **búcsút** vett a hölgyektől. Mert korán egyedül volt kénytelen másnapi egy gerenda mögé, naplója mellé, aztán anélkül, hogy valakitől **búcsút** vett volna, indult a Sétátér felé. Szerencséje volt, egyetlen , vagy tudomásuk van hasonló jellegű bűncselekményről. Végso **búcsút** vettek Fejtő Ferentől Budapestén Végso búcsút vettek Fejtő kezdés: 22:00 kor.vártak benneteket! 2013-09-15 MTI Úszás: végso **búcsút** vettek Temes Juditól A Farkasréti temetőben hétfőn végso búcsút

## A fordított szórendű megoldások a célszó előtti **bal oldali szó szerinti** rendezéskor tűnnek elő:

tudományos  
személyes-közösségi  
sajtó  
szépirodalom  
szépirodalom  
szépirodalom  
szépirodalom  
beszélnyelvi  
személyes-közösségi  
személyes-közösségi  
sajtó  
sajtó  
szépirodalom  
szépirodalom  
szépirodalom  
személyes-közösségi  
szépirodalom  
sajtó  
sajtó  
személyes-közösségi  
szépirodalom  
sajtó  
tudományos  
sajtó

derekan helytállt, a bajnokság végén pedig emelt fővel inthetett **búcsút** a legjobb mezőnyének. A gárda menet közbeni fiatalítása miatt mendei Club Prince, ahol erőt próbáló táncos mulatsággal inthetsz **búcsút** a punnyadásnak - mindent teljesen INGYEN! [2011-10-29] Instant alkalom után körülbelül félkönyvnyi súlyfeleslegnek inthetsz **búcsút** . Vigyázz, mert az együttlétek után sokan farkasdegről számlolnak Katséré karjában. A csillagasszony meglehetősen volna, hogy ő is **búcsút** mond a földnek, de nem tudott elszakadni a kunyhótól, ahol madárdalaladdal, szerelembe futó lány ösvényeiddel. Ezer lengő karral **búcsút** ingetgető: utolszór látlak így, leomló ifjúság. Kiertem belőled viszont ha már így alakult, szerdán mindenképpen győzelemmel kell **búcsút** vennünk egymástól. Bánki József [origo] lkszelt az MTK Pécsset megmet ottan, ez a vizsgálat, már éppen az István király-napi **búcsút** vártuk rá pár napra. Na persze hogy mi is beszebelhetünk Sonával Több mint valószínű, hogy az egyik magyar érdekltségű klub **búcsút** mond a Premier League-nek, hogy ez ne a Bolton legyen, azért " 25 alatt! " -ígt; oké, Milán az egyik tombolától könnyes **búcsút** vett... " 10 alatt!!!! " Milán kiesett a játékból, ott maradt nézték, azok sajnos hamarosan kénytelenek lesznek tőle könnyes **búcsút** venni. Az ORIGO információ szerint a T-Kábel az éjszakai csatorna szögekben és magasságokban állítható monitorálványának köszönhetően **búcsút** inthetünk a nyakban tomboló fájdalomnak. Arra az esetre pedig Amolyan ezredvégi requiem lesz, melynek során újra átéljük, majd **búcsút** mondunk a huszadik századnak. - Milyennek tartja ezt az időszakot hányadszor...! Mit csináljak? Olvassam Buddhát, hogy könnyen mondjak **búcsút** e világnak? Menjek parasztnak, és higgyem el az ócska babonákat években tüdejének egy részét is eltávolították. 2003-ban mondott **búcsút** az elnöki székek nem sokkal azelőtt, hogy a cseh köztársaság szupermodellek fémjelzte Fashion Cafénak éppen Claudia Schiffer mondott **búcsút** tavaly szeptemberben. Nagy Gábor HVG 99/02. szám 1999. január 16. Bár, ráadásul most Szalóki Ági is jelen lesz, aki azóta már **búcsút** intett a kollégáknak. A hosszú pihenő előtt még négy koncertet ébredés halál volt. Veled fent voltam a napfényen, Most már **búcsút** intenek a rögök és a fák, a hajnali álommal eltűnök örökre, egyedül élt és még nem volt ötvenéves; másnap azonban mégis **búcsút** kellett vennie tőle; a kaszárnyában aztán az a kellemetlen lapnak adott - interjúban kilátásba helyezte, hogy a nyáron **búcsút** inthet az angol élvonalnak. Sporthircentrum Ha a sportvilág " küldötteit ", akik abban az évben is megjárták a pünkösdi **búcsút** . Az évente megszervezett Ezer Székely Leány Napja a széksely menet A várt hetvenezert helyett csak hétézren vettek személyesen **búcsút** Zámbo Jimmy popénekéstől, aki sikerei csúcán - a rendőrség kolostor- és börtönudvara, ahol aznap a pálos rendi szerzetesek **búcsút** tartanak. Ehhez a szakrális eseményhez csatlakoznak a performansz-előadók egy gerenda mögé, naplója mellé, aztán anélkül, hogy valakitől **búcsút** vett volna, indult a Sétátér felé. Szerencséje volt, egyetlen kilátástalansága, hogy szeptember 11-én elhagyta Wetzlart és levélben vett **búcsút** a jegyespártól. Az elválás irodalmi feldolgozása a Werther-regény erősítette a Kalapácsosokat, azaz 12 évnyi szolgálat után vett **búcsút** Londontól. Hurstöt 75 ezer fontért szerezték meg a piros-fehérek április elején láthatja a szabadban a hófehér bocot. Ezek vettek **búcsút** Sir Edmund Hillarytól Ezek vettek búcsút Új-Zélandon Sir Edmund gyászoló családtagok vasárnap magán istentiszteleten vettek **búcsút** Erzsébet királynétől. Kedden a Windsor-ház zászlajával letakart kedvező, kilátás pazar. Este hazautazás. Nehéz szívvel vettem **búcsút** kedves hegyeimtől, hisz vége már arany szabadságomnak s már és Hegedűs Zsuzsanna ajándékaival, édeségszomaggal vettünk **búcsút** erdélyi látogatóinktól. Hisszük, hogy barátaink, vendégeink

Itt is könnyen megtaláljuk az *int búcsút*, *mond búcsút* és a *vesz búcsút* állandósult szókapcsolatokat, sőt e rendezésnél feltűnik egy tipikus jelzős szerkezet is, a *könnyes búcsút (vesz)*.

A képen a találatoknak csupán a töredékét látjuk, mivel a 7811 találatból mindössze 50 elemű véletlen mintát vettünk, s a találatoknak még így is csak egy részét másoltuk ki.

Természetesen a kollokációk vizsgálatára más módszerek is vannak, amint erről alább szó lesz. Ezt az opciót elsősorban akkor érdemes alkalmazni, ha nem kifejezetten a kollokációkkal foglalkozunk, ám valami miatt mégis gyorsan át szeretnénk tekinteni a találatokat a kollokációk szempontjából.

Ha csak olvasgatni akarjuk a találatokat, akkor a legjobb a **sor elején látható infó** szerint rendezni, mert akkor **az azonos stílusrétegbe tartozó találatok egymás után következnek**.



Az alábbi képen a *múlvára* szóalakra vonatkozó 45 találatot látjuk, infó szerint rendezve, karakterszámmal meghatározott szövegtérletben (KWIC-megjelenítés).

találatok száma: 45	
beszéltnyelvi	legfontosabb gazdasági erőteréről van szó, és hogyha ma már 13 <b>múlvára</b> , pontosabban 13 év múlva itt ülünk ebben a stúdióban, akkor
beszéltnyelvi	egyik településen, ahol nagyon komoly bűnözés folyik ott 3 hónap <b>múlvára</b> ígérték meg, hogy lesz valami rendőrfjór, vagy éppen föllállítanak
beszéltnyelvi	visszaigényelni az összeget. Baranyai Mária: - Tehát ha én mondjuk fél év <b>múlvára</b> vettem volna mondjuk Amszterdamba egy Malév jegyet, és azt
beszéltnyelvi	illetőnek, hogy az én kis kertemet ne tudta elvállalni csak egy év <b>múlvára</b> . Én ebből vontam le azt a következtetést, hogy itt nagyon sok
beszéltnyelvi	forgatókönyvek, hogy a május közepén ülésező választmány 45-50 nap <b>múlvára</b> összehívhatja a teljes tisztújító kongresszust. Ebből az is
beszéltnyelvi	volna a fogszabályozásra, de a legközelebbi időpontot két év <b>múlvára</b> tudták megadni, hogy körülbelül addigra lesz arra alkalom,
beszéltnyelvi	az igényes zenének van kelletje és azt gondolom, hogy egy év <b>múlvára</b> olyan helyzetben lesz a produkció, hogy nemzetközi folyamatok
beszéltnyelvi	forgatókönyvek, hogy a május közepén ülésező választmány 40-50 nap <b>múlvára</b> összehívhatja a teljes tisztújító kongresszus. Ebből az is
beszéltnyelvi	forgatókönyvek, hogy a május közepén ülésező választmány 45-50 nap <b>múlvára</b> összehívhatja a teljes tisztújító kongresszust. Ebből az is
beszéltnyelvi	fel a világnyelvekben, igaz, hogy Örkény is csak százhusz év <b>múlvára</b> jósolta mindezt. Tehát a magyar nem világnyelv, bár a világon
beszéltnyelvi	létrejövő bankfelügyelet. Tehát... Mv: Dehát ők nyilván nem 15-20 év <b>múlvára</b> tervezik ezt. Torba Tamás, közigazdász: De, körülbelül ilyen
sajtó	birtokában már el is röpben az újabb másfél óra, amikor végre öt hét <b>múlvára</b> , délután negyed ötre van egy alkalom. Ha nem jó, tessék szíves
sajtó	energiastratégia eredetileg támogatta is a kukoricabenzint. Nyolc év <b>múlvára</b> 10 százalékra kellett volna növelni a biüzemanyagok részesedését
sajtó	üzemeltető Fusion Rt. marketingigazgatóját is, aki néhány nap <b>múlvára</b> ígért választ, mert előbb egyeztetnie kell a vállalat külföldi
sajtó	ügyben ", ugyanakkor további tájékoztatást csupán " néhány nap <b>múlvára</b> " ígértek a kárelhárítási munkák előrehaladásáról és a szükséges
sajtó	adójóváírás összege 277,3 milliárd forint volt, tehát két év <b>múlvára</b> ez a kedvezmény lényegében megszűnik - írja a Napi Gazdaság
sajtó	szigorítást kényszeríthet ki a piaci várakozások szerint. Fél év <b>múlvára</b> 7,5 százalékra várja a piac a kamatot, amely most 6 százalékon
személyes-fórum	... fekvőverseny igen jó és kreatív Slet, elraktározom, pár év <b>múlvára</b> , természetesen tiéd a kopirajtí üdv bambi heki 01/21/99 10:41:00
személyes-fórum	sikerült végigjönni, nem igaz? A következő túrát kb. két hét <b>múlvára</b> tervezem, lényegesen hosszabb lesz. Aki úgy gondolja, hogy
személyes-fórum	térképen! Most a Downloading Components részről járok, 4 óra <b>múlvára</b> ígéri a 17 mega lehozatalát... Letöltött Mogorva 03/29/99 01:42:03
személyes-fórum	Lakatos még nem látta, mert mindenki elküld a fenébe, és 3 hónap <b>múlvára</b> vállalnak. Én olyan 150-et szánnék max a dologra, mert ennyire
személyes-fórum	, hogy mindkét fél igényeit kielégíti. Mondjuk ez a 20-30 év <b>múlvára</b> vonatkozik, déva pedig még nem tart ott. flyOrDie 09/13/99
személyes-közösségi	hogy ott volt, vert a picit szíve és jó helyen volt. Két hét <b>múlvára</b> rendelt vissza. Aznap délután egyetlen embernek mondtuk
személyes-közösségi	vége pár óra alatt mint egy randinak, hanem mindenképp 1 hónap <b>múlvára</b> van haza repjegyem. Mondjuk aki nem olyan csóró mint én, az
személyes-közösségi	vége pár óra alatt mint egy randinak, hanem mindenképp 1 hónap <b>múlvára</b> van haza repjegyem. Mondjuk aki nem olyan csóró mint én, az
személyes-közösségi	napiját amúgyis Bangkokban töltjük, megnézzük az időjárás két hét <b>múlvára</b> , döntünk és meg is vesszük az irodában a jegyet. Ennél korábbam
személyes-közösségi	pattozani. ha szerencsénk van, az öskommunisták is kihallnak négy év <b>múlvára</b> . talán 2018 után. addig meg, viseljük el ezeket a lábszagú
személyes-közösségi	pattozani. ha szerencsénk van, az öskommunisták is kihallnak négy év <b>múlvára</b> . talán 2018 után. addig meg, viseljük el ezeket a lábszagú
személyes-közösségi	pattozani. ha szerencsénk van, az öskommunisták is kihallnak négy év <b>múlvára</b> . talán 2018 után. addig meg, viseljük el ezeket a lábszagú
személyes-közösségi	pattozani. ha szerencsénk van, az öskommunisták is kihallnak négy év <b>múlvára</b> . talán 2018 után. addig meg, viseljük el ezeket a lábszagú
személyes-közösségi	pattozani. ha szerencsénk van, az öskommunisták is kihallnak négy év <b>múlvára</b> . talán 2018 után. addig meg, viseljük el ezeket a lábszagú
személyes-közösségi	többen bízni benne, hogy irreálisan nagy flottát ígérnek pár év <b>múlvára</b> , (közben meg a létező legolcsóbban beszerezhető gépekkel repkednek
személyes-közösségi	tudom mi van de ha vmi fertőző szart szedtem össze 2-3hónap <b>múlvára</b> kihullhat a fél fejem ráadásul sanszosan megfertőztem rakás
személyes-közösségi	tudom mi van de ha vmi fertőző szart szedtem össze 2-3hónap <b>múlvára</b> kihullhat a fél fejem ráadásul sanszosan megfertőztem rakás
személyes-közösségi	gyereksébeszetre de épp elmaradt a rendelés mikor mentünk és két hét <b>múlvára</b> kaptuk a következő időpontot. Azt mondták, hogy addig kezeljem
személyes-közösségi	hallásvizsgálatot? Hát jól sejted... Szóval inkább úgy 10 év <b>múlvára</b> vártam hasonló eseményeket. Inkább ezért megy a brüühü. (Bár
személyes-közösségi	levonni a következtetést. Tudod, erőforrás. Nekem nincs. 20 év <b>múlvára</b> felnevelem a családom, rá 5 évre meglesz a gyakorlatom...: Gondolom
személyes-közösségi	megtervezéséhez nem kell erőforrás, csak szellemi munka." 20 év <b>múlvára</b> felnevelem a családom, rá 5 évre meglesz a gyakorlatom. " Lol
személyes-közösségi	szerintem teljesen normális... Ha én kérek ambulanter, én is 3-4 hét <b>múlvára</b> kapok. 2013-10-30 8emi95@kriszti203 mert utána két hétig sehová
szépirodalom	influenzán, engem félt, hogy nehezen fogok kilábolni három nap <b>múlvára</b> , amikor indulnom kell Firenzébe. Nekem meg eszemben jár apám
szépirodalom	meg, hogy a legjobb barátja volt. Akkor tűzték ki, hatvan nap <b>múlvára</b> , az ő operációját. Az elmúlt két hétben kívül-belül rágott
szépirodalom	". Félmillió Rákóczi Ferenc fekszik most kiterítve. (Egy hét <b>múlvára</b> eltűnnek a kurdok az újságok első oldaláról. Sehoh se lesznek
szépirodalom	a következő kutyákat terveztük, úgy körülbelül tizenégy év <b>múlvára</b> : egy tibeti terrier, ugyanis másikkunknak egy puli-pumi keverék
tudományos	szájából ez a mondat hangzik föl: " A kormány csak másfél hónap <b>múlvára</b> ígéri, hogy... " Sajátságos egy időhatározó ez a múlvára. Hasonlót
tudományos	hónap múlvára ígéri. hové... " Sajatságos evv időhatározó ez a <b>múlvára</b> . Hasonlót hallok evv tévéinterjú-аланtvól is. és azonmód fölvevzem

Így rendezve a találatokat azonnal szembeötlik, hogy a *múlvára* szóalakra hat alkorpuszból vannak adatok, betűrend szerint a „beszélt nyelviből”, a „sajtónyelviből”, a „személyes-fórumból”, a „személyes-közösségből”, a „szépirodalmiból” és a „tudományosból”. Egyedül a „hivatalos” alkorpuszból nincs találatunk.

Azt, hogy a *múlvára* használata mely alkorpusz(ok)ra jellemző, ránézésre nem tudjuk megállapítani, ahhoz relatív gyakoriságot kellene számolni; ezzel a *Stílusréteg szerinti mennyiségi vizsgálat* című fejezetben foglalkozunk.

Mivel azonban ritka előfordulású nyelvi formáról van szó, a relatív gyakoriság sem ad igazán megbízható fogódzót, ugyanis az alacsony elemszám miatt nagy a véletlen szerepe.

Ha nagyon sok a találatunk, és nem fogjuk tudni az összeset áttekinteni, a legegyszerűbb, ha a „keverés” opciót választjuk. Ekkor **véletlenszerű sorrendben** jelennek meg a találatok, így nem kell attól tartanunk, hogy például csak bizonyos stílusréteg szavaira vagy bizonyos toldalékkal ellátott szavakra vonatkozó találatokat fogunk látni.

Az alábbi képen a *romlás* szótőre vonatkozó találatok első oldalának elejét látjuk: mindegyik a „szépirodalmi” alkorpuszból való:



## ÷ 5. Véletlen minta

Ha a keresés **nagyon nagy mennyiségű** (mondjuk több ezer, több tízezer, több százezer vagy akár több millió) találatot ad, érdemes véletlen mintát készíteni, s azon végezni el a vizsgálatokat. Az, hogy mekkora legyen a véletlen minta, az elvégzendő feladattól függ. Általában **néhány száz** találat még viszonylag kényelmesen kezelhető.

The screenshot shows the MNSZ2 web application interface. At the top, the user is logged in as 'gramma' and the corpus is 'MNSZ2'. The version is 'v2.0.5 - friss verzió 2018. június 6-tól'. The main content area is titled 'Véletlen minta' and contains a text box with the text 'Véletlen minta az aktuális konkordanciából.' and a form field for 'A minta elemszáma: 350'. Below the form is a button labeled 'Véletlen minta'. On the right side of the interface, there is a footer with the text 'Lexical Computing Ltd. Sketch Engine (ver:open-2.59.1-open-2.91.13) A Nemzeti Korpuszportál tagja. MNSZ2 - v2.0.5 - 2018-06-08 Nyelvváltás: English | magyar'.

Ha véletlen mintát kértünk, akkor a példamondatok gyűjtéséhez elegendő ezt használni; ha nem találunk a véletlen mintában megfelelőket, akkor folyamodhatunk a teljes mintához vagy pedig nagyobb véletlen mintát állíthatunk be.

Amikor viszont szám adatokkal dolgozunk, akkor természetesen nem használhatjuk a véletlen minta adatait, hanem a teljes mintával kell dolgoznunk. Mivel a tisztítást nagy mintán nagyon időrabló volna elvégezni, ha a találatok száma 350-400-nál magasabb, a tisztítást – ha szükség van rá – a véletlen mintán végezzük, s kiszámoljuk, az adatok hány százaléka releváns. Utána ugyanezt a számítást elvégezzük minden további adatnál, amely a teljes minta valamely részletére vonatkozik.

A tisztításnak csak akkor van értelme, ha abszolút számokkal dolgozunk. Amikor arra vagyunk kíváncsiak, az adatbázisnak (vagy valamely alkorpusznak) mekkora részét teszik ki a találatok, főleg az adatokat tisztítani, mivel az adatbázisnak, ill. **az egyes alkorpuszoknak a nagyságára vonatkozóan nem rendelkezünk tisztított adatokkal.**

## ÷ 6. Szűrés

A szűrés funkciót mélyebb elemzésekhez szükséges használni. Házi feladathoz, házi dolgozathoz vagy szemináriumi dolgozathoz ennek használata nem követelmény. Ha mégis szeretnénk használni ezt a funkciót, a súgóból megismerhetjük működését. Itt csak a bevezető részt idézzük a súgóból:

A korpuszlekérdező nagyon erős eszköze a szűrés. Egy lekérdezés eredményét szűrhetjük egy újabb lekérdezéssel, ilyenkor az első lekérdezésből kapott konkordancia anyagán futtatjuk a szűrőlekérdezést. Az így kapott eredményt ismét szűrhetjük, és így tovább. Bonyolult lekérdezéseket szűrőlekérdezések soraként fogalmazhatunk meg, illetve egy nagy méretű konkordanciát feltérképezhetünk szűrések által. Az űrlap alsó részében egy keresést adhatunk meg. Ez a

felület azonos a keresésnél használt összetett felülettel, annak minden lehetőségét használhatjuk itt is.

Keresés

---

< Konkordancia

Véletlen minta

Szűrés

Gyakoriságok

szóalakok

szótövek

elemzések

dokumentumok

Kollokációk

Haladóknak:

Alkorpuszok

### Konkordancia szűrése

**Szűrés:**  pozitív  negatív

többszavas találat esetén a találat:  első  utolsó szava alapján

keresési ablak: eleje:  to   a találati szót beleértve

lekérdezés típusa: részletes keresés

részletes keresés: szóalak   szófaj: TETSZ   összetett szó | morf:  morféma: TETSZ

## ÷ 7. Gyakoriságok

Gyakorisági lista készíthető a találati szóalakok, szótövek, morfológiai elemzések és a dokumentumok szerint. Az ilyen listák nagyon hasznosak, ám ezeknek a használata sem kötelező, így elkészítésük módjával itt nem foglalkozunk.

Keresés

---

< Konkordancia

Véletlen minta

Szűrés

Gyakoriságok

szóalakok

szótövek

elemzések

dokumentumok

Kollokációk

### Többszintű gyakorisági eloszlás

gyakorisági küszöb:

1. szint

attribútum: word

kisbetű/nagybetű nem számít

6L  
5L  
4L  
3L  
2L  
1L  
Node  
1R

pozíció: 2R

2. szint

attribútum: word

kisbetű/nagybetű nem számít

6L  
5L  
4L  
3L  
2L  
1L  
Node  
1R

pozíció: 2R

3. szint

attribútum: word

kisbetű/nagybetű nem számít

6L  
5L  
4L  
3L  
2L  
1L  
Node  
1R

pozíció: 2R

4. szint

attribútum: word

kisbetű/nagybetű nem számít

6L  
5L  
4L  
3L  
2L  
1L  
Node  
1R

pozíció: 2R

### Alkorpuszok szerinti gyakorisági eloszlás

*Figyelem! Ez a funkció még nem teljes.*

gyakorisági küszöb:

jelenjen meg az a kategória is, ahol nincs találat:

doc.file  
doc.region  
doc.style  
div.column  
div.type

## ÷ 8. Kollokációk

Ez az opció lehetővé teszi, hogy fölfedezzük a célformára jellemző kollokációkat. Ez főleg a szerkezetek vizsgálatában lehet hasznos. A „**Kollokációk**” menüpontra rákattintva megjelenő űrlapon a „**Kollokáció-jelöltek**” opció mutatja meg a lehetséges kollokációkat, a program által kiszámított gyakoriság szerint.

Ha például arra vagyunk kíváncsiak, hogy az *éti* szó vajon csak a *csiga* jelzője lehet-e, vagy más szavaké is, rákeresünk az *éti* szóra. Mivel jelzői funkcióban érdekel minket az *éti* használata, jobb, ha szóalakra keresünk rá:

Ha nem változtatunk az alapbeállításokon, ilyen találatokat kapunk:

**Kollokáció-jelöltek megtekintése**

	<u>Freq</u>	<u>T-score</u>
p/n csiga	79	8.888
p/n osztriga	8	2.828
p/n spárga	4	2.000
p/n települési	3	1.731
p/n kagyló	2	1.414
p/n gomba	2	1.413
p/n franc	1	0.999
p/n ~	1	0.993
p/n tanács	1	0.990
p/n -	1	0.114
p/n ,	1	-6.700

A kimutatás alapján könnyen megállapíthatjuk, hogy a *települési*, a *franc* meg a *tanács* irreleváns találat, de a biztonság kedvéért rákattinthatunk a sor elején álló *p* betűre: ez lehetővé teszi, hogy a maga eredeti kontextusában nézzük meg a találatot vagy a találatokat (az *n* betűre kattintva az összes találat megjelenik, nem csak az adott kollokációra vonatkozók).

Azokat az eseteket, amikor írásjel alkotja a „kollokáció” második tagját, mindenképpen meg kell néznünk kontextusban, a *p*-re való kattintással.

Csak így derül ki, hogy a hullámvonal az *éti* szót helyettesíti, és a *csiga* szóhoz kell hozzáérteni, tehát ez is releváns találat, a *csiga* találatai közé számítható (még ha metanyelvi jellegű is):

Keresés ?	korpusz: MNSZ2 találatok száma: 1 tudományos	csiga: escargot ( <i>éti</i> ~ )/ limace (meztelen~)
--------------	--	--

Az *éti* + gondolatjelként funkcionáló kiskötőjel mint „kollokáció” mögött is releváns találat húzódik meg:

Keresés ?	korpusz: MNSZ2 találatok száma: 1 sajtó	Az emberi fogyasztásra alkalmas <i>éti</i> - vagy európai - osztriga teljes kifejlődését - amikor már alkalmas étkezési célokra is - évekkal a világrajövele után éri el.
--------------	---	---

Amint a képről leolvasható, ez a találat az *éti osztriga* kollokációt képviseli.

Ami viszont az *éti* szó + vesszőt mint „kollokációt” illeti, azt látjuk, hogy ez irreleváns, nincs köze az *éti* szóhoz, hanem elírás *érti* helyett:

Keresés ?	korpusz: MNSZ2 találatok száma: 1 sajtó	Az indiai alakulat főnöke, Vijay Mallya nem <i>éti</i> , hogy miért három évvel a szerződés megkötése után merült fel ez a kérdés.
--------------	---	--

Persze könnyebb volna a dolgunk, ha a program volna annyira intelligens, hogy egy kivételjegyzék alapján nem hozna létre szó + írásjel „kollokációkat”, hanem az írásjelet átugorva a következő szót tekintené relevánsnak.

Mivel a releváns találatok száma nem annyira magas, hogy ne volnának könnyen áttekinthetők, érdemes őket azért is megnézni, hogy nincs-e köztük ismétlődő. Az *éti csiga* esetében egyetlen ismétlődő találatot fedeztünk fel (ami azt jelenti, hogy a korpusz két példányban tartalmazza az érintett dokut), a többi kollokáció esetében nem találtunk ismétlődést.

Ezek után megállapíthatjuk, hogy az *éti* melléknév legjellemzőbb kollokációja az *éti csiga* 79 találattal (79 + 1 - 1), sokkal ritkább az *éti osztriga* 9 találattal (8 + 1), ezt követi az *éti spárga* 4, az *éti gomba* 2, illetve az *éti kagyló* ugyancsak 2 találattal.

Meg kell még jegyezni, hogy a hétköztetes értelmező szótár és az értelmező kéziszótár csak az *éti csiga* kollokációt ismeri.

Az értelmező szótár **éti** szócikke:

**éti** mn (*Áll*)

⟨Csak ebben az állandósult szókapcsolatban:⟩ ~ *csiga*: a tüdős csigák rendjébe tartozó, sárgás v. barnás héjú, sárgás testű csiga (*Helix pomatia*); nálunk a legnagyobb csiga; a nyugati népek ősidőktől fogva eszik.

Az értelmező kézisótár **éti** szócikke:

**éti** ⑤ mn *Áll*: ~ *csiga*: emberi táplálékul is haszn., tüdős, nagy csiga (*Helix pomatia*).

Ez is egy példa arra, hogy az MNSZ2 segítségével olyan lexikális információkhoz juthatunk, melyeket a szótárakbn nem találunk meg.

## & A KONTEXTUS MEGADÁSA

A keresőfelületről elérhető „**Kontextus**” menüpontra kattintva meg tudjuk szabni, hogy a célszó előtt (azaz tőle „**balra**”) vagy után (azaz tőle „**jobbra**”) vagy a célszó előtt is, után is (azaz „**mindkét irányban**”) milyen szavak forduljanak vagy ne forduljanak elő.

Ilyenkor ezeket a szavakat a „**szótövek**” mezőbe írjuk be, utána pedig (a feladat jellegétől függően) a „**mind**” vagy „**bármelyike**” opciót választjuk. Fontos, hogy ebbe a mezőbe a szótövet kell beírni: ha olyan szóalakot írunk be, amely eltér a szótőtől, a program nem ad találatot (pl. ha a *látunk* alakot szeretnénk kontextusnak, erre sajnos nincs lehetőség, mert a program csak a *lát* szótövet fogadja el kontextusként).

Ha azt akarjuk, hogy egy vagy több szó NE forduljon elő a célszó kontextusában, a „**közül egy se**” opciót választjuk.

Az ilyen keresésnél azt is meg kell adnunk, hogy a célszótól milyen távolságban forduljanak vagy ne forduljanak elő a megadott szavak: **1, 2, 3, 4, 5, 6, 7, 10** vagy **15** szó távolságban.

Az alábbi képen az alapbeállítás látszik, ami mindkét irányban 5 szó:

MNSZ2  
Felhasználó: gramma korpusz: MNSZ2 Sűgő

Keresés  
Lehetőségek:  
Kontextus  
Alkorpuszok

korpusz: MNSZ2  
lekérdezés típusa: részletes keresés  
részletes keresés: szóalak  szófaj: TETSZ  
 összetett szó | morf: morféma: TETSZ

**Kontextus**

Szűrés szótövekre  
szűrés ablak: mindkét irányban 5 szó.  
szótövek: forduljon elő.

Konkordancia készítése Törlés

Az együtt-előfordulásra való keresés többek között az **állandósult szókapcsolatok** vizsgálatában lehet hasznos. Ha például az *éti csiga* szókapcsolatra vagyunk kíváncsiak, beírhatjuk a keresőmezőbe az *éti* szót (szóalakként), majd a kontextusnál a *csiga* szót (szótóként) úgy, hogy az közvetlenül az *éti* után forduljon elő (azaz 1 szó távolságban, tőle jobbra). (Mivel csak egy szót adunk meg kontextusként, mindegy, hogy a „mind” vagy a „bármelyike” opciót választjuk.)

Keresés  
Lehetőségek:  
Kontextus  
Alkorpuszok

korpusz: MNSZ2  
lekérdezés típusa: részletes keresés  
részletes keresés: szóalak  szófaj: TETSZ  
 összetett szó | morf: morféma: TETSZ

**Kontextus**

Szűrés szótövekre  
szűrés ablak: jobbra 1 szó.  
szótövek: csiga mind forduljon elő.

Konkordancia készítése Törlés



Az *éti csiga* szókapcsolatra 79 találatot kapunk, például ilyeneket:

<p>Keresés</p> <p>Mentés</p> <p>Megjelenítés KWIC/mondat</p> <p>Rendezés bal   jobb találat infó keverés</p> <p>Véletlen minta Szűrés</p> <p>Gyakoriságok szóalakok szótövek elemzések dokumentumok</p> <p>Kollokációk</p> <p>Próbálja ki a Mazsola eszközt igék és ...</p>	<p>korpusz: MNSZ2 találatok száma: 79</p> <p><b>szépirodalom</b> - Leonének hívják - felelte -, apró csizmája van, <b>éti csigát</b> tud sütni, ért a klasszikus zenéhez.</p> <p><b>szépirodalom</b> De nem baj, hoztattak <b>éti csigát</b>, tápnak a nyulaknak, az a legdrágább, ezt kinézték egy könyvből.</p> <p><b>szépirodalom</b> De ezek rossz erkölcsű nyulak voltak, nem ették meg az <b>éti csigát</b>, és felfordult mindvalahány.</p> <p><b>szépirodalom</b> Ami nem baj, mert az elnök bácsi, az bennmaradt, és most <b>éti csigát</b> akar tenyészteni, mert abból maradt, csak előbb le kell bontani a templomot, hogy tudjanak venni nyulakat, éticsiga-tápnak.</p> <p><b>hivatalos</b> Az <b>éti csiga</b> szervezettana.</p> <p><b>személyes-fórum</b> Fókusz! (vérekes szem, habzó száj) Az <b>éti csiga</b> és az élő hal párbaját a gyomrodban ne várd meg!</p> <p><b>személyes-fórum</b> Lebunkóztam V.-t, ráléptem egy <b>éti csigára</b> és halva született topicot hoztam létre.</p> <p><b>személyes-fórum</b> Ne az <b>éti csiga</b> házára gondoljatok, hanem olyan kis lapos 1-2 cm-es csigaházra.</p> <p><b>személyes-fórum</b> " Álljon meg a menet! " - gondolhatta magában - " ez a hapsi meztelen mint az <b>éti csiga</b> !</p> <p><b>személyes-fórum</b> Bár nagyon szeretem a házas csigákat, inkább pofára esek, mint elgázoljak egy <b>éti csigát</b>, mi több egyszer majdnem a kocsi is összetörtém miatta, de sajnos itt nem tudtam mit tenni, annyit volt...</p> <p><b>sajtó</b> Tízezeret ér egy <b>éti csiga</b></p> <p><b>sajtó</b> Védettséget kapott a veréb, az <b>éti csiga</b>, valamint a hangyabolycsapat.</p> <p><b>sajtó</b> Védett lett több tömegesen előforduló állat, így az <b>éti csiga</b>, a veréb, valamint a hangyaboly.</p> <p><b>sajtó</b> Az exportálók országos szinten évente mindössze háromezer tonna <b>éti csigát</b> szállíthatnak külföldre.</p> <p><b>sajtó</b> Eszik az <b>éti csigát</b></p> <p><b>sajtó</b> Rémulten olvasom az újságban, hogy csökken az <b>éti csiga</b>, a veréb és a hangyák száma.</p> <p><b>sajtó</b> Az idény az <b>éti csigával</b> kezdődik - bár ez nem gyógynövény -, majd a gyermekláncfű és a bojtorján gyökerével folytatódik.</p> <p><b>sajtó</b> Nem mintha Miskolcon nem ettek volna például nyáron sültöt kutyát, macskadögöt Skriba módra, borban áztatott sáskát, sózott tücsköt, hangyás szendvicset, a házából harapófóvával kihúzott <b>éti csigát</b>, félig tollas lopott tyúkot gyerekkoromban az Avason, ám a tökök, herék, gyomrok, belek, tarajok és körmök arsenálja, ami az Alföldön fogadott, felülmúlt minden képzeletet.</p> <p><b>sajtó</b> Nemhogy engedélyhez nem kötik, de még kvóta sem szabályozza a gyűjtését, pedig mindez az <b>éti csigáknál</b> már bevált.</p>
---	--

Az így kapott eredmény teljesen megegyezik azzal, amit fentebb a kollokációs vizsgálattal kaptunk.

Ha ezzel a módszerrel szeretnénk kideríteni, hogy az *éti* jelző a magyarban csak a csigával kapcsolatban használatos-e, vagy másra is vonatkozhat (azaz más szónak is lehet-e a jelzője az *éti* melléknév), megszabhatjuk, hogy a program csak azokat a találatokat adja ki, amelyekben az *éti* után *nem* áll a *csiga*. Ehhez a *csiga* beírás után a „közül egy se” opciót választjuk, s megszabjuk, hogy a *csiga* szó ne forduljon elő az *éti*től jobbra 1 szó távolságban (azaz közvetlenül mellette):

<p>Keresés</p> <p>Lehetőségek: Kontextus Alkorpuszok</p>	<p>korpusz: MNSZ2</p> <p>lekérdezés típusa: részletes keresés</p> <p>részletes keresés: szóalak <input type="checkbox"/> szófaj: TETTSZ <input type="checkbox"/> összetett szó   morf: morféma: TETTSZ</p> <p><b>Kontextus</b></p> <p>Szűrés szótövekre</p> <p>szűrési ablak: jobbra 1 szó.</p> <p>szótövek: csiga <input type="checkbox"/> közül egy se <input type="checkbox"/> forduljon elő.</p> <p>Konkordancia készítése Törlés</p>
--	---

Ilyen beállításnál 24 találatot kapunk.

Szigoríthatjuk a keresést úgy, hogy azt sem engedjük meg, hogy a *csiga* 5 szó távolságban forduljon elő, az *éti* szó előtt sem:

Keresés

Lehetőségek:  
Kontextus  
Alkorpuszok

korpusz: MNSZ2  
lekérdezés típusa: részletes keresés  
részletes keresés: szóalak  éti  szófaj: TETSZ  összetett szó | morf: morféma: TETSZ

Kontextus

Szűrés szótövekre  
szűrési ablak: mindkét irányban 5 szó.  
szótövek: csiga közül egy se forduljon elő.

Konkordancia készítése Törlés

A találatok száma így 23 lesz. Ez elég alacsony szám ahhoz, hogy könnyen megállapítsuk: bár az *éti csiga* szókapcsolat a legjellemzőbb, előfordul *éti spárga*, *éti kagyló*, *éti osztriga* és *éti gomba* is; ezenkívül jó pár találat elírásnak köszönhető (ezeket természetesen a statisztika-készítés előtt, a másodlagos tisztítás során eltüntetjük).

Az alábbi képernyőkép csaknem az egész találati felületet tartalmazza:

Keresés

találatok száma: 23

Mentés  
Megjelenítés KWIC/mondat  
Rendezés bal | jobb találat infó keverés  
Véletlen minta Szűrés  
Gyakoriságok szóalakok elemzések dokumentumok  
Kollokációk

Viszajelzés  
Próbálja ki a Mazzola eszközt ígék és argumentumaik közvetlen vizsgálatához!

szépirodalom	Tudod mit, ezt másokra hagyom, fedezzék fel rótam, fedezzék fel bennem a megszakadó szívú, a madárszólókat-dobó-szívú E-l-a-v-u-l-t E-m-b-e-r-i-s-é-g-e-t, mely a Modern Múzeumból kijövet átvág egy zöldégyplacon, spárgát venne - nem, hogy felkösse magát! spárgát, <b>éti</b> spárgát, hogy elkészítse - hogy F. elkészítse! - egy villegjuifi konyhában, végül kiköt egy pár banánnal és egy önkiszolgálóban az Ötödik György sugárpalota mélyszomszédságában.
hivatalos	Ezen díjszabás alkalmazása kötelező az Önkormányzat tulajdonát k é pezó Nádas-r <b>éti</b> települési szilárd hulladéklerakó telepre beszállított hulladékok ártalmatlanítása esetén:
hivatalos	Ezen díjszabás alkalmazása kötelező az Önkormányzat tulajdonát k é pezó Nádas-r <b>éti</b> települési szilárd hulladéklerakó telepre beszállított hulladékok ártalmatlanítása esetén:
hivatalos	Ezen díjszabás alkalmazása kötelező az Önkormányzat tulajdonát k é pezó Nádas-r <b>éti</b> települési szilárd hulladéklerakó telepre beszállított hulladékok ártalmatlanítása esetén:
hivatalos	l/ a városkert! l <b>éti</b> tanácsok (megalakításuk esetén) és azok végrehajtó szervei jogosítványai terjedelmének és korlátainak megállapítása a városkerületi területiális közösségek érdekében;
szépirodalom	Vásároltam sajtos rudat, vettem <b>éti</b> spárgát, krumplit (ez utóbbit Totyínak), Totyínak fürödni kell, ez is eszembe jutott, adtam neki fürdővizet, mire kellene vágnom?
szépirodalom	Már a Vadász utca környékén el kellett mennem... kis szekrényszerű hely volt, papírt a közértben vettem, mai vásárlásom egyébként három újság, egy detektívregény, fél kiló <b>éti</b> spárga, egy köteg kapor, két saláta, két üveg fürödi rizling, két tekercs vécepapír, nekik egy doboz mirelit málna és magamnak egy kis tok kávétableta volt... és az előbb voltam... de lehet, hogy csak a rendőrségi tárgyalást olvastam, a gyermekgyilkosságait... bocsánat, bíróságit, rendőrségi hírt, olyan mindegy... a legutóbb az irodalommal is így vannak, ez bizonyos... olvasnak valamit... honnét? hogyan? mit?... elveszesül, olvasnak valamit.
szépirodalom	Szép szál <b>éti</b> spárga, mondta róla, saját beismerése szerint (mindenfajta szexuális melléközöngé nélkül), egy osztálytársa.
hivatalos	Rákok, kagylók, <b>éti</b> kagylók és egyéb vízi állatok, beleértve a belőlük készült termékeket: 227.4. verbar; 2 [9]. verbar; EN 1786 EN 1788 L 12.01-1. verbar;
sajtó	Az indiai alakulat főnöke, Vijay Mallya nem <b>éti</b> , hogy miért három évvel a szerződés megkötése után merült fel ez a kérdés.
sajtó	A módszer hasonló a Paulnál beválthoz: a polip <b>éti</b> kagylót választott két dobozból, amelyeken az adott csapatok nemzeti zászlói díszlettek.
sajtó	Az emberi fogyasztásra alkalmas <b>éti</b> - vagy európai - osztriga teljes kifejlődését - amikor már alkalmas étkezési célokra is - évekkal a világrajövele után éri el.
tudományos	Néhány osztrigafaj emberi fogyasztásra is alkalmas, például az <b>éti</b> osztriga ( <i>Ostrea edulis</i> ), amelyet alacsony zsír- és szénhidráttartalma miatt szívesen fogyasztanak; más fajok gyöngyöt növesztenek (gyöngyosztriga).
tudományos	Az <b>éti</b> osztrigák különösen táplálók és egészségesek.
tudományos	A tenyésztett osztrigák 99 százaléka csendes-óceáni osztriga ( <i>Crassostrea gigas</i> ), 1 százaléka <b>éti</b> osztriga ( <i>Ostrea edulis</i> ).
tudományos	Az <b>éti</b> osztriga vagy európai lapos osztriga ( <i>Ostrea edulis</i> ) a kagylók ( <i>Bivalvia</i> ) osztályának a fonálkopoltyúsok ( <i>Ostreoida</i> ) rendjébe, ezen belül az osztrigafélék ( <i>Ostreidae</i> ) családjába tartozó faj.
tudományos	Az <b>éti</b> osztriga minden európai tengerparton előfordul, ahol a víz hőmérséklete a nyári hónapokban nem süllyed 15 Celsius fok alá, és a sótartalom legalább 19 ezrelék; vagyis csupán a hideg északi vidékeken, a Balti-tengerben és a Fekete-tengerben nem található.
tudományos	Az 1940- és 1950-es években, az <b>éti</b> osztriga letelepedett a Maine és a Rhode Island államok partmenti vizeibe is.
tudományos	Az <b>éti</b> osztriga 10 - 15 centiméter hosszú, meglehetősen csúnya és durva szerkezetű teknőit tetőcsérépszerűen egymásra boruló koncentrikus pikkelyek vagy lemezek fedik, amelyek felületét sugaras irányú, szabálytalan bordák tagolják.
tudományos	Az <b>éti</b> osztriga szilárd aljazaton él a parti hullámverésnek kitett sziklákon, gazdag állományai néhol egész padokat alkotnak.
tudományos	Közeli rokonaival téveszthető össze, mint például az ízletes rizike vagy a lucfenyvesi rizike, ezek azonban mind kitűnő <b>éti</b> gombák.
tudományos	Nem téveszthető össze semmivel, kitűnő <b>éti</b> gomba.

" halál-öröm, élet-temető " érintkezési hibák elhárítója ha lel áramlót, tömíti A doktor úrnak hoztuk, de leejtettük hálalöre, mi lett -

Ilyen alacsony elemszám mellett könnyen elvégezzük az elsődleges, másodlagos és harmadlagos tisztítást.

A releváns találatok megvizsgálásához legjobb, ha azokat a találati szótól, az *étitől* jobbra álló szó szerint rendezzük, s a KWIC-beállítást alkalmazzuk, mert így az azonos kollokációk egymás alá kerülnek.

Az említett beállításokról fentebb volt szó, a *Megjelenési beállítások* című fejezetben.

MNSZ2	
Felhasználó: <b>gramma</b>	korpusz: <b>MNSZ2</b>
>> v2.0.5 - friss verzió 2018. június 6-tól	
<p><b>Keresés</b></p> <p><b>Mentés</b></p> <p><b>Megjelenítés</b></p> <p><b>KWIC/mondát</b></p> <p><b>Rendezés</b></p> <p>bal   jobb</p> <p><b>találat</b></p> <p>infó</p> <p>keverés</p> <p><b>Véletlen minta</b></p> <p><b>Szűrés</b></p> <p><b>Gyakoriságok</b></p> <p>szóalakok</p> <p>szótövek</p> <p>elemzések</p> <p>dokumentumok</p> <p><b>Kollokációk</b></p> <p>Visszajelzés</p> <p>Próbálja ki a Mazzola eszközt ígék és argumentumaik közvetlen vizslatához!</p>	<p>korpusz: MNSZ2 találatok száma: 23</p> <p><b>sajtó</b></p> <p><b>sajtó</b></p> <p><b>személyes-közösségi</b></p> <p><b>tudományos</b></p> <p><b>tudományos</b></p> <p><b>hivatalos</b></p> <p><b>sajtó</b></p> <p><b>tudományos</b></p> <p><b>tudományos</b></p> <p><b>tudományos</b></p> <p><b>infó</b></p> <p><b>keverés</b></p> <p><b>tudományos</b></p> <p><b>tudományos</b></p> <p><b>Szűrés</b></p> <p><b>tudományos</b></p> <p><b>tudományos</b></p> <p><b>tudományos</b></p> <p><b>szépirodalom</b></p> <p><b>szépirodalom</b></p> <p><b>szépirodalom</b></p> <p><b>szépirodalom</b></p> <p><b>hivatalos</b></p> <p><b>hivatalos</b></p> <p><b>hivatalos</b></p> <p><b>hivatalos</b></p>

két csapat között. Az indiai alakulat főnöke, Vijay Mallya nem **éti**, hogy miért három évvel a szerződés megkötése után merült fel sőtartalom legalább 19 ezreléknyi. Az emberi fogyasztásra alkalmas **éti** - vagy európai - osztriga teljes kifejlődését - amikor már lilára mulatt ámitó előtelnek zsirponty hal előre; molett, ám **éti** francba, a cigányokról én akartam beszélni Höl Lelő romalét-témát illatára emlékeztet. Nem téveszthető össze semmivel, kitűnő **éti** gomba. Fiatal példány Forrás Dr. Batlő Béla: Ismerjük meg Gombáinkat ízletes rizike vagy a lucfenyvesi rizike, ezek azonban mind kitűnő **éti** gombák. Mérgező gombával nem lehet összetéveszteni. Forrás Thaiföld)őverbar; EN 1784 EN 1786 EN 1788őverbar; Rákok, kagylók, **éti** kagylók és egyéb vízi állatok, beleértve a belőlük készült állatjóságokat. A módszer hasonló a Paulnál beválthoz: a polip **éti** kagylót választott két dobozból, amelyeken az adott csapatok Néhány osztrigafaj emberi fogyasztásra is alkalmas, például az **éti** osztriga (*Ostrea edulis*), amelyet alacsony zsír- és szénhidrátartalma százaléka csendes-óceáni osztriga (*Crassostrea gigas*), 1 százaléka **éti** osztriga (*Ostrea edulis*). Az osztrigák kizárólag friss árúként a Rhode Island államok partmenti vizelbe is. Megjelenése Az **éti** osztriga 10 - 15 centiméter hosszú, meglehetősen csúnya és északnyugati partjainál is fellelhető. Az 1940- és 1950-es években, az **éti** osztriga letelepedett a Maine és a Rhode Island államok partmenti osztrigafélék (*Ostreidae*) családjába tartozó faj. Előfordulása Az **éti** osztriga minden európai tengerparton előfordul, ahol a víz szürkésbarna, az alsón gyakran vöröses foltok láthatók. Életmódja Az **éti** osztriga szilárd aljzaton él a parti hullámerésnek kitett honlapján Források Dán kultúra Rapperek Atobot Éti osztriga Az **éti** osztriga vagy európai lapos osztriga (*Ostrea edulis*) a kagylók cukorbeteg, májbeteg, HIV-fertőzötték) vérmérgezést okozhat. Az **éti** osztrigák különösen táplálók és egészségesek. Energiatartalma vásárlásom egyébként három újsg, egy detektívregény, fél kiló **éti** spárga, egy köteg kapor, két saláta, két üveg fűre di rizingl ígérkezik semmiféle tündökítés érdemével. Fád. Matt. Spröd. Szép szál **éti** spárga, mondta róla, saját beismerése szerint (mindenfajta zöldségpiacon, spárgát venne - nem, hogy felkösse magát! spárgát, **éti** spárgát, hogy elkészítse - hogy F. elkészítse! - egy villegjűff összejöttek 4-en az Astoriában. Vásároltam sajtos rudat, vettem **éti** spárgát, krumplit (ez utóbbit Totyinak), Totyinak fürödni plenáris ülésén döntenek a következő kérdésekben; / / a városkert! **éti** tanácsok (megalakításuk esetén) és azok végrehajtó szervei alkalmazása kötelező az Önkormányzat tulajdonát k é pező Nadas-r **éti** települési szilárd hulladéklerakó telepre beszállított hulladékok alkalmazása kötelező az Önkormányzat tulajdonát k é pező Nadas-r **éti** települési szilárd hulladéklerakó telepre beszállított hulladékok

Lexical Corp  
Sketch Engine (ver:open-2.59)  
[A Nemzeti Korpusz](#)  
MNSZ2 - v2.  
Nyelváltás: J

Ezzel a módszerrel is – ilyen alacsony elemszám mellett – percek alatt megkapjuk az állandó-sult szókapcsolatok egyszerű statisztikáját, amely azonos lesz a fentebb bemutatott kollokációs vizsgálat statisztikájával: az *éti osztriga* 9-szer fordul elő, az *éti spárga* 4-szer, az *éti kagyló* 2-szer, és az *éti gomba* is 2-szer. A többi találat nyilvánvalóan irreleváns.

A helyesírási szabályzat legújabb kiadása óta az *éticsiga* egybeírandó; mivel a többi *éti* előtagú szó nincs említve, feltehetőleg azokat továbbra is külön kell írunk. Ezzel tehát a magyar helyesírási szabályai egy újabb kivétellel gazdagodtak.

Bár az MNSZ2 szövegeinek összeállításakor az egybeírt *éticsiga* forma még helyesírási hibának számított, az MNSZ2 az *éticsiga* szótőre 68 találatot ad, ebből az elsődleges tisztítás után 64 marad.

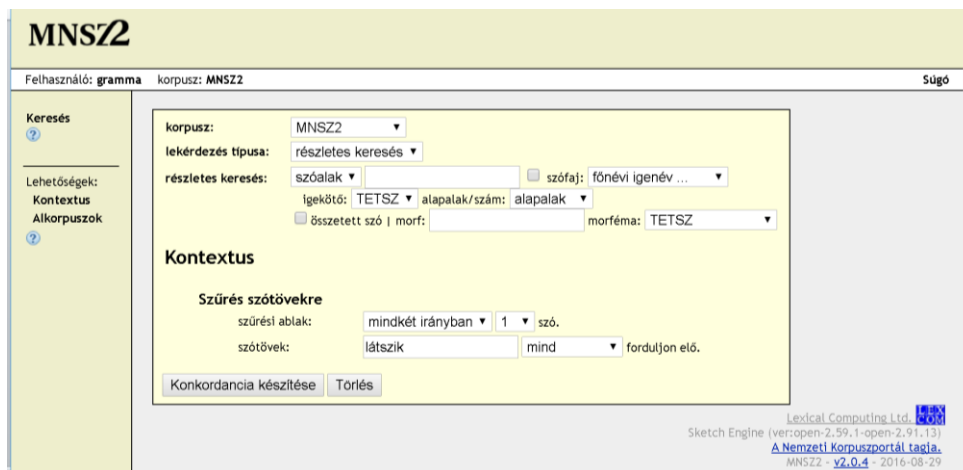
Az alábbi képen ezek egy részét látjuk, a bal oldali kontextus szerinti rendezésben:



ként pedig beírjuk a *látszik* szótövet, úgy, hogy az mindkét oldalon előfordulhasson 1 szó távol-ságra (azaz egymás mellett).

Így az MNSZ2-ben található összes olyan főnévi igenevet megkapjuk találatként, melynek közvetlen környezetében a *látszik* ige valamely alakja áll.

Az alábbi ábra a szükséges beállításokat mutatja:



Nézzük meg az első néhány találatot a 17 053 közül, úgy, hogy a „keverés” opció kiválasztásával különböző alkorpuszok találatai jelenjenek meg:

Keresés

korpusz: MNSZ2  
találatok száma: 17053  
oldalszám: 1 / 12 Ugrás következő utolsó

személyes-közösségi	Az úgy 2013. szeptemberében kapott levelben 0 Ft -os sztorószámra és bocsánatkérés mellett <b>rendeződni látszott</b> , ám Önök ma ismét tértivevényes levelet küldtek, mely 10204 Ft befizetés elmaradása esetén behajtási csoportnak való átadásról rendelkezett.
személyes-fórum	<b>Beindulni látszik</b> némi moderálás is, ami szintén jó lesz, mert a túl durva topicok eltűnnek.
sajtó	Az első generációs cégekből csak néhány, szerencsés pillanatban indult és kiválóan menedzselte cég <b>látszik megragadni</b> a piacon, viszont velőt rázó trombitálással jelentek meg a régi elefántok.
sajtó	Az elmúlt napok meglehetősen borongós, csípős időjárása a hét végére <b>enyhülni látszik</b> : a legmagasabb nappali hőmérséklet 6-8 fok között várható, de az ország egyes részein akár a 10 fokot is elérheti.
személyes-közösségi	Sajnos <b>beigazolódni látszik</b> a rossz hír az időjárással kapcsolatban.
szépirodalom	Az az Arany, akit <b>folytatni látszott</b> , sem azonos azzal, aki a köztudatban él.
szépirodalom	A Rimbaud-i mű hazai, majd azt követően összeurópai utókorra csak híven <b>követni látszik</b> az eredeti létrejöttének körülményeit.
sajtó	A mostani eredmények ennek mégis <b>ellentmondani látszanak</b> .
szépirodalom	Október közepén, feltehetőleg a lengyel helyzettel összefüggésben, fokozott készültséget rendeltek el, bár ennek <b>ellentmondani látszik</b> , hogy a katonák egy része mezőgazdasági munkára volt kivezényelve.
sajtó	Az ausztrál cenzorok <b>megszelidülni látszanak</b> , viszont új mumus tűnt fel a színén.
sajtó	E precedenstértékű döntéssel <b>eldölni látszik</b> az a jogértelmezési vita, amely arra irányult, hogy milyen létesítményi körre terjedhet ki az önkormányzatok igénye.
sajtó	Ezt <b>látszik igazolni</b> az is, hogy a tavaszon olyasmir történt, amit korábban a helybeliek nem mertek megtenni: március 15-e alkalmából nyílt ünnepséget tartottunk a vidéken, koszorúzással egybekötve.
szépirodalom	Áron nézi ezt, az újját szopogatója és <b>elborzadni látszik</b> a cselekvés hiábavalóságán, majd azon gondolkodik, hogy fér majd be a szobájába a zsiráf, ha látogatóba jön.
hivatalos	Műndezeknek azonban <b>ellentmondani látszik</b> , hogy az elmúlt hetekben házkutatást tartottak a Kereskedelmi Bank főrészvényesének számító belga KBC Bank mostani és előző elnökének, valamint hat vezetőjének otthonában és irodájában.
sajtó	Románia teljesítette az IMF-célokat, a gazdasági egyensúly <b>helyreállni látszik</b> .
sajtó	A szintén kiöltözött műsorvezetők figyelmeztetik a pezsgőjét kavargató, <b>ellazulni látszó</b> Csabát, hogy ne igya le magát teljesen, a másik két finalista rendkívüli idegességről, szívproblémákról számol be.
személyes-fórum	Mintha <b>megfelekedezni látszaná</b> arról, hogy Koszovóban KÉT népirtó aktív: a szerbek és az UCK.
sajtó	Miközben a kormányzati célok szintjén igen nagyra törő tervek <b>látszanak kibontakozni</b> , 4-5 százalékos növekedés, egymillió munkahely, tartósan a bruttó hazai termék 3 százaléka alatti államháztartási hiány, 60 százalékra csökkenő államadósság.
tudományos	= Talán felesleges, de célszerűnek <b>látszik leszögezni</b> még mielőtt a témával tovább haladnánk: a szerb nemzetnek természetesen éppúgy joga van az önrendelkezéshez, mint bármely más nemzetnek.
személyes-közösségi	<b>Befagyni látszik</b> az okostelefon-piac: Lassan lezárul a piac felosztásának időszaka, a...
sajtó	Ezt azonban több programozó nem hagyta annyiban, így örömmel jelenthetem, hogy <b>megoldódni látszik</b> a probléma, így a Windows XP tulajoknak nem lesz feltétlenül kötelező átváltani az új operációs rendszerre csak és kizárólag az új, DirectX 10-es játékok miatt.
sajtó	Bár már minden <b>eldölni látszik</b> , hiszen Magyarországnak új köznevelési és felsőoktatási törvénye van, a Hálózat a Tanszabadságért (HAT) Egyesület mégis egyúttműködésre hívja a kutatókat, a pedagógusokat, a szülőket és a diákokat.
szépirodalom	A nép iránti hűség kérdése ezen a ponton <b>látszik szembeszegülni</b> a legmagasabb szellemi eszményekkel.
sajtó	Több jel is <b>igazolni látszik</b> az ellenzék aggodalmát.

isszajelzés  
\*próbálja ki a Mazzola szközt igek és rgumentumaik özetlen ízgalatához!

A képen látjuk, hogy mind az elsődleges találati szó, a főnévi igenév, mind pedig a kontextusként beírt *látszik* ragozott és jelezett alakja ki van emelve, mégpedig a szokásos módon: vastagítással és piros betűszínnel. A kettő úgy különböztethető meg egymástól, hogy az elsődleges találati szó álló betűvel van szedve, a másodlagos, a kontextus pedig dőlt betűvel.

Ha a **bővebb kontextust** akarjuk látni, az **elsődleges találati szóra** kell rákattintani.

Jegyezzük még meg, hogy ha a megfelelő **jelentés** kiválasztása végett szűrünk, akkor a **szűrési ablakban** mindenképpen a legmagasabb értéket kell beállítani, a 15-öt (valójában még az is kevés). Ellenben ha **szókapcsolatot** keresünk, akkor célszerű minél alacsonyabb értéket beállítani, számolva azonban azzal, hogy a szókapcsolat két szava közé más szó is beékelődhet (pl. *állandó lakhely*: az *állandó és átmeneti lakhely* kifejezésben két szó ékelődik a vizsgált szavak közé).

Ilyenkor minél nagyobb számot adunk be, annál több lesz az irreleváns találat, minél alacsonyabb számot írunk be, annál több releváns találat marad ki; ezért fontos megtalálni a legjobb kompromisszumos megoldást.

## & AZ ALKORPUSZOK JELLEMZŐI

A program kétfajta alkorpuszt ajánl föl a felhasználónak: **régiók** szerintit és **stílusrétegek** szerintit. Ezek a lehetőségek a fő keresőfelületen a baloldali menüpontok közül az „Alkorpuszok” menüpontra kattintva válnak elérhetővé. Itt nyílik lehetőség arra is, hogy saját alkorpuszokat készítsünk magunknak.

Az első kép azt mutatja, milyen a keresőfelület lényegi része az „Alkorpuszok” menüpontra való kattintás előtt:

The screenshot shows the search interface with the following elements:

- Keresés** (Search) with a help icon.
- Lehetőségek:** Kontextus, Alkorpuszok (highlighted).
- Form fields:**
  - korpusz: MNSZ2
  - lekérdezés típusa: részletes keresés
  - részletes keresés: szóalak
  - szó faj: TETSZ
  - összetett szó | morf: (empty)
  - morféma: TETSZ
- Buttons:** Konkordancia készítése, Törlés

A második képen azt látjuk, ami elénk tárul az „Alkorpuszok” menüpontra való kattintás után:

The screenshot shows the search interface after clicking on the 'Alkorpuszok' menu item. The main content area is titled 'Alkorpuszok' and contains the following elements:

- alkorpusz:** a teljes korpusz (dropdown), [info új alkorpusz készítése](#)
- doc.file** (empty text input)
- doc.region** (checkboxes):
  - erdélyi
  - kárpátaljai
  - magyarországi
  - szlovákiai
  - vajdasági
  - Mindent
- doc.style** (checkboxes):
  - beszéltnyelvi
  - hivatalos
  - sajtó
  - személyes-fórum
  - személyes-közösségi
  - szépirodalom
  - tudományos
  - Mindent
- div.column** (empty text input)
- div.type** (empty text input)
- Buttons:** Konkordancia készítése, Törlés

Visszajelzés

Próbálja ki a Mazsola eszközt ígék és argumentumaik közvetlen vizsgálatához!

Amint látjuk, régió szerinti bontásban öt alkorpusz jelenik meg, betűrendben: 1. „erdélyi”, 2. „kárpátaljai”, 3. „magyarországi”, 4. „szlovákiai”, 5. „vajdasági”.

Stílusrétegek szerinti alkorpuszokból a legújabb verzióban hét van, a következők, ugyancsak betűrendben: 1. „beszélt nyelvi”, 2. „hivatalos”, 3. „sajtó”, „személyes-fórum”, 4. „személyes-közösségi”, 5. „szépirodalom”, 6. „tudományos”.

A „személyes-közösségi” alkorpuszban meglévő nagyszámú azonos dokumentum annyira torzítja a statisztikai számításokat, **hogy azokban az esetekben, amikor relatív gyakoriságot kell számítanunk, ezt az alkorpuszt nem fogjuk használni.**

Annak érdekében, hogy ne kelljen minden kereséskor külön kijelölni a többi korpuszt, **saját korpuszt** hozunk létre úgy, hogy a keresőfelületen az „Alkorpuszok” felirat alatt található sorban rákattintunk az „Új alkorpusz készítése opcióra”.

## Alkorpuszok

alkorpusz:  [infó új alkorpusz készítése](#)

Kattintás után ezt látjuk:

Felhasználó: **gramma** korpusz: **MNSZ2**

**Keresés**

alkorpusz:

[Document counts](#) [Tokens](#)

doc.region	#
<input type="checkbox"/> erdélyi	4755179
<input type="checkbox"/> kárpátaljai	3006537
<input type="checkbox"/> magyarországi	1222641139
<input type="checkbox"/> szlovákiai	20662244
<input type="checkbox"/> vajdasági	2472153
<input type="button" value="Mindet"/>	

doc.style	#
<input type="checkbox"/> beszélt nyelvi	92537513
<input type="checkbox"/> hivatalos	118311718
<input type="checkbox"/> sajtó	430303070
<input type="checkbox"/> személyes-fórum	57735231
<input type="checkbox"/> személyes-közösségi	308795170
<input type="checkbox"/> szépirodalom	101726086
<input type="checkbox"/> tudományos	144128464
<input type="button" value="Mindet"/>	

doc.region	#
<input type="checkbox"/> erdélyi	4755179
<input type="checkbox"/> kárpátaljai	3006537
<input type="checkbox"/> magyarországi	1222641139
<input type="checkbox"/> szlovákiai	20662244
<input type="checkbox"/> vajdasági	2472153
<input type="button" value="Mindet"/>	

doc.style	#
<input checked="" type="checkbox"/> beszélt nyelvi	92537513
<input checked="" type="checkbox"/> hivatalos	118311718
<input checked="" type="checkbox"/> sajtó	430303070
<input checked="" type="checkbox"/> személyes-fórum	57735231
<input type="checkbox"/> személyes-közösségi	308795170
<input checked="" type="checkbox"/> szépirodalom	101726086
<input checked="" type="checkbox"/> tudományos	144128464
<input type="button" value="Mindet"/>	

Visszajelzés

Próbálja ki a Mazzola eszközt ígék és argumentumaik közvetlen vizsgálatához!



Itt „az új alkorpusz neve” mezőbe beírjuk az általunk készített új korpusz nevét (ezt mi választjuk meg), majd a „doc.style” alatt kipipáljuk az összes jelölőnégyzetet a „személyes-közösségi” kivételével, ahogy azt a jobb oldali képernyőképen látjuk. Végül rákattintunk az „Alkorpusz készítése” opcióra, s létrehozuk az új, saját korpuszunkat, amely már nem fogja tartalmazni a „személyes-közösségi” alkorpuszt.

Amikor kereséseket indítunk, és a saját korpuszunkat akarjuk használni, az „Alkorpuszok” felirat alatti mezőre rákattintva megjelenik a „teljes korpusz” opció mellett az a név is, amit mi adtunk a saját korpuszunknak, második opcióként. Erre rákattintva már használhatjuk is a saját alkorpuszunkat.

Mivel az MNSZ2 erős túlzással sem mondható felhasználóbarátnak, minden új keresés indításakor érdemes egy pillantást vetni erre a mezőre, ugyanis bizonyos esetekben visszaállítódik a „teljes korpusz”. Ilyenkor újra ki kell választanunk a „saját alkorpusz” opciót.

Az egyes alkorpuszok **méretére** vonatkozó adatokat – regionális bontásban is – a következő, az MNSZ2 oldaláról vett táblázat mutatja (<https://clara.nytud.hu/mnsz2-dev/stat.html>). A számszerű adatok millió szóban vannak megadva, százezer szóra kerekítve):

	magyarországi	erdélyi	szlovákiai	vajdasági	kárpátalji	összesen
sajtó	350,5	0,6	11,6	1,5	0,7	364,8
szépirodalom	77,0	0,8	2,3	0,2	0,4	80,6
tudományos	112,0	1,6	3,3	0,3	0,7	117,9
hivatalos	98,0	0,6	0,2	0,1>	0,3	99,0
személyes	300,3	0,4	–	0,1>	0,4	301,1
beszélt nyelvi	76,2	–	–	–	–	76,2
összesen	1013,9	3,9	17,3	2,0	2,5	1039,7

A táblázat a (v2.0.5.) verzióra vonatkozó adatokat tartalmazza. Egyetlen „személyes” alkorpusz szerepel benne.

Újabbban e táblázat alatt a **„személyes” alkorpusz két részére vonatkozó becsült adatok** is szerepelnek: a „személyes-fórum” alkorpusz becsült mérete 57,9 millió szövegszó, a „személyes-közösségi” alkorpuszé pedig 243,2 millió szövegszó.

Arra nézve viszont, hogy ezek hogyan oszlanak meg régiók szerint, sajnos továbbra sem rendelkezünk adatokkal.

A táblázatban közölt méretadatokról **jelentősen eltérnek** azok, amelyek akkor jelennek meg, amikor az „Alkorpuszok” menüpont alatt az „új alkorpusz készítése” opciót választjuk: a táblázatos számadatok **jóval kisebbnek** mutatják csaknem az összes alkorpuszt, mint az „Alkorpuszok” menüpont alatti „alkorpuszok készítése” opció választásakor megjelenő adatok (kivételek a „személyes-fórum” alkorpusz).

**MNSZ2**

Felhasználó: gramma korpusz: MNSZ2

Keresés

Lehetőségek:  
Kontextus  
Alkorpuszok

Alkorpuszok

alkorpusz: a teljes korpusz [infó új alkorpusz készítése](#)

doc.file

doc.region

erdélyi  
 kárpátaljai  
 magyarországi  
 szlovákiai  
 vajdasági

Mindent

doc.style

beszéltnyelvi  
 hivatalos  
 sajtó  
 személyes-fórum  
 személyes-közösségi  
 szépirodalom  
 tudományos

Mindent

div.column

div.type

Konkordancia készítése Torlés

korpusz: MNSZ2

az új alkorpusz neve:

[Document counts](#) [Tokens](#)

doc.file

doc.region

	#
<input type="checkbox"/> erdélyi	4755179
<input type="checkbox"/> kárpátaljai	3006537
<input type="checkbox"/> magyarországi	1222641139
<input type="checkbox"/> szlovákiai	20662244
<input type="checkbox"/> vajdasági	2472153

Mindent

doc.style

	#
<input type="checkbox"/> beszéltnyelvi	92537513
<input type="checkbox"/> hivatalos	118311718
<input type="checkbox"/> sajtó	430303070
<input type="checkbox"/> személyes-fórum	57735231
<input type="checkbox"/> személyes-közösségi	308795170
<input type="checkbox"/> szépirodalom	101726086
<input type="checkbox"/> tudományos	144128464

Mindent

div.column

div.type

Alkorpusz készítése

Csak találgathatunk, hogy a különbség esetleg abból adódik, hogy a korpusz továbbra is bővül, miközben a méretadatokra vonatkozó táblázat nem frissül (bár annak frissítése még egy nagyon lassú és fáradt munkatárssal számolva sem vehet igénybe tíz percnél többet).

### ÷ 1. Régió szerinti alkorpuszok

Ez az opció lehetővé teszi, hogy külön-külön megvizsgáljuk a **Magyarországra** és a határon túli nagyrégiókra – **Erdélyre, Szlovákiára, a Vajdaságra, a Kárpátaljára** – vonatkozó adatokat, sőt ezeket össze is hasonlítsuk egymással. (A keresőfelületen a régiók – amint fentebb is láttuk – betűrendben szerepelnek, mi nagyságrendben fogjuk őket emlegetni.)

**Alkorpuszok**

alkorpusz: a teljes korpusz [infó új alkorpusz készítése](#)

doc.file

doc.region

erdélyi  
 kárpátaljai  
 magyarországi  
 szlovákiai  
 vajdasági

Mindent

Régió szerinti keresésre olyankor van szükségünk, amikor a célforma másképp használatos Magyarországon és másképp egy vagy több határon túli régióban.

Például a *-nök/-nők* ragos formák mai meglétének vizsgálatakor feltétlenül szükséges a régiók szerinti vizsgálat, mivel tudjuk, hogy ezek leginkább Erdélyben maradtak meg. De az ún. suksükölés vizsgálatában is hasznos a régiók szerinti vizsgálat, mivel a suksükölés valószínűleg erősebben jelentkezik a határon túli magyarok nyelvhasználatában, mint a magyarországiakéban.

Az MNSZ-nek ez a része későbbi keletkezésű, a külső régiók szövegei ugyanis csak 2002-től kerültek be a szövegtárba oly módon, hogy határon túli mivoltuk is kódolva lett.

Mivel a külső régiókban keletkezett, ill. a külső régiókban született vagy élő emberek által írott szövegek (főleg szépirodalmiak) korábban is kerültek be a korpuszba, a határon túli szavak vizsgálatánál egyik **torzító tényező** az, hogy a **szövegtárba 2002 előtt felvett művek, pl. Tamási Áron, Sütő András stb. művei a „magyarországi” alkorpusz részeként jelennek meg.** Ezenkívül a külső régiók sajtótermékeiből is került valamennyi bele a még osztatlan korpuszba, s ezek jelenleg ugyancsak magyarországiaként vannak annotálva.

A megfelelő régiónál található jelölőnégyzetre kattintva, azt ezzel kipipálva máris leszűkítjük a keresést az adott régióra. Az alábbi képen a magyarországi és a vajdasági alkorpusz van kijelölve:

Alkorpuszok  
alkorpusz: a teljes korpusz [infó új alkorpusz készítése](#)

doc.file

doc.region

- erdélyi
- kárpátaljai
- magyarországi
- szlovákiai
- vajdasági

Mindent

Ha nem jelölünk be egyetlen nyelvváltozatot sem, az a keresés szempontjából olyan, mintha mindegyiket bejelöltük volna, és a találatok az egész korpuszból fognak megjelenni.

Alkorpuszok  
alkorpusz: a teljes korpusz [infó új alkorpusz készítése](#)

doc.file

doc.region

- erdélyi
- kárpátaljai
- magyarországi
- szlovákiai
- vajdasági

Mindent

=

Alkorpuszok  
alkorpusz: a teljes korpusz [infó új alkorpusz készítése](#)

doc.file

doc.region

- erdélyi
- kárpátaljai
- magyarországi
- szlovákiai
- vajdasági

Egyiket se

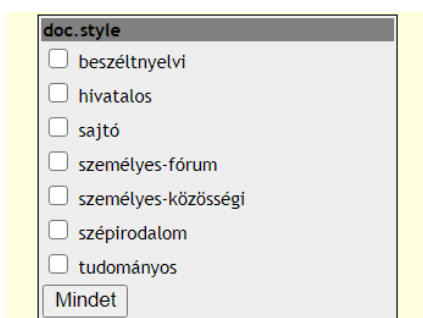
Ezt azok számára kell hangsúlyozni, akik már korábban is használták az MNSZ2-t, ugyanis a v2.0.4-nél ez még nem így volt. Akkor a szövegtár számos annotálatlan dokumentumot tartalmazott, melyek csak akkor jelentek meg, ha nem jelöltük ki az egyik alkorpuszt sem. Ha kijelöltük az összeset, a találatok száma alacsonyabb volt, mert az annotálatlan dokuk nem jelentek meg. Öröndetes, hogy a fejlesztők ezt a problémát kiküszöbölték.

Amint az MNSZ2 méretadataira vonatkozó fenti táblázatból is láthattuk, a határon túli anyag mennyisége az egész anyagon belül csaknem elenyésző, **inkább szimbolikus jelentősége** van. Éppen ezért **a régiókra vonatkozó adatokból csak nagy-nagy óvatossággal vonhatók le (bizonytalan) következtetések.**

## ÷ 2. Stílusréteg szerinti alkorpuszok

A stílusrétegek szerinti alkorpuszok – amint ezt már fentebb is láttuk – a következők: „szépirodalom”, „sajtó(nyelv)”, „tudományos”, „hivatalos”, „beszélt nyelvi”, valamint „személyes-fórum” és „személyes-közösségi”.

A keresőfelületen az egyes alkorpuszok betűrendben így szerepelnek:



Amint említettük, a keresőfelületen az egyes stílusrétegek (alkorpuszok) is betűrendben szerepelnek, akárcsak a régiók. Mi azonban nem betűrendben fogjuk sorra venni az egyes alkorpuszokat, hanem a következő logika szerint:

1. az első helyre a jellemzően választékos stílusú „**szépirodalmi**” alkorpuszt tesszük;
2. a másodikra a részben választékos, részben formális, részben informális stílusú „**sajtó-nyelvi**” alkorpuszt;
3. a harmadik helyre az elsősorban formális stílusú „**tudományos**” alkorpusz kerül;
4. a negyedikre a szintén formális stílusú „**hivatali nyelvi**” alkorpusz;
5. az ötödik helyre a részben formális, részben bizalmas stílusú „**beszélt nyelvi**” alkorpuszt tesszük;
6. a hatodikra az informális, bizalmas stílusú „**személyes-fórum**” alkorpuszt;
7. a hetedikre a szintén informális, bizalmas stílusú „**személyes-közösségi**” alkorpuszt, ám mivel ez az alkorpusz zavaróan nagy számban tartalmaz azonos tartalmú, ill. szövegű dokumentumokat, ezt az alkorpuszt a vizsgálataink során csak olyankor fogjuk használni, amikor nem számítunk relatív gyakoriságot. Amikor számítunk, egy olyan új korpuszt hozunk létre a magunk számára, amely ezt az alkorpuszt nem tartalmazza (l. fentebb).

Ahhoz, hogy az alkorpuszok szerinti elemzés során nyert adatokat megfelelően tudjuk **értékelni**, figyelembe kell vennünk az egyes alkorpuszok sajátosságait. Ezekről adunk most egy rövid áttekintést.

Ahogy fentebb jeleztük, az alkorpuszok méretére nézve két adatot adunk meg. Az első a weboldalon közzétett táblázatból való, amely még a szövegtár (v2.0.4.) verziójához készült, de utólag ki lett egészítve a két „személyes” alkorpusz méretére vonatkozó „becsült” adatokkal. A

második adatsor akkor jelenik meg, ha a keresőfelületen az alkorpuszok megnyitása után az „Az új alkorpuszok készítése” lehetőségre kattintunk.

### ÷ A) A „szépirodalmi” alkorpusz sajátosságai

A „szépirodalmi” alkorpusz mérete **80,6**, ill. **101,7 millió szövegszó**; ez a második legkisebb terjedelmű alkorpusz, nála csak a „beszélt nyelvi” alkorpusz kisebb. Az alkorpusz a Digitális Irodalmi Akadémia (<https://dia.hu/>) anyagának egy részét tartalmazza.

Az akadémia honlapjáról megtudhatjuk, jelenleg kik a tagjai az akadémiának, de azt nem, kik voltak a tagjai akkor, amikor az MNSZ2-be bekerült az akadémia anyaga, így nem is derül ki, kiknek a műveit tartalmazza a „szépirodalmi” alkorpusz. Csak az MNSZ2 konkrét találataiból látjuk, kik azok, akiknek legalább egy műve biztosan benne van a Digitális Irodalmi Akadémiában.

Az alkorpusz adatainak értékelésénél ezeket a tényeket kell figyelembe venni:

**a)** A régiók szerinti vizsgálatot nehezíti az a tény, hogy **a határon túli szerzők egy része a „magyarországi” alkorpuszban szerepel**, mivel ezek még akkor kerültek be az MNSZ-be, amikor még nem voltak határon túli alkorpuszok, s úgy látszik, rendkívül bonyolult, nehezen megoldható feladat módosítani a határon túli írókra vonatkozó dokumentumok annotációját.

**b)** Objektív nehézséget jelent, hogy vannak írók, akik **életük egy részét Magyarországon kívül, más részét Magyarországon** élték le, így besorolásuk eleve problematikus.

**c)** További probléma, hogy **az adatbázisba bevett írók köre elég szűk**: a Digitális Irodalmi Akadémia jelenleg 170 tagot számlál, a 10-es évek közepén, amikor a szövegtárat bővítették, ez a szám 100 körüli lehetett. Ez látszólag nem is olyan kevés, ám egyfelől egyes szerzők műveiből valószínűleg nem sok tartozik az adatbázisba (ezt abból gondoljuk, hogy nevük viszonylag ritkán bukkan fel kereséskor), másrészt vannak köztük olyanok is, akiknek a nyelvhasználata sehogyan sem tekinthető mai magyar nyelvinek.

Így például benne van az MNSZ2-ben Áprily Lajos (1887–1967), Kassák Lajos (1887–1967), Füst Milán (1888–1967), Tersánszky Józsi Jenő (1888–1969), Déry Tibor (1894–1977), Tamási Áron (1897–1966), Németh László (1901–1975), Illyés Gyula (1902–1983), és a sort még sokáig folytathatnánk további olyan írókkal, akik ugyan valamivel később születtek és haltak meg, de műveik **zavaró heterogenitást** visznek az adatbázisba, melyben több alkorpusz is ténylegesen a mai magyar nyelvet képviseli.

**d)** Számos írói mű **többször is benne van** az adatbázisban, de az ismétlődések – mint a többi alkorpuszban is – az elsődleges tisztítás során kiküszöbölhetők.

**e)** Más tévedések is akadnak. Például, amint már fentebb is említettük, egy magyarországi születésű, Nyugaton élő írónak, Balogh Bertalannak ugyanaz a műve kétszer is bent van az adatbázisban, egyszer szlovákiai, másodszer erdélyi magyarként.

Balogh Bertalan egyébként nem tagja a Digitális Irodalmi Akadémiának, amiből arra következ-  
tethetünk, hogy más hasonló írók is akadnak az adatbázisban. Ezek nyilván a bővítés során ke-  
rültek be, de hogy kikről van szó, és honnan származnak a bevett dokumentumok, arról az  
MNSZ2 weboldalán nem kapunk tájékoztatást.

**f)** Az alkorpuszba belekerültek irodalomtudományi szaknyelvi szövegek is, amennyiben ezek  
szerzői írók voltak, s a korpusz összeállítói valószínűleg nem távolították el a kötetek esetle-  
ges bevezető vagy záró tanulmányát sem, így kisebb mértékben olyan szavak is előfordulnak  
ebben a stílusrétegben, amelyek inkább a szaknyelvi regiszterek jellemzői.

**g)** A szaknyelvekkel azért is érintkezik a „szépirodalmi” alkorpusz, mert történelmi visszaem-  
lékezések is vannak benne, s ez a szövegeket számos, amúgy inkább a hivatali nyelvre vagy a  
történettudomány szaknyelvére jellemző szavakkal, kifejezésekkel telíti.

Ez objektív nehézség: azt a tényt tükrözi, hogy az egyes nyelvi regiszterek nem különíthetők  
el élesen egymástól.

**h)** Az adatok értékeléséhez minden esetben figyelembe kell venni, hogy írói narrációról  
van-e szó, vagy pedig az író egy regényhőst beszéltet; a regényhős beszédstílusa radikálisan  
eltérhet a regényíróétól.

A regényhős beszédstílusa nem a szépirodalmi stílust képviseli, hanem egyfajta vendégszö-  
vegnek tekinthető még akkor is, ha azt az író alkotta.

Nézzünk meg néhány találatot az irodalmi stílusértékű *bámész* szó-  
ra a „szépirodalmi” alkorpuszból, a mondatkontextus megjeleníté-  
sével.

Mivel most csak a „szépirodalmi” alkorpusz találataira vagyunk kí-  
váncsiak, a stílusok szerinti alkorpuszoknál kipipáljuk a „szépiroda-  
lom” melletti jelölőnégyzetet, hogy csak ennek az alkorpusznak a  
találatai jelenjenek meg.

Keresés	korpusz: MNSZ2 találatok száma: 265
Mentés	<b>szépirodalom</b> S áldott lesz a <b>bámész</b> leánykák között a paradicsomos kertben eltűnő, az ivarérett és isteni ünő.
Megjelenítés KWIC/mondat	<b>szépirodalom</b> Úgy hull a sajtó semmisség ölebe A hang, a lélek ismeretlenül, Mint elfelejtett védő, senki sarja, Fölösleges <b>bámész</b> lenyilazott - S a mindenség pusztán süvit tovább...
Rendezés bal   jobb találat info keverés	<b>szépirodalom</b> Akkor nem bántam, nyár jött, fellázadtunk, a Sétatér szegletébe szaladtunk, volt nyulánk bokrunk, romladék, volt forrás, <b>bámész</b> közösség; próbahelyre lettünk!
Véletlen minta Szűrés	<b>szépirodalom</b> Irodalmi Szemle, Pozsony (2004/8) A <b>bámész</b> civil, Szlovákiai magyar szép versek 2005
Gyakoriságok szóalakok szótövek elemzések dokumentumok	<b>szépirodalom</b> A köznapi szóhasználatban az Ali Baba egyszerűen tolvajt jelent, s én ebben a kevés bizalommal megelégedeztem, ám légkondicionált, vezetékes vizzel, hűtőszekrényvel, társalgóval, előzékeny személyzettel ellátott " luxusszállóban " laktam több mint egy évig, itt edződtem <b>bámész</b> vendégből szakavatott helybelivé.
Kollokációk	<b>szépirodalom</b> A tisztára sepert kavicsos út két oldalán <b>bámész</b> cselédek várokoztak, hogy az úri vendégek után betödulhassanak Isten hajlékába.
	<b>szépirodalom</b> A hazainduló násznép még csomagolt és búcsúzkodott, de a kapu előtt már újabb csoportok tüntek fel: rendőrök, <b>bámész</b> falusiak, mindenekelött azonban újságírók és fotóriporterek, a fővárosi és megyei lapok tudósítói.
	<b>szépirodalom</b> Sok <b>bámész</b> ember ott áll körül.
	<b>szépirodalom</b> És az álékra kiváltó <b>bámész</b> vaddal is.
	<b>szépirodalom</b> <b>Bámész</b> tehének édeskes szagával az a lány üzent újra: a havasi gyopár száraz csillagát ő küldte utánad a világ tetejéről.
	<b>szépirodalom</b> Ha botránnyhős kell a világnak, aki a <b>bámész</b> fejek különlegesség iránti érdeklődését kielégíti, odaáll ő maga: Közép-Európa és a világ tánctermeinek padlóját ugyanolyan intenzitással tapodja, mint bármely parasztház szobájának döngölt agyagját.
	<b>szépirodalom</b> A különbség talán az, hogy míg a <b>bámész</b> sétáló nem álmodik különleges kirakatokról, elfogadja, amit lát, én folyton valami meglepőt várok az újságtól.
	<b>szépirodalom</b> A két <b>bámész</b> lány, s nyomukban a két kenyeret majszoló kölök is ott áll szájátva a létra alatt.
	<b>szépirodalom</b> Ha valaki fölvilágosít akkor, miben veszek részt, milyen hagyományt " ápolok ", talán kiállok a sorból, és a magyarázatot éppoly elképedő arccal hallgatom, mint amilyen átéléssel ijesztettem az anyjuk szoknyájába kapaszkodó, <b>bámész</b> kislányokat.
	<b>szépirodalom</b> Csak palánkrágó ugatás, csak partra vágó hullámok kísérték hazáig, mert mégsem igazi úrleány, mert mégiscsak a kupleros leánya - azután, mert fülében nem szünik a deszpa, s mert itthon is arcát sütik azok a hiperanganzsínú lámpaernyők, hát minden holmit ledobva magáról, a vizzel teli nyagylavóba áll - két ifjú mell, két hosszú, duzzadó comb s a szerelem lakatlan fecskéfszke a <b>bámész</b> has szép erese alatt -
	<b>szépirodalom</b> Az öreg Harasztai a <b>bámész</b> csoport szélén eregeti a füstöt, ha már nem férnek mögöttes, előbbre lép egy rövidet, mintha ő lenne a vezér és a hirtelen támadt kíváncsiság merészsége.
	<b>szépirodalom</b> Aztán meg részvét tárgya, hirdették, dobolták jövetelüket, mégis több a <b>bámész</b> , mint a segítő.

## ÷ B) A „sajtónyelvi” alkorpusz sajátosságai

A „sajtónyelvi” alkorpusz mérete **364,8**, ill. **430,3 millió szövegszó**; ez az összes alkorpusz közül a legterjedelmesebb.

Az MNSZ weboldaláról mindössze annyit tudunk meg az ide tartozó szövegekről, hogy „[s]zéles skáláját mutatják be a nyelvi változatoknak, vertikálisan és horizontálisan is”. Az adatok igazán alapos értékeléséhez ennél jóval több információra volna szükségünk.

**a)** Amikor a „sajtónyelvi” alkorpuszban talált adatokat használjuk fel a szó stílusértékének megállapításához, figyelembe kell venni, hogy **az írott sajtónyelven belül vannak többé-kevésbé informális stílusú műfajok is** (pl. a glossza); ezenkívül pedig a sajtóban sok olyan idézet olvasható, ami szóban hangzott el (pl. amikor a riporter kérdést tesz föl egy politikusnak).

Az elmondottakból következik, hogy ha egy szóra sok a „sajtónyelvi” korpuszból származó példa, ez nem azt jelenti, hogy ne lehetne informális, bizalmas stílusértékű vagy akár szleng jellegű is.

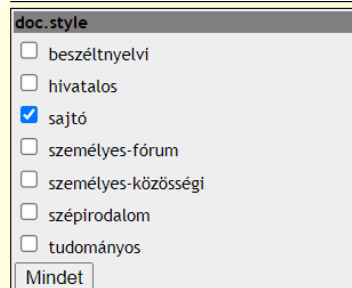
**b)** Sokat segíthetne az adatok értékelésében, ha tudni lehetne például, milyen szövegeket kell érteni „sajtószövegeken”, **milyen típusú lapok** lettek az alkorpuszba beválogatva.

Fontos volna például tudni, hogy vannak-e köztük vidéki lapok is, vagy csak országosak (nem tekintve most vidékinek a határon túliakat); milyen arányban vannak képviselve a napi-, a heti-, a havilapok; továbbá vannak-e köztük magazinok, szakmai lapok, bulvárlapok stb., mivel ezek részben eltérő műfajú szövegeket tartalmaznak.

Sajnos az egyes dokuk annotációjában sem találjuk meg a sajtótermékek nevét, megjelenési helyét, idejét és egyéb releváns adatokat, csak egy kódot, amelyből jó esetben kiolvashatjuk a lap címét.

**c)** Bizonyos találatokból arra következtethetünk, hogy a „sajtónyelvi” alkorpuszba (olykor?) bekerültek a cikkek internetes változatához írott olvasói kommentek is, amelyek egészen más stílust képviselnek, mint maga a cikk, melyhez a kommentelők hozzászóltak: nyilván jóval nagyobb arányban informálisak, mint a sajtónyelvi szövegek többsége.

A kommenteket ezért célszerű lett volna a „személyes-fórum” alkorpuszhoz tartozóként annotálni.



A screenshot of a document style selection menu. The title bar reads "doc.style". Below it, there is a list of options with checkboxes:

- beszélnyelvi
- hivatalos
- sajtó
- személyes-fórum
- személyes-közösségi
- szépirodalom
- tudományos

At the bottom of the list is a button labeled "Mindet".

Néhány találat a „sajtónyelvi” stílusértékű *csomagterv* szóra a „sajtónyelvi” alkorpuszból, egy mondatnyi kontextusban való megjelenítéssel:





Ezek NEM a tudományos szaknyelv elemei, hanem a szerzők **mintegy idézik őket meta-nyelvi adatokként** a tudományos közleményekben, így ezeket nem szabad figyelembe venni a szó stílusértékének megállapításakor.

Az ilyen adatokat a másodlagos tisztítás során feltétlenül el kell távolítani.

**d)** Az alkorpuszban **Wikipédia-szócikkek** is vannak. Ezek egy részét nem szakemberek, hanem **lelkes laikusok** írták. Nem csoda hát, hogy ezekben is előfordulnak olyan nyelvi formák, amelyek egyébként nem jellemzőek a tudományos stílusra.

**e)** Úgy tűnik, hogy a „tudományos” alkorpuszban **túl vannak reprezentálva a társadalom-tudományok**, s nagyon alacsony a természettudományos szakszövegek aránya.

Erre nézve az MNSZ weboldala nem közöl adatokat, csak a tapasztalatok alapján mondjuk: a találatok közt jóval ritkábban jelennek meg természettudományi, mint társadalomtudományi szövegek.

Néhány találat a „tudományos” minősítésű **decimális** szóra a „tudományos” alkorpuszból, mondatkontextus megjelenítésével:

**doc. style**

beszélnyelvi

hivatalos

sajtó

személyes-fórum

személyes-közösségi

szépirodalom

tudományos

Mindent

<p>Keresés</p> <hr/> <p>Mentés</p> <p>Megjelenítés KWIC/mondát</p> <p>Rendezés bal   jobb találat infó keverés</p> <p>Véletlen minta Szűrés</p> <p>Gyakoriságok szóalakok szótövek elemzések dokumentumok</p> <p>Kollokációk</p> <hr/> <p>Visszajelzés</p> <p>Problálja ki a Mazzola eszközt igék és argumentumaik közvetlen vizsgálatához!</p>	<p>korpusz: MNSZ2 találatok száma: 252</p> <p><b>tudományos</b> A métermértékekkel együtt rendelték el a <b>decimális</b> rendszer, a tizes váltószámok alkalmazását.</p> <p><b>tudományos</b> Már nagyon régen gondoltak rá, hogy el kellene térni az eddig használt időmértéktől: az időmérésben is át kellene térni a <b>decimális</b> rendszerre.</p> <p><b>tudományos</b> A másodperc tört részét már a világon mindenütt <b>decimális</b> rendszerben számolják, ám ennél tovább egészen mostanáig nem jutottunk.</p> <p><b>tudományos</b> Végül is a bielik nem bizottságosoztak, mint valamikor az elődök a zónaidő ügyében, hanem egyszerűen megtekették városukat a " világ közepének ", a Greenwich Mean Time (GMT) helyett bevezették a Biel Mean Timeot (BMT), s a <b>decimális</b> időszámítás problémáját is megoldották.</p> <p><b>tudományos</b> Káplány Géza: Az egyetemes <b>decimális</b> osztályozórendszer rövidített magyar nyelvű kiadása.</p> <p><b>tudományos</b> - Esetenként szükség van a bináris kódok <b>decimális</b> (tizes rendszerbe való) alakítására: </p> <p><b>tudományos</b> - egy globális méretu koordináta-rendszerben, ahol a terület mérete 10.000 km2 és a pontosság 1 mm, 10 <b>decimális</b> számjegyre és 30 bináris számjegyre volna szükség</p> <p><b>tudományos</b> Mind a három koordináta halmazra becsülje meg, hogy (a) mennyi <b>decimális</b> és (b) mennyi bináris számjegy szükséges a válasz tárolására.</p> <p><b>tudományos</b> - ez az eljárás minden következő pontnál, lehetővé teszi, hogy négy <b>decimális</b> számjegyet használjunk 12 helyett, vagy megközelítően 12 bit-et, 36 helyett</p> <p><b>tudományos</b> - két levél szomszédos, ha bináris megfelelőik bináris 1-el vagy bináris 10-el különböznek ( <b>decimális</b> 1 vagy 2) a mozaik aritmetikában</p> <p><b>tudományos</b> - példa: 01 és 03 szomszédos, mert a 0001 és 0011 bináris 10-el, illetve <b>decimális</b> 2-vel különböznek</p> <p><b>tudományos</b> - valós ( <b>decimális</b> ) számok</p> <p><b>tudományos</b> A mező általában két részből áll: - a z o n o s í t ó (vagy hívójel) <b>decimális</b> szám: 01-től n-ig (rendszer/softver függő, hogy hány mezőt enged meg -, tehát mekkora a rekord terjedelme), - t a r t a l o m A mezők osztályozhatók ismételtetőség szempontjából is: - i s m é t e l h e t ő (ha 1 rekordban többször fordul elő), - n e m i s m é t e l h e t ő (ha nem, vagy csak 1-szer fordul elő).</p> <p><b>tudományos</b> A számítógép " lelke " a mikroprocesszor, egyetlen integrált áramkör, amely képes valamennyi aritmetikai és logikai műveletet elvégezni egy 8 vagy 16 jegyű bináris (kettes számrendszerbeli) számon, ami vagy egy, vagy pedig két <b>decimális</b> számot, betűt vagy írásjelet jelöl, valamint vezérli az egész gép működését.</p> <p><b>tudományos</b> A lebegőpontos értéket <b>decimális</b> megjelenésére konvertálja.</p> <p><b>tudományos</b> Az adott lebegőpontos értéket <b>decimális</b> megjelenésére konvertálja.</p> <p><b>tudományos</b> Az adott lebegőpontos értéket <b>decimális</b> megjelenésére konvertálja.</p> <p><b>tudományos</b> Egy ilyen összeköttetés egy híd azért nem képes megvalósítani, mert a helyi Ethernet hálózat címei 48 bitesek, míg egy X.25-ös WAN címei 14 <b>decimális</b> jegyből állnak, és ezek egymással inkompatibilisek.</p> <p><b>tudományos</b> IP address - az Internet hálózathoz kapcsolt gép azonosító kódszáma, rendszerint négy <b>decimális</b> számból áll</p> <p><b>tudományos</b> Írjuk fel példaként a 39715 <b>decimális</b> számot bináris formában:</p> <p><b>tudományos</b> <b>Decimális</b> számrendszerből hexadecimálisra hasonló eljárással válthatunk, mint a kettes számrendszer esetén, csak az osztó itt a 16 lesz.</p> <p><b>tudományos</b> érték % előjel nélküli, egész típusú <b>decimális</b> szám és százatékjel, amely egy százalékos arányt fejez ki.</p> <p><b>tudományos</b> Aszerint, hogy e regiszterekben melyik bitet állítjuk vagy töröljük a megfelelő <b>decimális</b> szám beírásával, különböző modemparancsot adhatunk ki.</p> <p><b>tudományos</b> S3 elhelyezi a carriage return ASCII kódját <b>decimális</b> alakban.</p> <p><b>tudományos</b> S4 elhelyezi a line feed ASCII kódját <b>decimális</b> alakban.</p> <p><b>tudományos</b> S5 elhelyezi a backspace ASCII kódját (8) <b>decimális</b> alakban.</p> <p><b>tudományos</b> Az Internetto Hírlevele a webes ékezeteket használja, amelyek a CWI kódolástól eltérően a <b>decimális</b> 193 és 254 számú kódok között helyezkednek el.</p> <p><b>tudományos</b> Ez a következő lesz: eredm= 16 Hex= 00000010 Octal= 00000000020 azaz a számítási eredmény <b>decimális</b> , hexadecimális és oktális formában lett az eredm szimbólumhoz hozzárendelve.</p> <p><b>tudományos</b> Amikor valaki számot, teljesen természetesen alkalmazza a <b>decimális</b> , vagyis a tizes számrendszert.</p>
---	--

## ÷ D) A „hivatalos” alkorpusz sajátosságai

A „hivatalos” alkorpusz mérete **99,0**, ill. **118,3 millió szövegszó**; ezzel a terjedelemmel a korpuszok sorában a negyedik helyen áll (a „szépirodalmi”, a „beszélt nyelvi” és a „személyes-fórum” kisebb még nála).

**a)** Az alkorpusz szövegei az MNSZ weboldalán található leírás szerint „szabályokat, törvényeket, rendeleteket, parlamenti vitákat tartalmaznak”.

Nem tudhatni, a felsorolt szövegtípusok milyen arányban fordulnak elő, abból viszont, hogy a találatok közt leginkább a parlamenti és a fővárosi önkormányzati jegyzőkönyvekkel találkozunk, **gyanítható, hogy ezek vannak túlsúlyban**.

**b)** A parlamenti és önkormányzati beszédek **szóban** hangzottak el, s bár a felszólalók nyilván többnyire **írott szöveg** alapján mondták el őket, **nem feltétlenül mindig olvasták őket fel szó szerint**.

Az adatok megfelelő értékeléséhez hasznos volna tudni, hogy a szövegtárba bekerült szövegek a ténylegesen elhangzott felszólalások átíratái-e, vagy esetleg a képviselők által készített írott változatok.

Az is lehet, hogy az átíratot készítőknél az írott változat rendelkezésükre állt, s átvették belőle a szó szerint felolvasott részeket; erre nézve is jó lenne, ha a korpuszleírás tartalmazna információt.

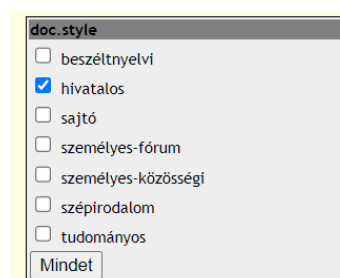
Tapasztalataink szerint az MNSZ2 „hivatalos” alkorpuszának szövegeiben bőven akadnak **expresszív** és más **beszélt nyelvi** szavak és kifejezések is, olyanok, amelyek egyáltalán nem tekinthetők a hivatali nyelvi stílusra jellemzőnek.

Az ilyen expresszív vagy más beszélt nyelvi szavak előfordulása a hivatalos nyelvi korpuszban tehát nem bizonyítéka annak, hogy ezek a formális írott nyelv, ill. a hivatali szaknyelv részei volnának.

Az informális stílusú nyelvi formákat a hivatali nyelvi szövegekben vagy vendégnyelvi szövegekként értékeljük, és kihagyjuk őket a vizsgálatból, vagy pedig a hivatali nyelvnek mint regiszternek különféle stílusú válfajait különböztetjük meg: a formális mellett a bizalmas stílusút, de hasznos lehet külön számontartani kissé formális vagy kiegyenlített hivatali nyelvi stílust is.

Az előbbi megoldást (az informális stílusú szövegek kizárása) akkor érdemes alkalmazni, ha a hivatali nyelvi regiszterek képezik a vizsgálat tárgyát, az utóbbit (stílusrétegek megkülönböztetését) pedig akkor, ha az érdekel bennünket, hogy egy olyan szintéren, mint amilyen a parlament ülésterme, milyen stílusú diskurzusok fordulnak elő.

Nézzünk meg néhány példát a formális stílusértékű, a hivatali (és más) szaknyelvi regiszterekhez tartozó *dologi* szóra a „hivatalos” alkorpuszból, egy mondatnyi kontextusban való megjelenítéssel:



doc. style	
<input type="checkbox"/>	beszéltnyelvi
<input checked="" type="checkbox"/>	hivatalos
<input type="checkbox"/>	sajtó
<input type="checkbox"/>	személyes-fórum
<input type="checkbox"/>	személyes-közösségi
<input type="checkbox"/>	szépirodalom
<input type="checkbox"/>	tudományos
Mindet	

<b>hivatalos</b>	El kell mondani, hogy a <b>dologi</b> jellegű működési kiadások szinten tartására, az inflációs hatások ellensúlyozására a 2004. évre benyújtott előterjesztés nem számol dologi automatizmussal.
<b>hivatalos</b>	A megyei önkormányzat a kulturális szférával 106 millió forintos költségcsökkentésben egyezett meg, amely személyi leépítésekben és <b>dologi</b> kiadások jelentős csökkentéséből takarítható meg.
<b>hivatalos</b>	- a biztosítékként adott vagy <b>dologi</b> jogok tárgyát képező közösségi növényfajta-oltalom, vagy
<b>hivatalos</b>	<b>Dologi</b> kiadásra 2-3 százalékos növekedés: ez egyszerűen riasztó, hiszen ismerjük az intézmények, a laktanyák, a haditechnika állapotát, erre 2-3 százalékos növekedést számolni - több mint 20 százalékos infláció mellett - elképesztő.
<b>hivatalos</b>	Azonban tudnunk kell, hogy a fent említett normatívaemelés az intézmények többségében nem fedi le a közalkalmazottak átlagos 50 százalékos béremelését, illetve a <b>dologi</b> kiadásokat.
<b>hivatalos</b>	A bevételek alakulása segítette az intézményeknél a többletfeladatok fedezetének részbeni megteremtését és az inflációval összefüggő <b>dologi</b> , felhalmozási kiadásnövekedések többleteinek ellentételezését.
<b>hivatalos</b>	h) <b>dologi</b> jog alapítását vagy átruházását a rendelet 19. cikkének megfelelően, valamint a dologi jog természetét;
<b>hivatalos</b>	Azok támadnak ma bennünket a <b>dologi</b> kiadások 2,5 százalékos növekedése miatt, akik az elmúlt négy évben egyetlenegy fillért nem tettek be ebbe az egészségügybe.
<b>hivatalos</b>	az " 5501 Budapesti Történeti Múzeum " cím <b>dologi</b> kiadási előirányzatát
<b>hivatalos</b>	Örömmel mondhatom, hogy itt a tárca olyan tervezést hajtott végre, amelyben 99 milliárd 540 millió forintról 121 milliárd 763 millió forintra emelte a <b>dologi</b> kiadásokat.
<b>hivatalos</b>	A másik a <b>dologi</b> kiadások forrásainak megnövelése.
<b>hivatalos</b>	Tessék már megnézni a költségvetés különböző fejezeteit, hogy a <b>dologi</b> kiadások milyen mértékben lecsökkennek, és tessék nekem egy dolgot elhinni: ha veszem Ajkát vagy Ajka térségét, ott ötven intézmény működik, a tűzoltóság dologi kiadása nagymértékben lecsökkent, a kórház esetében a bér jellegű alaptámogatás 60 százalékról fölment 80 százalékra; tehát lecsökkent, 20 százalék maradt a dologi kiadásokra.
<b>hivatalos</b>	(3) A <b>dologi</b> jellegű kiadások előirányzata 6.277.929 ezer Ft
<b>hivatalos</b>	S a <b>dologi</b> megszorítások miatt felmerült, hogy ha a laboratóriumokat nem tudja fenntartani az állam, akkor ugyan már, dobjanak össze egy kft.-t, és végezzék állami feladataikat vállalkozásban.
<b>hivatalos</b>	Magyarul, a <b>dologi</b> lassan már a nullához fog konvergálni.
<b>hivatalos</b>	Ugyanakkor az meglehetősen aggályos, hogy a <b>dologi</b> kiadások központi költségvetési hozzájárulásból történő növelésére nincs mód.
<b>hivatalos</b>	A <b>dologi</b> kiadásoknál sajnálatos módon már tragikusabb a helyzet, ott sajnos az idei, 2006-os 19,4 milliárdos bázisához képest mintegy 600 millió forintos csökkenés tapasztalható.
<b>hivatalos</b>	Nagyon sok hozzászólásban kitértek képviselőtársaim - akár a mezőgazdaság, akár az önkormányzatok, akár más intézmények esetében - a <b>dologi</b> kiadásra.
<b>hivatalos</b>	168. § (1) A természetbeni juttatások értékét, valamint a <b>dologi</b> kár összegét a kártérítés megállapításakor irányadó fogyasztói ár alapján kell meghatározni.
<b>hivatalos</b>	Ha nem indítanak másodlagos eljárást, a <b>dologi</b> joggal érintett vagyontárgyak értékesítéséből keletkező többletet a főeljárás felszámolója számára kell kifizetni.
<b>hivatalos</b>	De hogy Budapesten maradnak a központi szervek ugyanolyan létszámmal, ugyanannyi objektummal, ugyanolyan <b>dologi</b> kiadással, ebből mintha nem akarnánk problémát csinálni.
<b>hivatalos</b>	A 143. pontban Dobos Krisztina a felsőoktatási speciális programok <b>dologi</b> kiadásainak előirányzatát kívánja megemelni.
<b>hivatalos</b>	Magyarul, a <b>dologi</b> lassan már a nullához fog konvergálni.

## ÷ E) A „beszélt nyelvi” alkorpusz sajátosságai

A „beszélt nyelvi” alkorpusz terjedelme **76,2**, ill. **92,5 millió szövegszó**.

Az alkorpusz az MNSZ2 weboldalán található tájékoztatás szerint „az MR1 Kossuth rádió (*sic!*) bizonyos anyagait öleli fel az (*sic!*) 2004–2012 évekből, felolvasott szöveget (hírek) és spontán beszélgetést (riportok) vegyesen”.

Ebből a leírásból az is kiderül, hogy ez az alkorpusz nem tartalmaz határon túli anyagot, pedig van rádiózás mind a négy bevont határon túli régióban.

**a)** A felolvasott szövegek – ahogy azt tanultuk – az **írott nyelvet** képviselik, csupán hangosítva vannak, ezért a szerkesztők **rendkívül súlyos szakmai hibát** követtek el ezek bevonásával az MNSZ2 „beszélt nyelvi” alkorpuszába.

**b)** A riportok szövegei viszont valóban tartalmaznak beszélt nyelvi elemeket, bár a szövegek jellegzetességeinek vizsgálatakor figyelembe kell venni azt is, hogy a riportalanyok általában **formálisabb nyelvhasználatra** törekszenek, így szövegeik nem az informális, spontán beszélt nyelvet képviselik.

Ezért van az, hogy a „személyes” alkorpusz találatai sokszor sokkal inkább tükrözik a beszélt nyelvet, mint a „beszélt nyelvi” alkorpusz találatai.

c) A találatok alapján úgy tűnik, hogy a beszélgetések tematikailag nagyon egyoldalúak: politikai és közéleti jellegűek.

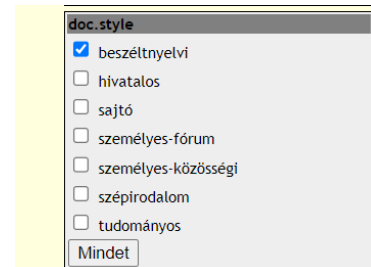
Ez súlyos torzulással járhat, mert az adatokból nemegyszer úgy tűnik, mintha a „beszélt nyelv” a hivatali nyelvhez állna legközelebb. Ehhez természetesen hozzájárul, hogy valószínűleg a felolvasott hírek is ilyen témájúak.

d) A vizsgálatok folyamán nagyon hasonló szövegeket is találtunk. A köztük lévő különbségekből arra következtettünk, hogy ugyanazt a szöveget két különböző személy írta át, és (tévedésből?) mindkettő belekerült az adatbázisba.

Az ilyen ismétlődés ugyanolyan elbírálás alá esik, mint az, amikor az ismétlődő szövegek teljesen azonos szövegezésűek, vagyis ezeket az elsődleges tisztítás során ki kell hagyni a statisztikából.

Nézzünk meg néhány példát a bizalmas stílusértékű *drukk* szóra a „beszélt nyelvi” alkorpuszból!

A *drukk* szótöpre kerestünk rá, a „beszélt nyelvi” alkorpusz kipipálásával, a találatokat egy mondatnyi kontextusban jelenítettük meg.



<b>Keresés</b> ?	korpusz: MNSZ2 találatok száma: 40
<b>Mentés</b>	<b>beszélt nyelvi</b> De lehet hogy pont az együttes <b>drukkunk</b> is hátrátatja őket.
<b>Megjelenítés</b> KWIC/mondat	<b>beszélt nyelvi</b> " Persze végig futott a fejemben, hogy hát ha épp ebbe a buszba szállunk be, akkor ez velünk történik, de annyira más állapotba kerülünk ilyenkor, hogy ennek a <b>druknak</b> köszönhetően, hogy elszalad a fejben és már másra koncentrálnunk utána.
<b>Rendezés</b> bal   jobb	<b>beszélt nyelvi</b> Mi most azért nem szorulunk, még valószínűleg a kedvesem sem, mert túl vagyunk ezen a <b>drukkon</b> , hogy sikerüljön a foglaló szerződés lejárta előtt, ezért szerintem még nem értünk rá ezzel foglalkozni. "
<b>találat</b> infó	<b>beszélt nyelvi</b> Meccs előtt volt egy picit <b>drukk</b> nyilvánvalóan az emberben, de az ember a parton egész mást érez, egész mást csinál, de én azt gondolom, hogy nem is lehet hasonló.
<b>keverés</b>	<b>beszélt nyelvi</b> Van bennünk izgalom és van bennünk egy kis <b>drukk</b> emiatt, hiszen ez egy nagyon nagy felelősség, hogy magát Pécsset is képviselni egy nagy felelősség általában, de azzal hogy ez tetőződik azzal hogy egy nemzetet egy államot és a nemzet alatt nemcsak a magyar határokon belüli magyarságot értem hanem az egész egyetemes magyarságot ha szabad így fogalmaznom.
<b>Véletlen minta</b>	<b>beszélt nyelvi</b> R.: Szilvi olyan, mint nagyjából az élsportolók, tehát hogy ilyenkor mutat egy olyan közepes formát és amikor jön a <b>drukk</b> és akkor jön a szereplés, akkor mindig jobb.
<b>Szűrés</b>	<b>beszélt nyelvi</b> Akkor még amatőr színjátszó voltam a híres-hírhedt Aréna színjátszó csoportnak Veszprémben a tagjaként, és hát iszonyú félelemmel és <b>drukkal</b> jöttem föl Pestre én a kis fiatal, huszoneves leányka és bementem a toalettire emlékszem, és ott is hát reszkettem.
<b>Gyakoriságok</b> szóalakok	<b>beszélt nyelvi</b> Hát azt még nem mondhatjuk, mert ugye még <b>drukk</b> is van bennünk, azért megpróbáljuk a legjobbat kihozni magunkból.
<b>szótövek</b> elemzések	<b>beszélt nyelvi</b> Sramó Gábor, a Bóbita Bábszínház vezetője.: - Számunkra nagyon fontos és nagyon izgalmas, hogy egyrészt lássuk egymás előadásait, a másik, ami mindig minden bábszínháznak egy misszió, hogy felnőttekkel megismertetni a bábművészetet és ennek a műfajnak a sokszínűségét, és ilyenkor kettős <b>drukk</b> és kíváncsiság munkál az emberben, hogy hogy sikerül a kollegáknak és hogy fogadja a közönség.
<b>dokumentumok</b>	<b>beszélt nyelvi</b> Az egész családom elég vallásos, katolikusok vagyunk, ez egy olyan dolog amikor évente többször is föl kell menni a ringbe, rengeteg ember előtt kell boksolni, és ez eléggé nagy stresszel jár, bár én szeretem, de olyankor izgul mégis az ember, tehát bármilyen mérkőzést is boksoltam, mindig ingulok egy kicsit előtte, tehát a <b>drukk</b> bennem van, és hát amikor fölmegek persze eszembe jut.
<b>Kollokációk</b> ?	<b>beszélt nyelvi</b> Önben is van egy kis <b>drukk</b> ?
	<b>beszélt nyelvi</b> Félni nem féltünk, tudtuk, hogy a mi csapatunk jobb, de azért ki-ki mérkőzésen mindenkiben van egy picit egészséges <b>drukk</b> , és szerintem, ez mindenkiben benne volt, ezt a feszültséget lehetett is érezni, egész nap a csapaton, de úgy látszik, hogy ez nagyon jól jött, hogy a csapat délután ilyen jó teljesítményt nyújtott. "
	<b>beszélt nyelvi</b> <hi rend="1"> Kalapos Mihály </hi> :- Mekkora a <b>drukk</b> most benned?
	<b>beszélt nyelvi</b> Van önben egy kis <b>drukk</b> ?
	<b>beszélt nyelvi</b> Róth Antal edző: - Persze én azt gondolom, h ogy mindenkiben kell, h ogy legyen <b>drukk</b> , hiszen az embert ez élteti, hogy ha csak megszokásból dolgozik az ember és rutinszerűen, akkor az kb. olyan, hogy nem igazából érdekli a feladat.
	<b>beszélt nyelvi</b> Persze érdekel és ez egyfajta <b>drukkot</b> nyilván jelent is az embernek, de ez olyan, mint a játékosnak a mérkőzés előtti lámpaláz.
	<b>beszélt nyelvi</b> Zubai Szabolcs: - Igen, ez egy jogos kérdésnek tűnik, de én úgy gondolom, hogy ezeket a mérkőzéseket mindenki annyira várja, hogy el se tudja képzelni azt, hogy netalán rosszul sülné el, netalán túlpörögné ezeket a mérkőzéseket, hanem egy olyan természetes és megfelelő <b>drukk</b> van, aminek viszont meg kell lenni, és úgy gondolom, hogy ez a csapat minden egyes tagján megvan, és ez szerintem nagyon jó dolog, hogy átérezzük ennek a súlyát és felelősségét.
	<b>beszélt nyelvi</b> Félni nem féltünk, tudtuk, hogy a mi csapatunk jobb, de azért ki-ki mérkőzésen mindenkiben van egy picit egészséges <b>drukk</b> , és szerintem, ez mindenkiben benne volt, ezt a feszültséget lehetett is érezni, egész nap a csapaton, de úgy látszik, hogy ez nagyon jól jött, hogy a csapat délután ilyen jó teljesítményt nyújtott. "
	<b>beszélt nyelvi</b> " Félni nem féltünk, tudtuk, hogy a mi csapatunk jobb, de azért ki-ki mérkőzésen mindenkiben van egy picit egészséges <b>drukk</b> , és szerintem, ez mindenkiben benne volt, ezt a feszültséget lehetett is érezni, egész nap a csapaton, de úgy látszik, hogy ez nagyon jól jött, hogy a csapat délután ilyen jó teljesítményt nyújtott. "
	<b>beszélt nyelvi</b> Az éles mérkőzés ugyanolyan minden kispadon, de ha valaki a Ferencváros kispadjáról várja az első éles mérkőzését biztos bennem van egy bizonyos egészséges <b>drukk</b> , de ez normális és nincs bennem semmi idegeskedés ami arra utalna hogy nem tudnám normálisan végezni a munkámat.
	<b>beszélt nyelvi</b> Hajdú B. István: - És volt benned egy kis <b>drukk</b> , hiszen név évvel ezelőtt az egész csapatnak az olimpia nem úgy sikerült, ahogy azt terveztük?

## ÷ F) A két „személyes” alkorpusz sajátosságai

A „személyes” alkorpusz az MNSZ2 legújabb verziójában (v2.0.5) két részből áll: a „**személyes-fórum**” alkorpusz internetes fórumok beszélgetéseit tartalmazza, a „**személyes-közösségi**” pedig a közösségi médiából származó szövegeket tartalmaz.

A két „személyes” alkorpusznak együttesen **301,1**, ill. **366,5 millió szövegszó** a terjedelme, vagyis – így egybevéve – a második legnagyobb alkorpuszt alkotnák, csak a „sajtónyelvi” előzi meg őket.

Ez a két alkorpusz **áll legközelebb a beszélt nyelvhez** (közelebb, mint az ún. „beszélt nyelvi” alkorpusz szövegei), de természetesen nem tekinthető beszélt nyelvnek: a beszélt nyelvihez legközelebb eső írott nyelvnek kell minősíteni. Ennek ellenére sokszor reális következtetéseket vonhatunk le az ebben az alkorpuszban előforduló szavak alapján a valódi beszélt nyelvre nézve.

### ÷ a) A „személyes-fórum” alkorpusz sajátosságai

A „személyes-fórum” alkorpusz mérete **57,9**, ill. **57,7 millió szövegszó**; a második adat „becsült” adat. Ez az egyedüli alkorpusz, melynek nagyobb a mérete a táblázat alatt közölt adat szerint, mint az „Alkorpuszok” menüpont alatti „alkorpuszok készítése” opció választásakor megjelenő adat szerint. Ez az összes alkorpusz közül a legkisebb terjedelmű.

**a)** Az adatok értékelésekor figyelembe kell venni, hogy az, hogy a fórumozó milyen nyelvváltozatban ír, nagyban függ egyéniségétől és társadalmi státuszától is, tehát az, hogy ő éppen fórumozik, korántsem határozza meg abszolút módon a stílusát. Nemritkán találunk a fórumokon egészen **választékos** vagy éppen **formális stílusú** hozzászólásokat is, melyeket szerzőik **esetleg előzőleg gondosan megfogalmaztak és megszerkesztettek**, s csak utána illesztették be a fórumba hozzászólásként.

**b)** Azt sem szabad elfelejtenünk, hogy a fórumozók **nemritkán idéznek külső forrásokból**, vagy átfogalmazva, vagy szó szerint (s nem is feltétlenül különítik el idézőjellel az idézetet), emiatt pedig **kisebb számban bármilyen stílusértékű szavak** előfordulhatnak ebben az alkorpuszban. A találatok tisztítása során az idézetekből származókat el kell távolítani.

Ehhez célszerű beállítani minél nagyobb szöveggörnyezet megjelenítését, hogy lehetőleg egyértelműen kiderüljön, idézetről van-e szó.

Az idézőjel nélküli egyenes idézetek az eredeti szerző stílusát tükrözik, s így nem feltétlenül illenek bele a „személyes” alkorpuszba.

**c)** Bár – amint jeleztük – ez az alkorpusz korántsem csupán bizalmas stílusú szövegeket tartalmaz, mégis talán ez alkalmas leginkább a bizalmas stílus, ill. a bizalmas stílusértékű nyelvi formák (vagy akár a tipikus helyesírási hibák) vizsgálatára.

d) Az adatok értékelésekor figyelembe kell venni, hogy a bizalmas stílussal nemegyszer együtt jár a **nyelvi játék**: a hozzászólók szándékosan, stíluselemként is „elkövethetnek” ún. nyelvhelyességi hibákat, vagy használhatnak tájnyelvi formákat, máskor készakarva írhatnak le hibásan bizonyos szavakat, kifejezéseket.

A szándékosságra utalhat az **irónia**, olykor az **érzelmi töltés** s persze az, ha a szándékosan „rosszul” használt nyelvi forma **idézőjelbe** van téve, **emotikon** áll mellette, vagy **metanyelvi megjegyzéssel** van ellátva.

A nyelvi játék további jele az, ha az érintett fórumozó felszólalásai amúgy rendezettek, nincs bennük stílustörés, kevés bennük az elírás, helyesírási hiba stb.

Előfordulhat a szövegekben **szerepjátszás** is, amikor a szerző a nyelvhasználatával valaki mást utánoz, karikíroz, gúnyol (ezzel olykor a sajtónyelvi, sőt a hivatali nyelvi alkorpuszban is találkozhatunk). Az adatok elemezésekor ezt a lehetőséget is szem előtt kell tartani.

Néhány példa a bizalmas stílusértékű **szívat** szóra a „személyes-fórum” alkorpuszból:

korpusz: MNSZ2 találatok száma: 200	
személyes-fórum	Bár lehet, hogy kényelmesebb lenne otthon ülni a hátsómon, és <b>szívatni</b> a kollégákat, hogy, hej, beh rosszul csináljátok, gyerekek!
személyes-fórum	És ezek <b>szívatnak</b> , te beindulsz, levágsz egy oldalas hegyibeszédet, amin lehet nagyokat szórakozni.
személyes-fórum	Én azt mondtam, hogy élvezem Hunszkörpió dumáját, mert szerintem <b>szívatja</b> a népet a góbédumákkal.
személyes-fórum	Nekem speciel az nem tetszett, ahogy velem érvelt anno, de az, hogy most ő is <b>szívat</b> titeket, annál inkább.
személyes-fórum	Én nem titkolom, hogy ki vagyok, aki akarja, megtudhatja ki van ultrabrowney mögött, álnevet azért használok, mert nem muszáj minden hülye bejelentkezzen álneven, és engem <b>szívasson</b> , csak azért, mert kiteszem a valódi nevemet.
személyes-fórum	Te Veridicus, te ugye csak <b>szívatod</b> itt az embert?
személyes-fórum	Bicskásnak most írtam magánlevélben, hogy kizárólag egy céllal jövök fel a Disputára: oly keveset <b>szívatallak</b> az utóbbi időben, hogy az már kezd menthetetlenül kínossá válni.
személyes-fórum	Hogy az a Kóbrával a Zalkából mit bírta egymást <b>szívatni</b> !
személyes-fórum	Vagy <b>szívatják</b> , vagy buta, vagy hazudik!
személyes-fórum	Már másodszer <b>szívassák</b> meg azt akinek szurkolok a franc beléjük!
személyes-fórum	ERIKÁVAL <b>SZÍVASSA</b> KI!
személyes-fórum	Legfeljebb a Torgyival fogják <b>szívatni</b> , pedig a Magyar Bálintot nem szívatják a Hornnal..
személyes-fórum	Legfeljebb a Torgyival fogják szívatni, pedig a Magyar Bálintot nem <b>szívatják</b> a Hornnal..
személyes-fórum	Nem olyanok tűnik a fiú, mint aki a hűjét jácra, vagy <b>szívatni</b> akar valakit.
személyes-fórum	Az indulási gond lehet a szívatótól is (ilyenkor már csak Téged <b>szívat</b> :)).
személyes-fórum	Vagy mit tud OV. csinálni? b) nem a miniszterelnök nem pontosan idézett, de ott és akkor jogos kijelentése miatt <b>szívatják</b> a magyarokat.
személyes-fórum	Nem tudnák az ottani magyarokat <b>szívatni</b> , ha Erdély beljebb lenne.
személyes-fórum	Mi nem <b>szívatnánk</b> a románokat, akár mit mondana a miniszterelnökük.
személyes-fórum	Olyan jó volt, amikor mindenem rendben volt és a bunkó rendőr (mert az!), semmivel nem tud <b>szívatni</b> és ez szívás (!) neki.
személyes-fórum	Mit <b>szívat</b> sz itt engem, rosszéletű?
személyes-fórum	- Szeretnénk a többgyerekesek esetében növelni az egy főre jutó nettó jövedelmet, ezáltal ösztönöznünk a gyermekvállalást, de nem akarjuk nagyon <b>szívatni</b> az átlagpolgárt.
személyes-fórum	MuszáLY neked rögtön az elején <b>szívatnod</b> szegényt?
személyes-fórum	Én biztosan nem költöttem volna 40 milla USD-t a megtakarított pénzből, hogy <b>szívassam</b> az elnököm.
személyes-fórum	Általában ott vannak a trafipaxok, ahol teljesen irreálisan alacsony a sebességkorlátozás és direkt lehet <b>szívatni</b> a népet.
személyes-fórum	(Hogy minket egy tizennyóc éves vidéki paraszt (nem földművelő) próbált <b>szívatni</b> (má' kapásból zavarta, hogy mi " pöstiek " vagyunk, ez valami fixáció arrafele), el se tudom mesélni milyen sötét vót szerencsétlen, leszarom, vigyorogtam rá tudod, nekem nem ott kellett megtanulnom azt, hogy a konyhámon kívül hogy kell beszélni a bácsikkal; szerinted ez kellett az " élethez "?)
személyes-fórum	Ha tudtak, utána jöjjenek rá ki <b>szívat</b> ta meg őket...
személyes-fórum	Azt kérdezed, miért magyar lakosok nagyrészt, azok akik <b>szívatják</b> , ezt a vén eszementet?
személyes-fórum	A rendszámot nem veszik le, csak <b>szívatják</b> a faszit, de azt nagyon.
személyes-fórum	Szóval ezt a jogalkotó kifejezetten azért rakta a törvénybe, hogy <b>szívasson</b> .
személyes-fórum	Mindig szívesen hagytam őket magukra kettesben a tükörrel, és röhögtem rajtuk ezerral, amikor - tisztán sportból - nekiláttak volna feltekerni az ujjaik közé, viszont igen soxor felbosszantottak, amikor óvatlanabb haverjaimat <b>szívat</b> ták.
személyes-fórum	Sokat <b>szívat</b> tak - szerencsére - és ettől sokat keményedtem.
személyes-fórum	Ha gyermekkoromban valahol <b>szívat</b> tak - pl. edzőtáborban - arról mindig szívesen meséltem mint kellemes élményről, mindenesetre a következő évben újra mentem fekvőtámaszozni, lepedővel csapkodni, cipőt fogpasztával kibélelni, mert élveztem amikor az edzőt tudtuk heccelni és ő elismerése jeléül 50 fekvőtámaszt kiosztott.
személyes-fórum	Mégpedig azért szervokormány, mert mindenkit csak <b>szívat</b> ... csak szívatóval indul be szegény.... nem hiába, ez itt már a polgártél...
személyes-fórum	Mégpedig azért szervokormány, mert mindenkit csak <b>szívat</b> ... csak szívatóval indul be szegény.... nem hiába, ez itt már a polgártél...
személyes-fórum	Úgy látszik, még nem elég ír utast <b>szívat</b> ott a cég, van mit pótolni....

## ÷ b) A „személyes-közösségi” alkorpusz sajátosságai

a) A „személyes-közösségi” alkorpusz mérete **243,2**, ill. **308,8** millió szövegszó. Különálló alkorpusznak tekintve a számadatok szerint ez a második legnagyobb alkorpusz a „sajtó-nyelvi” után, s talán a valóságban is, annak ellenére, hogy nagyon nagy számban tartalmaz teljesen azonos tartalmú és szövegű dokumentumokat.

b) A közösségi média szövegeiben még **vegyesebb anyagot** találunk, mint a „személyes-fórum” alkorpuszban: a posztolt szövegek a legkülönbélebb forrásokból származhatnak.

Épp ezért talán ez az az alkorpusz, melynek adatait a legnehezebb elemezni, értékelni. Ehhez mindig figyelembe kellene venni, hogy az adott oldalon közzétett eredeti szövegről van-e szó, vagy pedig máshonnan származó, megosztott dokuról, dokurészletről.

c) Ennél is jóval nagyobb a már jelzett probléma, hogy ti. az MNSZ2 a megosztásokat is tartalmazza, így számos esetben ugyanaz a szöveg több tucatszor, de akár néhány százszor is megjelenik a találatok közt.

**Mivel ezek az ismétlődések súlyosan torzítják a kapott eredményeket, ezt az alkorpuszt nem fogjuk használni olyankor, amikor az adatokat statisztikailag dolgozzuk fel, viszont ha csak példákat keresünk, vagy a nyelvi formák stílusértékét csupán a szöveggörnyezetből kívánjuk megállapítani, ezt az alkorpuszt is bevonhatjuk a vizsgálatba.**

Néhány találat ugyanarra a *szívat* szóra a „személyes-közösségi” alkorpuszból:

korpusz: MNSZ2	
találatok száma: 1622	
oldalszám: 1 / 2	Ugrás <a href="#">következő</a>   <a href="#">utolsó</a>
személyes-közösségi	(Egyszer haverom <b>szívatam</b> be, hogy szilveszter péntek 13 lesz... és miért is ne, bevette..!
személyes-közösségi	Azért hogy mindenkinek össze jön minden oké elfogadom örülök neki, de mikor engem <b>szívat</b> az isten akkor már kezdek morcos lenni.
személyes-közösségi	Azért hogy mindenkinek össze jön minden oké elfogadom örülök neki, de mikor engem <b>szívat</b> az isten akkor már kezdek morcos lenni.
személyes-közösségi	(Egyszer haverom <b>szívatam</b> be, hogy szilveszter péntek 13 lesz... és miért is ne, bevette..!
személyes-közösségi	Meg fogom <b>szívatni</b> a drágát!!!!!! ugy tervezem, hogy jól elszigetelve a földtől, kikötök 1 gyöngyit a lábán villanypásztor vezetékevel Ez lesz a meglepi.!!!!
személyes-közösségi	(64. nap) Vajszínú (65. nap) Lövészet (66. nap) Végre otthon (67. nap) Otthon (68. nap) Otthon még ma is (69. nap) Elfagyott a fülem (70. nap) Hómérő (71. nap) Hadd legyek katona (72. nap) Mondja, miért <b>szívatják</b> magát annyira?
személyes-közösségi	Ki ez a ember és meddig akarja a dohányosokat és a nemzetet <b>szívatni</b> ??
személyes-közösségi	Ki ez a ember és meddig akarja a dohányosokat és a nemzetet <b>szívatni</b> ??
személyes-közösségi	Azért hogy mindenkinek össze jön minden oké elfogadom örülök neki, de mikor engem <b>szívat</b> az isten akkor már kezdek morcos lenni.
személyes-közösségi	Ki ez a ember és meddig akarja a dohányosokat és a nemzetet <b>szívatni</b> ??
személyes-közösségi	Ki ez a ember és meddig akarja a dohányosokat és a nemzetet <b>szívatni</b> ??
személyes-közösségi	Ki ez a ember és meddig akarja a dohányosokat és a nemzetet <b>szívatni</b> ??
személyes-közösségi	Ki ez a ember és meddig akarja a dohányosokat és a nemzetet <b>szívatni</b> ??
személyes-közösségi	Miután agyra <b>szívatott</b> az atlanticom töltése, és elolvastam pár fórum hsz-t a típusról úgy éreztem, hogy ez mellé nyúlás volt.
személyes-közösségi	És ami a legviccesebb a legjobban azt <b>szívatják</b> vele, aki pl. idézni akar a szövegből, vagy azért kijelölni, hogy rákeressen egy szóra, esetleg lefordítsa, ha idegennyelvű.
személyes-közösségi	Ribery a Bayern edzésén <b>szívatva</b> a stábot, Djokovics előnyben, Cavani közel a PSG-hez.
személyes-közösségi	Ki ez a ember és meddig akarja a dohányosokat és a nemzetet <b>szívatni</b> ??
személyes-közösségi	Remélem a többi égbőlkapott szerződésekkkel <b>szívató</b> bandát is hasonló elbánásban részesítik.
személyes-közösségi	Ribery a Bayern edzésén <b>szívatva</b> a stábot, Djokovics előnyben, Cavani közel a PSG-hez.
személyes-közösségi	- Dehogy, csak <b>szívatlak</b> .
személyes-közösségi	Ribery a Bayern edzésén <b>szívatva</b> a stábot, Djokovics előnyben, Cavani közel a PSG-hez.
személyes-közösségi	-Hayley felhívott, hogy... csak <b>szívatatok</b> , ugye?
személyes-közösségi	Ki ez a ember és meddig akarja a dohányosokat és a nemzetet <b>szívatni</b> ??
személyes-közösségi	Mikor a mai szakaszon szanaszéjjel <b>szívatták</b> a sprintercsapatokat?
személyes-közösségi	Ribery a Bayern edzésén <b>szívatva</b> a stábot, Djokovics előnyben, Cavani közel a PSG-hez.
személyes-közösségi	Aggódjunk egymásért, <b>szíva</b> ssuk egymást, ellenőrizzük egymást.
személyes-közösségi	Vagy <b>szívatjuk</b> egymást.
személyes-közösségi	Meg a zenekart segítőkét <b>szívatjuk</b> közösen.
személyes-közösségi	Aggódjunk egymásért, <b>szíva</b> ssuk egymást, ellenőrizzük egymást.
személyes-közösségi	Úgy hallottam, hogy te nagy játékos vagy... szóval játsszunk! beszélgezzünk, veszekedjünk, találkozzunk. írjunk egymásnak jó reggelt és jó éjszakát üzeneteket. sétáljunk együtt, varázsolj el. aggódjunk egymásért, <b>szíva</b> ssuk egymást, ellenőrizzük is egymást. adjunk egymásnak beceneveket, emlékeket. mesél el a titkaidat, bizz bennem falak nélkül. beszéljük át az egész éjszakát. tegyük féltékennyé a másikat, csokolozzunk és ölelkezzünk, mintha többről szólna.és, hogy ki veszti?! aki először lesz szerelmes!.

# & RÉGIÓ SZERINTI MENNYISÉGI VIZSGÁLAT

## ÷ 1. Általános kérdések

Amint azt már fentebb említettük, a régió szerinti vizsgálat olyankor hasznos, amikor a magyarországi és a külső régiókbeli magyar nyelvváltozatok közti különbségeket szeretnénk feltárni.

E különbségek közt vannak abszolútak és relatívak. Az **abszolút különbségek** abban állnak, hogy egy bizonyos nyelvi forma (csaknem) kizárólag valamely régió(k)ban használatos, míg más régió(k)ban egy másik forma használatos (szinte) kizárólagosan; egyik régióban sincs (számottevő) váltakozás a két forma közt.

A zárójeles kiegészítésekre („csaknem”, „számottevő”) azért van szükség, mert kisebb mértékű váltakozással mindig számolhatunk. Nemcsak azért, mert a külső régióbeli magyaroknak legalább egy része ismeri a saját régiójukban használt nyelvi formák közmagyar megfelelőjét, különösen azok, akik járnak Magyarországra, hanem azért is, mert Magyarországon nem teljesen elhanyagolható azoknak az embereknek a száma, akik más régióból költöztek át, s részben megőrizték korábbi nyelvhasználatukat.

**Relatív különbségekről** akkor beszélünk, amikor a régiók közti nyelvi különbségek gyakoriságiak: vagyis mindkét nyelvi forma minden vizsgált régióban használatos, csak eltérő gyakorisággal. Természetesen ugyanez érvényes akkor is, ha kettőnél több nyelvi forma használatáról van szó.

A régióközi összehasonlítást jelentősen nehezíti, hogy a négy külső régió – Erdélynek, a Felvidéknek, a Vajdaságnak, a Kárpátaljának – a nyelvi anyaga **aránytalanul csekély terjedelmű** a mai magyarországgal összehasonlítva.

Az alkorpuszok méretére nézve fentebb közöltünk egy táblázatot, mely szerint a „magyarországi” alkorpusz 1013,9 millió szövegszót tartalmaz, az „erdélyi” 3,9 milliót, a „szlovákiai” 17,3 milliót, a „vajdasági” 2,0 milliót, a „kárpátalji” pedig 2,5 milliót. (Az adatok az MNSZ2 oldalán található táblázatból valók: <https://clara.nytud.hu/mnsz2-dev/stat.html>).

Ezekről az adatokról jelentősen eltérnek azok, melyek akkor válnak láthatóvá, amikor az „Alkorpuszok” alatt az „új alkorpusz készítése” menüpontot választjuk. Ezek szerint az adatok szerint a „magyarországi” alkorpusz 1222,6 millió szövegszót tartalmaz, az „erdélyi” 4,8 milliót, a „szlovákiai” 20,7 milliót, a „vajdasági” 2,5 milliót, a „kárpátalji” 3,0 milliót.

Az eltérések okáról az MNSZ2 készítői nem adnak felvilágosítást, csak feltételezhetjük, hogy ez utóbbi adatok a korpusz bővítésének eredményeként álltak elő, a táblázat pedig egy régebbi állapotot tükröz. Kérdés, mennyire korrekt olyan alapvető kérdésben bizonytalanságban hagyni a felhasználót, mint amilyen az alkorpuszok és ezáltal az egész korpusz mérete.

**A külső régiókon belül is van aránytalanság**, akár a táblázat adatait vesszük mérvadónak, akár az „új alkorpusz készítése” menüpont alatti adatokat. A „szlovákiai” alkorpusz mindkét esetben többszörösen nagyobb mind az „erdélyinél”, mind a „vajdaságinál”, mind a



„kárpátaljinál”, sőt a „szlovákiai” alkorpusz több mint kétszer annyi szövegszót tartalmaz, mint az „erdélyi”, a „vajdasági” és a „kárpátalji” alkorpusz együttvéve.

A táblázatos adatok szerint a „szlovákiai” 17,3 milliójával a többi külső régióbeli 8,4 milliója áll szemben; az „új alkorpusz készítése” menüpont alatti adatok szerint a „szlovákiai” 20,7 milliójával a többi külső régió 10,3 milliója áll szemben.

A könnyebb áttekinthetőség érdekében a kétféle adatsort táblázatba foglaltuk. Az oszlopfejlécben a „táblázat” oszlopcím az MNSZ2 munkatársai által készített és az MNSZ2 honlapján az „Alkorpuszok mérete” menüpont alatti adatokra utal, az „új alkorpusz” oszlopcím pedig azokra az adatokra, melyek akkor válnak láthatóvá, amikor új alkorpuszt szeretnénk készíteni.

	régió	az alkorpuszok mérete	
		táblázat	„új alkorpusz”
1.	magyarországi	1013,9	1222,6
2.	erdélyi	3,9	4,8
3.	szlovákiai	17,3	20,7
4.	vajdasági	2,0	2,5
5.	kárpátalji	2,5	3,0
Σ	<b>MNSZ2</b>	<b>1039,7</b>	<b>1253,6</b>

Ezeknek az aránytalanságoknak a torzító hatását enyhíthetjük a relatív gyakoriság számolásával, de ez nem segít olyankor, amikor a külső régiókból kapott találatok száma nagyon alacsony. Ilyenkor ugyanis zavaróan nagy a véletlen szerepe (pl. éppen milyen témájú újságcikk, poszt, parlamenti felszólalás stb. került be az érintett külső régióból az adatbázisba a gyűjtés időszakában). Ez a tényező különösen akkor torzítja az eredményeket, amikor nem közhasználatú lexémákat vizsgálunk.

## ÷ 2. A nyílt nap vs. nyitott nap példája

Az előrebocsátottak nem jelentik azt, hogy sohasem volna érdemes összehasonlítani a magyarországi és a külső régiókbeli számadatokat, csupán arról van szó, hogy az összehasonlítást kellő körültekintéssel kell végezni. A következőkben egy ilyen vizsgálatot mutatunk be, amely egy való életből vett példán alapul.

A magyar tanszék egyik egykori hallgatója, aki a rádiónál dolgozik, felhívta a figyelmünket a következő Facebook-bejegyzésre:

Nemcsak a politika és a közélet prominens képviselőitől, de sajnos újságíróktól, tanítóktól, sőt iskolaigazgatóktól is gyakorta hallani/látni a „nyitott nap” kifejezést. Azért is veszélyes és szomorú ez a felvidéki magyar „betegség”, mert amíg a „tyepláki”, a „horcsica” és társaik általában csak a rétegnyelvben fordulnak elő, addig a „nyitott nappal” a köznyelvbe helyezik ezt a szlovakizmust. Pedig van magyar köznyelvi megfelelője, sőt, alig tér el tőle: „nyílt nap”. Használjuk bátran!

Az első feladatunk az volt, hogy korpuszadatokkal igazoljuk (vagy cáfoljuk), hogy a *nyílt nap* köznyelvi, a *nyitott nap* forma pedig felvidéki magyar. A teljesség kedvéért megnéztük a *nyitott ajtó napja* formára vonatkozó adatokat is, bár ezt a változatot a Facebook-bejegyzés szerzője nem említi.

A „részletes keresés” opciót választottuk: a *nyílt/nyitott/nyitott* szóalakra kerestünk rá, kontextusként pedig a *nap/nap/ajtó*, *nap* szótöveket adtuk meg. A keresett szótól jobbra, közbeékelődő szó nélküli találatokat kértük, mivel feltételeztük, hogy a szavak közé a szerkezetben nem szoktak más szavak bekerülni.

A „szótövek” sorban a „mind forduljon elő” opciót választjuk, mivel a vizsgált kifejezéseknek minden szó kötelező eleme.

Az alábbi képen látjuk a *nyílt nap* beállításait:

korpusz:	MNSZ2
lekérdezés típusa:	részletes keresés
részletes keresés:	szóalak <input type="checkbox"/> nyílt <input type="checkbox"/> szófaj: TETSZ
	<input type="checkbox"/> összetett szó   morf: morféma: TETSZ
<b>Kontextus</b>	
<b>Szűrés szótövekre</b>	
szűrési ablak:	jobbra 1 szó.
szótövek:	nap mind forduljon elő.

Itt a *nyitott nap* beállításai következnek:

korpusz:	MNSZ2
lekérdezés típusa:	részletes keresés
részletes keresés:	szóalak <input type="checkbox"/> nyitott <input type="checkbox"/> szófaj: TETSZ
	<input type="checkbox"/> összetett szó   morf: morféma: TETSZ
<b>Kontextus</b>	
<b>Szűrés szótövekre</b>	
szűrési ablak:	jobbra 1 szó.
szótövek:	nap mind forduljon elő.

Végül nézzük meg a *nyitott ajtó napja* beállításait is:

korpusz:	MNSZ2
lekérdezés típusa:	részletes keresés
részletes keresés:	szóalak <input type="checkbox"/> nyitott <input type="checkbox"/> szófaj: TETSZ
	<input type="checkbox"/> összetett szó   morf: morféma: TETSZ
<b>Kontextus</b>	
<b>Szűrés szótövekre</b>	
szűrési ablak:	jobbra 2 szó.
szótövek:	ajtó nap mind forduljon elő.

A vizsgálatban a hivatalosan közölt táblázatból vett adatokat használtuk. A keresési eredményeket az alábbi táblázat tartalmazza, régiók szerinti bontásban:

	<i>régió</i>	<i>alkorpuszok mérete</i>	<i>nyílt nap</i>	<i>nyitott nap</i>	<i>nyitott ajtók napja</i>
1.	magyarországi	1013,9	2723	76	1
2.	erdélyi	3,9	2	1	0
3.	szlovákiai	17,3	76	33	1
4.	vajdasági	2,0	2	0	0
5.	kárpátalji	2,5	2	0	0
Σ	<b>MNSZ2</b>	<b>1039,7</b>	<b>2805</b>	<b>110</b>	<b>2</b>

A táblázat jól példázza a fentebb említett tényt, hogy ti. az adatokból, amikor a számuk alacsony, nem vonhatunk le megbízható következtetéseket. Látjuk, hogy Erdélyből, a Vajdaságból és a Kárpátaljáról alig néhány adatunk van a vizsgált formákra. Ez annak lehet a következménye, hogy ezeknek az alkorpuszoknak az anyaga nagyon szűk, s véletlenül nem kerültek be olyan íráskorok, melyekben említve volna a nyílt nap. Kevésbé valószínű, hogy e régiókban maga a fogalom volna ismeretlen, de az is, hogy teljesen más volna a neve.

Szemben a statisztikai vizsgálatokra alkalmatlan erdélyi, vajdasági és kárpátalji adatokkal, a szlovákiai magyar adatok száma nem annyira alacsony, hogy ne volna érdemes komolyan venni őket, és statisztikai számításokat végezni velük. A továbbiakban tehát csak a magyarországi és a szlovákiai magyar adatokkal foglalkozunk, ezeket hasonlítjuk össze egymással. Ezeket egy virtuális „magyarországi + szlovákiai” korpusz részének tekintjük (erre alább mint HUSZ-korpuszra utalunk).

	<i>régió</i>	<i>alkorpuszok mérete</i>	<i>nyílt nap</i>	<i>nyitott nap</i>	<i>nyitott ajtók napja</i>
1.	magyarországi	1013,9	2723	76	1
3.	szlovákiai	17,3	76	33	1
Σ	mo-i + szlki (HUSZ)	<b>1031,2</b>	<b>2799</b>	<b>109</b>	<b>2</b>

A táblázatra ránézve azonnal látjuk, hogy a vizsgált változatok közül sem a *nyílt nap*, sem a *nyitott nap* nem kötődik közvetlenül az egyik vagy a másik régióhoz.

A *nyitott ajtók napjára* kapott magyarországi találat tükörfordítás a németből, amint az a szövegkörnyezetből ki is derül:

Németországban működik egy olyan dolog, hogy „Tag der offenen Türe”, azaz a nyitott ajtók napja. Évente kb. 4-5 alkalommal rendezik meg, ilyenkor a megválasztott képviselőknek lehetőségük van (és élnek is vele) a választóikkal beszélgetni, egyeztetni, lemérni a támogatottságukat. Ott bizony elmagyarázzák a Stanci néninek, ha kell, többször is, miért van szükség bizonyos döntések meghozatalára. Olyan fórumféle az egész, messze nem hasonlít az itthoni gyakorlatban megszokott képviselői fogadóórákhoz. (magyarországi)

A találat hiteles, a beszélő nyilván magyarországi, de németországi reáliát ismertette használja a (nyilván alkalmi) tükörfordítást.

A szlovákiai példa újságcikkből való:

Robert Fico, aki a gyermeknap alkalmából megtartott nyitott ajtók napján, vasárnap látogatókat kalauzolt a pozsonyi kormány épületében, nagyobb választási részvételre ösztönözte a szlovákokat. „Vagy olyan képviselőink lesznek ott (Európai Parlamentben), akik Magyarországot és a Kárpát-

medencét fogják képviselni, vagy Szlovákiát és a szociális programokat. Néhányan azt mondják, hogy azok az emberek, akiket a Szlovák Köztársaságban választanak meg EP-képviselőnek, állítólag a Kárpát-medencét fogják képviselni. Hát már teljesen megbolondultunk?” – jelentette ki az emberek előtt Fico.

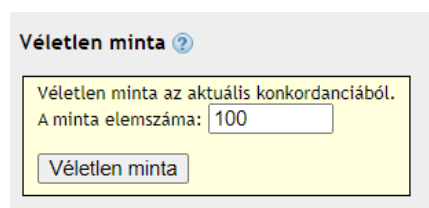
A szlovákiai magyar példa ugyan hiteles, de mivel csak egyetlen adatról van szó, nem vonhatunk le belőle megbízható következtetést. Az más kérdés, hogy a Google keresőjének segítségével további adatokat kaphatunk, ezeknek elemzése azonban nem témája ennek a jegyzetnek.

A továbbiakban ezért csak a *nyílt nap* és a *nyitott nap* változatokkal foglalkozunk.

A fenti táblázatból látjuk, hogy a *nyílt nap* formára a „magyarországi” alkorpusz 2723 találatot adott, míg a *nyitott nap* formára csak 76-ot; a „szlovákiai” alkorpuszban a *nyílt nap* forma 76, a *nyitott nap* forma pedig 33 találatot adott. Mivel a négy szám közül három alacsony, az adatokat tisztítani kell, hiszen nem mindegy, hogy mennyi az irreleváns találat.

Az elsődleges tisztítás során eltávolítjuk a korpusz gondatlan összeállítása miatti azonos találatok közül az ismétlődőket.

Ezt könnyű megtenni az alacsony számadatok esetében, mert 76, ill. 33 adatot gyorsan át lehet áttekinteni. A 2723 találat átnézése viszont túl nagy munka volna, ezért itt egy 100 találatot tartalmazó véletlen mintán végezzük el a tisztítást, és a releváns vs. irreleváns adatok arányát kivetítjük a 2723 találatra.



A 100 találatot a találati szó előtti szó szerint rendezzük, hogy az ismétlődő találatokat könnyű legyen felismerni. A 100 találatból 6 az ismétlődő, vagyis a találatok 6%-a. A 2723-nak a 6%-a 163,38; úgy tekinthetjük tehát, hogy a 2723 talált közül nagyjából 163 az irreleváns. Ezt levonva a 2723-ból 2560 találatot kapunk, vagyis ismétlődő találatok nélkül a *nyílt nap* forma előfordulása a magyarországi alkorpuszban 2560 körülre tehető, ami tehát becslés – nem valóságos – adat.

Gondosabb vizsgálat során átolvasnánk mind a száz találatot, hogy van-e köztük irreleváns, és ezek számát levonnánk a 2560-ból; ez volna a másodlagos tisztítás. (A tisztításról alább még bővebben is lesz szó.)

A másodlagos (és harmadlagos) tisztítástól azonban itt most eltekintünk, mert a tisztítással alább részletesebben is foglalkozunk, s a találatok felületes átnézésből úgy ítéltük meg, hogy a másodlagos tisztítás után csak nagyon csekély mértékben változna az adatok száma, ami elhanyagolható hatással volna az eredményekre.

A többi adatból nem kellett véletlen mintát készíteni, mivel számuk – ahogy említettük – alacsony. Az ismétlődő találatok eltávolítása után a következő adatokat kapjuk:

	<i>régió</i>	<i>alkorpuszok mérete</i>	<i>nyílt nap (tisztított)</i>	<i>nyitott nap (tisztított)</i>
1.	magyarországi	1013,9	2560	58
3.	szlovákiai	17,3	75	33
Σ	mo-i + szlklai (HUSZ)	<b>1031,2</b>	<b>2635</b>	<b>91</b>

A további számításokat ezekkel a számokkal végezzük, még ha ezek nem pontosak is. Például a *nyitott nap* magyarországi találatai közül 2 a szlovákiai magyar Új Szó napilapból való: ezeket – ha végeznénk másodlagos tisztítást – ugyancsak le kellene vonni a „magyarországi” alkorpusz adataiból.

Az Új Szóból származó némely adatok magyarországiént való annotálásának az az oka, hogy amikor még nem léteztek a külső régiók alkorpuszai, a külső régióknak a korpuszba bekerült (gyér) anyagát nem különítették el a magyarországitól (erről fentebb az alkorpuszok jellemzőségeiről már volt szó). Miután létrejöttek a külső régiók alkorpuszai, az MNSZ2 munkatársai ahelyett, hogy a korábban egységes korpuszban található külső régióbeli anyagot kiszűrték volna, és megváltoztatták volna az annotációját úgy, hogy utaljon az érintett külső régióbeli alkorpuszra, az „egyszerűség kedvéért” az egész addigi anyagot magyarországinak minősítették, szőröstül-bőröstül, vagyis a külső régiók íróinak műveivel és a külső régiók sajtótermékeivel együtt.

Más irreleváns találatok is akadhatnak; ilyen például egy Németh László-részlet, melyben a *nyitott nap* szókapcsolat egészen más jelentést hordoz:

Amíg az ember a természet, a szenvedélyei balekja, a sok holnaptól nem látja meg a napot. Megcsokolom, összeházasodom vele, gyerekeink lesznek, fölneveljük – rugdossa előre az egyik ösztön. Nagy ember leszek, nyomot hagyok a világban, be kell a piramisokat fejeztem – süketíti, vakítja el a még erősebb. Csak amikor ezek a sarkallások elégették benne a magukét, a holnap művei s gyermekei könyörtelen arccal tekintenek vissza rá, akkor látja meg az idő igazi szerkezetét, a haláltól mindenfelől körülvevett, de a magasság felé mindig nyitott napot. Ha egy kőfaragó követ farag, annak a kőszobor a célja – a szerzetesé a faragás. De a faragás sem mint egyfajta munka, hanem mint egyfajta áhítat. Műremek – annak minden ízét átjárva; egy festékrög, egy kődudor sincs benne, ami épp annyira ne volna eszme, mint anyag. Így kell a szerzetes napjának minden ízét, minden tevékenységét valamiféle pietásnak átjárni. A dolgokat inkább megtaláljuk, mint feltaláljuk. A művészet is kimondás inkább, mint kitalálás. (Németh László)

Az ilyen típusú irreleváns adatok természetesen nem az összeállító hibájából kerültek bele a korpuszba; ilyenekkel mindig számolnunk kell. Ha nagyon fontos a pontos eredmény, a másodlagos tisztítás során ezeket kiiktathatjuk, s ennek megfelelően módosíthatjuk a számadatokat (l. alább, *A stílusrétegek szerinti minőségi vizsgálat* című fejezetben).

### **Az adatok összehasonlítása**

Látjuk, hogy a *nyitott nap* forma több mint kétszer gyakoribb a „magyarországi” alkorpuszban, mint a „szlovákiai”. Ebből azonban természetesen szarvashiba volna azt a következtetést levonni, hogy a *nyitott nap* kétszer gyakoribb használatú a magyarországi magyarok, mint a szlovákiai magyarok körében, mivel a magyarországi alkorpusz összehasonlíthatatlanul nagyobb a szlovákiaiainál.

Ahogy az alábbi táblázatból kiderül, a HUSZ-korpusznak mindössze 1,68%-a tartalmaz szlovákiai magyar adatokat, a korpusz 98,32%-a magyarországi adatokból áll. Ez azt jelenti, hogy ha a *nyílt nap* egyformán gyakori volna a két alkorpuszban, akkor a találatoknak 1,68%-a volna szlovákiai és 98,32%-a magyarországi. Ezzel szemben a táblázat „nyílt nap” oszlopából látjuk, hogy a *nyílt nap* formára vonatkozó 2635 adatnak 2,85%-a szlovákiai magyar és 97,15%-a magyarországi magyar, a *nyitott nap* formára vonatkozó 91 adatnak viszont 36,26%-a szlovákiai magyar, és csak 63,74 százaléka magyarországi magyar:

	<i>régió</i>	<i>alkorpuszok mérete</i>		<i>nyílt nap</i>		<i>nyitott nap</i>	
1.	magyarországi	1013,9	<b>98,32</b>	2560	97,15	58	63,74
3.	szlovákiai	17,3	<b>1,68</b>	75	2,85	33	36,26
Σ	mo-i + szlki (HUSZ)	<b>1031,2</b>	<b>100%</b>	<b>2635</b>	<b>100%</b>	<b>91</b>	<b>100%</b>

Felületesen nézve úgy tűnhetne, hogy a Facebook-író szerint köznyelvinek tekintett *nyílt* forma nemhogy nem ritkább Szlovákiában, mint Magyarországon, hanem még gyakoribb is (több mint kétszer olyan gyakori). Ez azonban csak látszat, melynek nyilvánvalóan az az oka, hogy a sors úgy hozta, hogy maga a fogalom gyakrabban kerül elő a szlovákiai, mint a magyarországi alkorpuszban.

Ezt a feltevésünket számszerűen úgy tudjuk bizonyítani, hogy összeadjuk a két megnevezésre utaló adatokat (2560 + 58, ill. 75 + 33), és megnézzük, hány százalékuk magyarországi, hány százalékuk pedig szlovákiai magyar. Ha egyforma gyakori volna a fogalom a két alkorpuszban, akkor az előfordulások 1,68%-ának kellene szlovákiai magyarnak, 98,32%-ának pedig magyarországi magyarnak lennie. Ezzel szemben a helyzet a következő:

	<i>régió</i>	<i>alkorpuszok mérete</i>		<i>nyílt nap + nyitott nap</i>	
1.	magyarországi	1013,9	<b>98,32</b>	2618	96,4
3.	szlovákiai	17,3	<b>1,68</b>	108	3,96
Σ	mo-i + szlki (HUSZ)	<b>1031,2</b>	<b>100%</b>	<b>2726</b>	<b>100%</b>

Látjuk, hogy a két szó együttesére relatíve sokkal több a szlovákiai magyar adat, mint a magyarországi, hiszen míg a HUSZ-korpusznak csak 1,68%-a szlovákiai magyar, addig a *nyílt nap* + *nyitott nap* formákra vonatkozó adatoknak, 3,96%-a szlovákiai magyar, vagyis csaknem két és félszer több, mint amennyi lenne akkor, ha a fogalom egyformán gyakran fordulna elő a két alkorpuszban.

Ez a példa jól mutatja, hogy a nyelvi gyakorisági adatok vizsgálatakor mennyire fontos odafigyelni arra is, hogy maga a *fogalom* milyen gyakori az egyik vagy a másik alkorpuszban.

Ami a *nyitott nap* formára vonatkozó adatokat illeti, ezekből bátran levonhatjuk azt a következtetést, hogy a szlovákiai magyar nyelvhasználatra jellemzőbbek, mint a magyarországiakra, hiszen a *nyitott nap* formára kapott találatoknak 36,26%-a szlovákiai magyar, vagyis csaknem tízszer több, mint amilyennek lennie kellene, ha a két forma egyforma gyakori volna a két régióban (a két szóra együttesen vonatkozó szlovákiai magyar adatok a HUSZ-korpusz adatainak csupán 3,96%-át alkotják).

A korpuszadatok tehát igazolják, hogy a *nyitott nap* a szlovákiai magyar nyelvhasználatra jellemző, ugyanakkor nem teljesen ismeretlen Magyarországon sem.

Végezzük még el a két szóra vonatkozó adatok összehasonlítását más módszerrel is, mégpedig úgy, hogy megnézzük, hogy milyen a két formára vonatkozó találatok aránya az egyik és a másik régióban!

	<i>régió</i>	<i>nyílt nap</i>		<i>nyitott nap</i>		<i>nyílt nap + nyitott nap</i>	
1.	magyarországi	2560	97,78	58	2,22	2618	100%
3.	szlovákiai	75	69,44	33	30,56	108	100%
Σ	mo-i + szlkiai (HUSZ)	<b>2635</b>	<b>96,66</b>	<b>91</b>	<b>3,34</b>	<b>2726</b>	<b>100%</b>

E módszer egyik előnye, hogy kiküszöböljük vele a fogalomnak a két alkorpuszban való eltérő gyakoriságából adódó problémákat.

Az összehasonlításból kiderül, hogy a HUSZ egészében a *nyílt nap* forma előfordulási aránya 96,66%, míg a *nyitott nap* 3,34%. Ha a két forma egyformán gyakori volna a két régióban, akkor nagyjából ilyen arányt kellene látni az alkorpuszokban is.

Ehelyett azt látjuk, hogy a „magyarországi” alkorpuszban túlreprezentált a *nyílt nap* (97,78%, miközben ez az arány a HUSZ egészében csak 96,66) és alulreprezentált a *nyitott nap* (2,22%, miközben a HUSZ egészében ez az arány 3,34%). Ezzel szemben a „szlovákiai” alkorpuszban a *nyílt nap* erősen alulreprezentált (mindössze 69,44%, miközben a HUSZ egészében ez az arány 96,66%), a *nyitott nap* viszont erősen túlreprezentált (30,56%, holott a HUSZ egészében csupán 3,34).

Vagyis itt is azt látjuk, hogy a *nyitott nap* forma relatíve tízszer gyakoribb, mint amilyen a HUSZ-korpuszban, és csaknem 14-szer gyakoribb, mint a magyarországi alkorpuszban.

Ugyanakkor az is figyelemre méltó, hogy a „szlovákiai” alkorpuszban is több mint kétszer gyakoribb a köznyelvnek tekintett *nyílt nap* forma, mint a tipikusan szlovákiai magyar *nyitott nap*. Ez arra utal, hogy a közmagyar és a helyi forma közt széleskörű váltakozás figyelhető meg a szlovákiai magyarban.

### Honnan jöhet a *nyitott nap*?

A *nyitott nap* kifejezés szemléleti alapja az a kép lehet, amely a kifejezés szlovák és más (pl. német, francia, román, orosz) nyelvű megfelelőit ihlette: szlovákul *deň otvorených dverí*, vagyis szó szerint „a nyitott ajtók napja”, egy olyan nap, amikor az érintett intézménybe idegeneknek sem tilos a bemenet. Bár az MNSZ2 adatával – amint láttuk – a tükörfordításos *nyitott ajtók napja* kifejezés szlovákiai magyar használatát nem lehet bizonyítani, más forrásokból (a Google kereső által kapott adatokból és részt vevő megfigyelésből) tudjuk, hogy a *nyitott ajtók napja* is széleskörűen használatos a szlovákiai magyar sajtónyelvben, csak a bejegyzés szerzője valami miatt erre nem figyelt fel, vagy nem tartotta szükségesnek foglalkozni vele.

A *nyitott nap* kifejezést talán azok a szlovákiai magyarok hozták létre, akiknek tudomásuk volt arról, hogy Magyarországon ennek a fogalomnak *nyílt nap* a neve, de a pontos magyarországi formát nem jegyezték meg, a jelzőt mintegy a *nyitott ajtók napja* formából „kölcso-

nözték”. A „kölcsonzést” nyilván megkönnyítette, hogy a nyitott ajtók képét a *nyitott nap* forma jobban asszociálja, mint a *nyílt nap*. A nyitott ajtók képe pedig nem irtható ki a szlovákiai magyarok lelkéből csak azért, mert a magyarországi magyarok lelkében esetleg nincs ilyen.

Ezenkívül esetleg a közmagyar *nyitott valamire* is elősegíthette a *nyitott nap* megerősödését, hiszen ilyenkor az intézmény vezetői mintegy nyitottak arra, hogy kívülállókat is beengedjenek az intézménye, hogy azok megismerhessék annak működését.

A *nyitott nap* tehát egyfajta kompromisszumos megoldásnak is tekinthető, amely mind a magyarországitól, mind a szlováktól eltér, de egyszersmind hasonlít is mind a kettőre.

A *nyitott nap* magyarországi előfordulásai aligha magyarázhatók a szlovákiai magyar nyelvhasználat hatásával, sokkal inkább az lehet rájuk a magyarázat, mint a szlovákiai magyar *nyitott nap* formára: az idegen nyelvekben meglévő kifejezés alapján létrejövő *nyitott ajtók napja* tükörfordítás és a *nyílt nap* keveredéséből jöhetett létre a „kompromisszumos” *nyitott nap* kifejezés, függetlenül a szlovákiai magyar *nyitott naptól* (bár természetesen a beszélőknek egy kisebb hányadára a szlovákiai magyar forma is hathatott – elsősorban azokra, akik olvassák a szlovákiai magyar sajtót).

A szlovákiai magyar beszélőnek tehát háromféle kifejezés áll rendelkezésére, ha erre a rendezvényfajta szeretne utalni. Választhatja a közmagyarnak tekintett *nyílt nap* változatot; főleg olyankor érdemes ezt tennie, amikor nemcsak helyi olvasóközönsséggel számol. Dönthet továbbá a valószínűleg a szlovák kifejezés tükörfordításaként létrejött, szemléletes *nyitott ajtók napja* forma mellett; ezt a döntést indokolhatja többek között az, ha olyan közönységnek ír, amely valószínűleg sokkal inkább ismeri ezt a formát, mint a közmagyar megfelelőjét. Ezenkívül választhatja mintegy a kettő kompromisszumaként a *nyitott nap* kifejezést is, már amennyiben nem hagyja magát elkedvetleníteni az olyan jó szándékú, de kártékony tevékenységet folytató laikusoktól, mint amilyen a fejezet elején idézett Facebook-bejegyzés szerzője.

A bejegyzés szerzője nyilvánvalóan nyelvtudományi végzettséggel nem rendelkező laikus. Ugyanis csak egy olyan ember gondolhatja, hogy egy „ártatlan” (és „ártalmatlan”) kifejezés létrejötte és elterjedése „betegség”, aki nem ismeri a nyelvműködés általános törvényszerűségeit, s a kétnyelvűség lélektanáról, továbbá a nyelvi kontaktológiáról sincs semmilyen fogalma, ellenben hisz a nyelvi purizmus ideológiájában, melyet számos szlovákiai magyar iskolában a mai napig lelkesen terjesztenek a tanítók és tanárok, főleg az idősebb nemzedékhez tartozók.

Mindenkinek szíve joga hinni a nyelvi purizmusban vagy bármilyen más nyelvi ideológiában, s emiatt ellenszenvet érezni a *nyitott nap* iránt. Mivel azonban a kontaktusváltozat(nak gondolt forma) – ahogy a bejegyzés írója is látja – elterjedt a köznyelvben is, s mivel azért magyarországi példák is akadnak rá, egyáltalán nem indokolt kárhóztatni ezt a formát (és főleg a használóit).

Nyugodtan rá lehet bízni a beszélőkre, adott helyzetben melyik kifejezést szeretnék használni. Nem érdemes az emberek nyelvérzékét megbolygatni és elbizonytalanítani, „betegségnek” titulálni egy kifogástalan nyelvi formát csak azért, mert esetleg szlovák hatásra jött lét-



re, ill. szlovák hatásra gyakrabban használjuk, mint a magyarországiak. A kétnyelvű beszélők – ha van választási lehetőségük, és ha mentesek a nyelvi purizmustól és más hasonló nyelvi ideológiáktól – azokat a nyelvi formákat részesítik előnyben, amelyek a két nyelvben jobban hasonlítanak egymásra. Ez a magatartás van összhangban a nyelvi gazdaságosságra való törekvéssel, amely a nyelvi változásoknak is egyik fontos mozgatórugója.

A bejegyzés akkor lett volna hasznos és építő, ha csupán felhívta volna a figyelmet arra, hogy érdemes megismernünk és használnunk a *nyitott nap* közmagyar szinonimáját is, a *nyílt napot*. Ennek gyakorlati haszna lehetne annyiban, hogy megkönnyíthetné a magyar–magyar kommunikációt. Igaz, egy magyarországi beszélőnek aligha okoz gondot a *nyitott nap* kifejezés megértése a szövegkörnyezet segítségével, nem beszélve arról, hogy ezt a kifejezést nyilván Magyarországon sem mindenki ismeri, vagy ha hallotta/látta is, ő maga ritkán él vele.

Elvben mégis akadhat, aki esetleg arra gondol, hogy van valamilyen fogalmi különbség a „nyílt nap” és a „nyitott nap” között. Az ilyen apró megakadások azonban mindenfajta nyelvi kommunikációban, egy országon belül is természetesek. A beszélők többsége elsiklik fölöttük, azért is lehet a *nyitott napra* az MNSZ2-ben több mint félszáz példa.

Summa summarum: nem a *nyitott nap* kifejezés a beteg, ahogy a Facebook-bejegyzés szerzője állítja, hanem az a gondolkodás, mely szerint nyelvi kérdésben megfelelő szakmai felkészültség és a nyelvi adatok alapos megvizsgálása nélkül is lehet ítélni, politikusokat, újságírókat, pedagógusokat, iskolaigazgatókat pellengérré állítani.

# & STÍLUSRÉTEG SZERINTI MINŐSÉGI VIZSGÁLAT

## ÷ 1. Általános kérdések

A találatoknak a stílusrétegek szerinti vizsgálata főleg a szavak és szókapcsolatok **stílusértékének** megállapításakor hasznos. Az alkorpuszok adatai **kétféleképpen** segítik a találati szavak stílusértékének megállapítását:

**1. Egyrészt** úgy, hogy lehetővé teszik a **találati szó kontextusának** vizsgálatát, vagyis a találati szó kontextusaként megjelenő szövegrészletből többnyire kisebb-nagyobb mértékben megállapítható a szöveg(részlet) stílusa, s a szöveg(részlet) stílusából sok esetben megbízható következtetés vonható le a találati szó stílusértékére nézve.

**2. Másrészt** úgy, hogy megvizsgáljuk a találati szó **relatív előfordulási gyakoriságát** az egyes stílusrétegekben, s ebből következtést vonunk le a stílusértékére és használatára nézve. Erről a következő, a *Stílusrétegek szerinti mennyiségi vizsgálat* című fejezetben lesz szó.

Ennek megfelelően ezt a fejezetet a stílusrétegek szerinti minőségi vizsgálatnak szenteljük, vagyis figyelmünket a konkrét szövegrészletek stílusára összpontosítjuk, ily módon kívánjuk megállapítani a találati szó stílusértékét. A következő fejezetben pedig mennyiségi szempontból tanulmányozzuk a találatokat. Fő kérdésünk az lesz, hogy az egyes alkorpuszokban milyen a találati szó relatív gyakorisága, s ebből milyen következtetések vonhatók le a szó stílusértékére nézve.

## ÷ 2. A szok + főnévi igenév példája

A következőkben a *szok* igének és a főnévi igenévnek a kapcsolatával foglalkozunk, korpuszadatok minőségi elemzésével.

A keresést a „részleges keresés” opciót kiválasztásával végezzük. Szóalakként nem konkrét szót írunk be, hanem a „szófaj” menüpontnál kiválasztjuk a „főnévi igenév” opciót:

The screenshot shows a search interface with the following elements:

- korpusz:** MNSZ2
- lekérdezés típusa:** részletes keresés
- részletes keresés:** szóalak
- összetett szó | morf:** (empty)
- szófaj:** TETSZ (dropdown menu open)
- Kontextus**
  - Szűrési szótövekre**
    - szűrési ablak: mindkét irányban | 5 szó.
    - szótövek: szok | mind
- Alkorpuszok**
  - alkorpusz: a teljes korpusz
  - doc.file
  - doc.region

The dropdown menu for 'szófaj' contains the following items:

- TETSZ
- névszó jellegű ...
- névelő
- névutó
- 
- ige ...
- igekötő
- főnévi igenév** (highlighted)
- határozói igenév ...
- 
- kötőszó
- indulatszó
- mondatszó
- előtag
- 'e' klitikum
- írásjel ...

Kontextusként a *szok* szótövet adjuk meg; itt szóalakot nem lehet megadni (pedig arra volna szükségünk), szerencsére ezt a hangsort a program nem ismeri fel igeként, ezért nem ad ki toldalékos adatokat.

Mivel a *szok* és a főnévi igenév közé más szavak is ékelődhetnek, s a szerkezet tagjainak szórendje fordított is lehet, a „mindkét irányban 5 szó” opciót választjuk, ezzel elérve, hogy akár négy szó is beékelődjön a *szok* és a főnévi igenév, ill. a főnévi igenév és a *szok* közé:

The screenshot shows a search interface with the following elements:

- korpusz:** MNSZ2
- lekérdezés típusa:** részletes keresés
- részletes keresés:** szóalak (dropdown), szófaj: főnévi igenév ... (dropdown)
- igekötő:** TETSZ (dropdown), **alapalak/szám:** alapalak (dropdown)
- összetett szó | **morf:** (input), **morféma:** TETSZ (dropdown)
- Kontextus**
  - Szűrés szótövekre**
    - szűrési ablak:** mindkét irányban (dropdown), 5 (input) szó.
    - szótövek:** szok (input), mind (dropdown) forduljon elő.

Persze az ilyen „nagyvonalú” beállításnál sok irreleváns találatra számíthatunk, melyeket a másodlagos tisztítás során el kell távolítani. Ám tekintettel arra, hogy a vizsgált szerkezet erősen megbélyegzett, s ezért – írásban legalábbis – inkább csak a kevésbé iskolázott beszélők használják (akiknek a szövegei csak csekély mértékben kerülnek be a szövegtárba), nem számítunk sok találatra, így a másodlagos tisztítás feltételezésünk szerint nem lesz túlságosan időigényes feladat.

Mivel az összes stílusréteget meg szeretnénk vizsgálni minőségi szempontból, ezúttal a teljes korpuszal dolgozunk, nem hagyjuk ki a „személyes-közösségi” alkorpuszt sem. Ezért vagy az egyik stílusréteget sem jelöljük ki, vagy az összeset kijelöljük (ezt úgy is megtehetjük, hogy a „mindet” gombra kattintunk), mindkét esetben ugyanazt az eredményt kapjuk:

The image shows two screenshots of a document style selection interface, separated by an equals sign (=).

- Left screenshot:** A list of document styles under the heading 'doc.style'. All checkboxes are unchecked. At the bottom, there is a button labeled 'Mindet'.
- Right screenshot:** The same list of document styles, but all checkboxes are checked. At the bottom, there is a button labeled 'Egyiket se'.

Amint erről a régió szerinti alkorpuszok bemutatásánál szó volt, nem számít, hogy az egyik stílusréteg sincs-e kijelölve vagy az összes ki van jelölve, a „Konkordancia készítése” gombra kattintva mindkét esetben a találatok száma azonos: tisztítás nélkül 453.

Nézzük most meg az első néhány találatot, a bal oldali kontextus (vagyis a találat előtti szó) szerint rendezve, „mondat” (azaz nem KWIC) nézetben!

korpusz: MNSZ2	
találatok száma: 453	
személyes-fórum	Ő alá <b>szok ' írni</b> , hogy Soltika+ Fikikutya-aszörösgyerek (.cvnvnjfr88888+ -) aköbön. éljensáprlis.. semmintnemertő... délenkékazég..!!!
személyes-fórum	Ettől még lehetne jó, magam már régóta szeretnék befejezni egy kisregényt, nem ez a stílusom, de ilyet is <b>szok ' írni</b> .
sajtó	Hordót is <b>szok ' mosni</b> .
személyes-fórum	Az a fránya membrán is ki <b>szok ' repedni</b> !
személyes-fórum	Nem, csak a végén mindig késelés <b>szok ' lenni</b> .
személyes-közösségi	A padlástéri tűz nem kis eséllyel áramlopás miatt <b>szok ' keletkezni</b> amúgy, de persze a párásztok vegyenek mindenből újat nekik.
személyes-közösségi	A probléma ott <b>szok ' lenni</b> , hogy ha mégsem tudnánk, kik a hutuk/tuszik, akkor mindig kitermelődnek - helyesebben kitermelünk magunknak - olyanokat, akik megmondják nekünk, hogy kik a hutuk/tuszik.
személyes-közösségi	A probléma ott <b>szok ' lenni</b> , hogy ha mégsem tudnánk, kik a hutuk/tuszik, akkor mindig kitermelődnek - helyesebben kitermelünk magunknak - olyanokat, akik megmondják nekünk, hogy kik a hutuk/tuszik.
személyes-fórum	Én mennék, ha odaérek, remélem kitartotok, mer ' lapzártázom, és az sokáig <b>szok ' tartani</b> :-)
személyes-fórum	A kisebbik ma hatalmasat alakított: feleségem hatalmas pocakjával az ágyon <b>szok ' heverni</b> , és gyönyörködik, hogy a pici hogy mozog benne, rá-ráteszi a kezét.
személyes-fórum	Szerintem a Csülök most valami alkotóháza ' csücsül... és isz, ha van esze ( <b>lenni szok</b> neki)... de ha rólam is ír az írói munkásságába bele, hogy a felügyelő úr aztán mindenféleket kérdezzesen, akkor rágurítok egy hordót!
sajtó	Jelenleg legfontosabb feladatuk megvédeni a hidakat, <b>megakadályozni</b> a jégtorla <b>szok</b> kialakulását a folyókon, szükség esetén robbantásokkal megnyitni az utat a jégtablák levonulása előtt.
személyes-fórum	Ugrálni nem ugrálok, <b>úszni</b> a hógyerekek <b>szok</b> , én csak hazacipelem.
szépirodalom	<b>Vágni</b> .
személyes-fórum	<b>Énekelni</b> magyarul nem <b>szok</b> , mer ' erős az akcentusa...
személyes-fórum	demeter aszondta ő <b>szok</b> 100-al <b>húzni</b> felfelé az érdi emelkedőn... melleleg már 95 is sok:))
személyes-közösségi	Ahogy <b>mondani szok</b> ....
személyes-fórum	Na, visszajöttem a Pincéből, ahol egy kicsit pihentem, meg ittam is, ahogy <b>lenni szok</b> , de most sem értek egy mukkot sem.
személyes-fórum	Ha Söröshordó, akkor benne vagyok, mer'hogy megszomjaztam, ahogy <b>lenni szok</b> nálam...
személyes-fórum	Ezáltal teljesen olyan lesz minden, ahogy <b>lenni szok</b> , senki sem tud és nem ért semmit, én pedig rendesen be tudok rúgni:))) Na, áll az ajánlat?
személyes-fórum	Mint minden más, ez is majd helyrerázódik, ízé, helyreáll, vagy helyredől, na, ahogy <b>lenni szok</b> ....
személyes-fórum	Szögedében mindenki Lyól berúgott, oszt ' annyiban maradtunk... - utána a társaság társulatilag támolygott tovább szana is, meg széjel is, továbbá összességében vissza, ide ni, ahogy <b>lenni szok</b> kluboséknál, így aztán csak ezek a marhára bizalmas jelentések maradtak meg... ahogy lenni szok...
személyes-fórum	Az euroastra csak a magyar ferdítést közli, azt is lopták valahonnan, ahogy <b>lenni szok</b> , de az AltaVistával meg lehet találni az angol nyelvű " szabályokat " is, tessék, lehet válogatni: Netiquet 1, Internet Etiquette... van ott 666 találat, már ebből is látni, hogy Number of the Beast:)))
személyes-fórum	Szögedében mindenki Lyól berúgott, oszt ' annyiban maradtunk... - utána a társaság társulatilag támolygott tovább szana is, meg széjel is, továbbá összességében vissza, ide ni, ahogy lenni szok kluboséknál, így aztán csak ezek a marhára bizalmas jelentések maradtak meg... ahogy <b>lenni szok</b> ...
személyes-fórum	Na, pedig haza akartam <b>menni</b> .
személyes-közösségi	A ház nem <b>szok</b> 1 nap alatt <b>leamortizálódni</b> , talán, ha kijártak volna évekkal ezelőtt és segítettek volna a házat karbantartani.
személyes-közösségi	Ha mégis kedvem szottyanna valamiért alkoholt <b>inni</b> (ahogy sokan másoknak <b>szok</b> szottyanni), akkor bárban amúgy sem tenném, mert 1: nem járok bárkba 2: ott már most sem lehet rágyújtani mellé 3: diszkóba sem járok 4: ott sem lehet már rágyújtani mellé 4: kimondottan jól esik a hír, hogy csak otthon lehet piázni (hogy előadásokon is?? akkor iratkozunk be egyetemre, ott is előadnak, nem?), és azt is csak úgy, ha másokat nem zavarunk vele..
személyes-közösségi	Az a lényeg, ne szaporodjanak ivarosán, ami <b>teljesülni szok</b> .
személyes-fórum	Na, má ' megin ' televagyok ezzel a duplántriplántálással, meg olyan deliriumom is van, ami <b>lenni szok</b> , csupa ördög szaladgál itt körülöttem, még szerencse, hogy nem az én korsóm után kapkodnak:))
személyes-fórum	Nem pocsék az a cikk, csak olyan, mint amilyen <b>lenni szok</b> mindig, ha internetről van szó.
személyes-fórum	Amivel <b>pisilni szok</b> !
személyes-fórum	Viszont a wizard várban nem nagyon <b>szok</b> animate-dead <b>lenni</b> (csak resurrection).
személyes-fórum	Sajna nyáron a tengerparton biza fény van meg kékség, spec ez engem nem <b>szok</b> annyira <b>zavarni</b> , sőt...))) puszol: stewe

Ha most egymás után mindig csak egy-egy stílusréteget pipálunk ki (kivéve a „személyesnél”, ahol mind a két „személyeset” kipipáljuk), akkor az egyes stílusrétegekre nézve kapunk adatokat.

Az alkorpuszok szerinti vizsgálatot most tehát a **szok** + főnévi igenév példán szemléltetjük, stílusrétegenként haladva.

Mivel az adatok harmadlagos tisztításához a dokuazonosítóra is szükségünk van, a „Megjelenítés” menüpont alatt az eddig megjelenített „stílus” opció mellett a „dokuazonosító” opciót is kiválasztjuk. A jobb oldali képen ezt a beállítást látjuk.

## ÷ A) „szépirodalmi” alkorpusz találatainak vizsgálata

Az alábbi képernyőkép a „szépirodalmi” alkorpuszba tartozó összes találatot tartalmazza, bal oldali kontextus szerinti rendezésben:

korpusz: MNSZ2	
találatok száma: 23	
doc#2091,szépirodalom	Vágni .
doc#616,szépirodalom	Másokét, persze. Már ahogy ez <b>lenni szok</b> .
doc#1845,szépirodalom	Másokét, persze. Már ahogy ez <b>lenni szok</b> .
doc#2017,szépirodalom	Madame D'Array azt mondja: verebeket látott tegnap, ezt nagyon jól érti, szóval az átlagos veréblátásnál élesebben láthatta őket, s azt kérdi: milyen lenne, ha a mi két kis barátunk is úgy ugrálna, menne az úttestre, állna a járdaszélen, ázna, igazsá van -e Schmidt Egonnak, vagy csak kedves akar lenni velünk - és (Madame D'Array) két ilyen verébmentő meg is érdemli a kedvességét; s tudod, ő nem <b>szok</b> ilyeneket <b>mondani</b> ; erről most leszokom, mert a macskák különben kikezdi még Szokol nevét, holott az akkor... mert a kis rádió tegnap, képzeld, még a Labdáját is magával hozta a nagyobbik szobába, mert a magyar-jugoszláv vízilabdát hallgattam (ezek a közvetítések nekem olyan visszapergetők; nem is az ötvenes évekig, inkább úgy hat-hét évvel vissza), nem érdemli, hogy tréfálkozzanak vele, s ő oly komoly és elszármazott; hát a Madame D'Array, ha ő ilyet mond, akkor lehet is ilyet mondani -, nem, azt mondtam erre tegnap, tényleg
doc#613,szépirodalom	Úgy <b>szok</b> a kezük <b>remegni</b> , hogy a végén nem bírják.
doc#1703,szépirodalom	Ilyen kifordított szórendeket máshol nem használnak, ha úgy jön ki, még a " <b>szok lenni</b> " -t is használja.
doc#1310,szépirodalom	Így " <b>szok lenni</b> " általában, ha elsők a Zsuzsi lába.
doc#1936,szépirodalom	Addig <b>szok maradni</b> a vénjányánál.
doc#757,szépirodalom	Kár - teszi hozzá -, hogy a hallgató részeket nem hallotta, mert kissé nagyothall a bal fülére, s épp azon <b>szok feledni</b> .
doc#2070,szépirodalom	Márhogy amint bennünk a bölcsesség <b>szok támadni</b> .
doc#753,szépirodalom	Egy héten kétszer, pontosan szerda és szombat délután, két fiatalember <b>szok megjelenni</b> az udvar mélyén.
doc#529,szépirodalom	- Nem, ezen látszott, hogy <b>szok mosdani</b> .
doc#608,szépirodalom	S hogy na ja, éppenséggel van, aki ki <b>szok járni</b> , kíváncsiságból, vagy ki tudja.
doc#756,szépirodalom	Nagyon jól látom, hogy Kasosné mindent ezekkel a kocsisnékkel <b>szok elpletykázni</b> .
doc#1874,szépirodalom	- Mire megjöttem nagymamáméktól, már aludt, de máskor is korán <b>szok lefeküdni</b> , mert reggel ötre jár takarítani...
doc#1473,szépirodalom	Mindent <b>szok tudni</b> .
doc#356,szépirodalom	- Fejít, gyomrát, mindenit, ahogy már <b>szok lenni</b> .
doc#1936,szépirodalom	Ott <b>szok árulni</b> .
doc#1874,szépirodalom	- Nekünk sem <b>szok feledni</b> !
doc#613,szépirodalom	Sokat <b>szok verni</b> ?
doc#371,szépirodalom	Merthogy éjjel <b>szok festegetni</b> .
doc#373,szépirodalom	Merthogy éjjel <b>szok festegetni</b> .
doc#302,szépirodalom	- Mert az már csak mindig úgy <b>szok lenni</b> , hogy kin a nyelvüket köszörülgetik, az nem tud öröla.

Már a fentebb látható képen is feltűnik, hogy a második és a harmadik mondat azonos. Mind a kettő Petri Györgytől való, két különböző könyvből, de az egyik könyv a szerző összegyűjtött verseit tartalmazza, ez magyarázhatja az azonosságot.

Még van két további azonos találat is, mégpedig Jókai Anna két különböző művéből. A megjelenített mondat teljesen azonos. Mivel azonban a mondat elég rövid – mindössze négy szóból áll, felmerülhet a gyanú, hátha csak véletlenül azonosak. Ezért mindkettőnél a találati szóra (a főnévi igenévre) kattintva megnézzük a tágabb kontextust (a megjelenő mezőben a két + jellel tovább bővítjük a kontextust). Ezt az eredményt kapjuk:

doc.file lit\_hu\_dia\_Jokai\_Anna\_\_\_A\_labda\_\_\_1971.clean  
doc.region magyarországi  
doc.style szépirodalom  
div.type novella

- Mit gondolsz? - kérdezte Mari. - Ezt odaadják Neki? El kell olvasnia? Az asszony, a horgoló-túvel, felnézett rájuk egy pillanatra. - Hogyan? - szólt szolgálatkészen. - Tessék mondani... a művész, úgy értem, a festő... látta már...? A néni fölényesen mosolygott. - Minden este megtekinti. A Mester. Zárás után. Ötől hétig pedig ott ül, kinn, a foajében. - És mit csinál? - kérdezték rémülten, szinte egyszerre. - Alszik. Többnyire alszik. Merthogy éjjel szok **festegetni** . Nappal meg dolgozik. A Mester. Mint akármelyik rendes ember, bizony... Lábjujhegyen elosontak mellette, az előcsarnokban. Még az arcát se látták, belelógatta a kabát felhajtott nyakába. - Láttad? - Nagyokat szippantottak a friss őszi levegőn. - Fantaszi-

kus! - Esküszöm, fordítva volt rajta a kabát. Esküszöm... Magyar János és Magyar Jánosné is látták a férfit a padon, fél órával a zárás előtt, Magyar né oda is akart lépni. - Hátha várják valahol - mondta szelíden -, fiam, ki tudja, nem kéne-e felébreszteni

**doc.file** lit\_hu\_dia\_Jokai\_Anna\_\_\_Az\_ifju\_halasz\_es\_a\_to\_\_\_1992.clean

**doc.region** magyarországi

**doc.style** szépirodalom

**div.type** novella

- Mit gondolsz? - kérdezte Mari. - Ezt odaadják Neki? El kell olvasnia? Az asszony, a horgoló-túvel, felnézett rájuk egy pillanatra. - Hogyan? - szólt szolgálatkészen. - Tessék mondani... a művész, úgy értem, a festő... látta már...? A néni fölényesen mosolygott. - Minden este megtekinti. A Mester. Zárás után. Ötől hétig pedig ott ül, kinn, a foajében. - És mit csinál? - kérdezték rémülten, szinte egyszerre. - Alszik. Többnyire alszik. Merthogy éjjel szok **festegetni**. Nappal meg dolgozik. A Mester. Mint akármelyik rendes ember, bizony... Lábujjhegyen elosontak mellette, az előcsarnokban. Még az arcát se látták, belelógatta a kabát felhajtott nyakába. - Láttad? - Nagyokat szippantottak a friss őszi levegőn. - Fantasztikus! - Esküszöm, fordítva volt rajta a kabát. Esküszöm... Magyar János és Magyar Jánosné is látták a férfit a padon, fél órával a zárás előtt, Magyar né oda is akart lépni. - Hátha várják valahol - mondta szelíden -, fiam, ki tudja, nem kéne-e felébreszteni

Látjuk, hogy a nagyobb szövegrészletek is teljesen azonosak. A késsel szedett részre (a dokuazonosítóra, a régióra vagy a stílusrétegre) kattintva megjelennek a doku alapadatai. Ezekből kiderül, hogy két különböző munkáról van ugyan szó, de mivel a szöveg műfaja novella, feltehető, hogy ugyanaz a novella lett két különböző gyűjteményes kötetbe beválogatva. Ezért a tisztítás során az egyik találatot bátran törölhetjük.

Az elsődleges tisztítás eredményeként tehát 21 találatunk maradt. Az elsődleges tisztítás során mindazokat a találatokat eltávolítjuk, amelyekre **semmilyen vizsgálatban nem lehet szükségünk**, amelyek tehát a **szöveggyűjtés tökéletlenségének** a következményei (ti. hogy több azonos szöveg is bekerült, aminek nem szabadna előfordulnia).

A másodlagos tisztítás során azokat a találatokat távolítjuk el, amelyek nem a program hibájából keletkeztek, de **a mi vizsgálatuk szempontjából irrelevánsak**. A „szépirodalmi” alkorpuszban egyetlen irreleváns találat van az elsődleges tisztítás után maradt 21 találat közt, a következő:

**doc.file** lit\_hu\_dia\_Lazar\_Ervin-Naplo

**doc.region** magyarországi

**doc.style** szépirodalom

**div.type** dummy

vette fel a nevezetes sikerdíjat. Mintha bizony ez volna a kérdés. Tocsik csak ötödrendű cinkos egy rablásban, amelyet az SZDSZ és az MSZP követett el az állami vagyon - azaz közösségi vagyon kárára. Mit Tocsikoznak még mindig?! VISSZA a megírandóhoz! Hogy esetleg rákényszerítse! Guplitz [?] kivitte az állomásra, Halbax volt a gépész. Nem tudok írni, Aranyt olvastam. Keveháza, Buda halála (vagy az egésznek nem ez a címe). Ilyen kifordított szórendeket máshol nem használ, ha úgy jön ki, még a „szok **lenni**”-t is használja. Ám az abajgat ígét aki dédelget értelemben használja, itt feleletet kap: (talán 2. ének) épp ellenkezőleg, zavar, nem hagy nyugton. Tegnap, születésnap ajándék, egy nagy piros daganatot fedeztem föl a bal lábam nagyujján. Mintha egy csontgöb nőne a bőr alatt. Lehet, hogy ez közsvény?

Enyhe, égető érzés, nyomás, ma fájdalom. Módosnak kellene írni valamit az Európai Utasba. Tárgy? Sasóka. De tegnap láttam az Oktogonon a levágott kezű kislányt. 2-3 éves lehetett, az anyja

Ezt a találatot azért kell kihagyni a statisztikából, mert a kifejezés itt nem természetes módon van használva, hanem **metanyelvi értelemben**. A metanyelvi értelemben használt elemeket a statisztikából kihagyjuk, viszont **példamondatként nagyon hasznosak**, így ezek feltétlenül idézendők a dolgozatban, a többi találattól elkülönítve. Tanulságos ugyanis, hogy a szerzők milyen állításokat fogalmazznak meg egy-egy nyelvi formával kapcsolatban.

Harmadlagos tisztítást csak szigorúbb vizsgálat esetén végzünk, amikor **rendkívül fontos az adatok megbízhatósága**. Ilyenkor a **dokuazonosító** szerint rendezzük a találatokat (a „Rendezés” menüpont alatt az „infóra” kattintva), hogy könnyen észrevegyük, ha ugyanaból a dokuból két vagy több idézet is szerepel a találatok közt.

Most nem azonos idézetekről van szó, amelyek a teljesen azonos tartalmú és megfogalmazású dokuból származnának, hanem ugyanabból a munkából való eltérő idézetekről, melyekben csupán a találati szó azonos.

Ezzel az eljárással el tudjuk kerülni, hogy adatainkat torzítsa egy-egy szerző stílusának valamely sajátossága.

Meg kell azonban jegyezni, hogy **a doku szerinti rendezés mint opció sajnos már évek óta nem működik, hibajelentést ad**, ezért aki harmadlagos tisztítást szeretne végezni, az **elmentheti** a konkordanciát a fent már ismertetett módon, majd **táblázatba illesztheti, a táblázatban pedig a megfelelő oszlopban már betűrendbe állíthatja a dokuazonosítókat**.

Elvileg tovább szigoríthatnánk a szűrést, és egy-egy szerzőtől csak egyetlen találatot őrizhetnénk meg, ez is hasznos volna, ott, ahol még nagyobb jelentősége van a pontosságnak.

A mostani vizsgálatunkban a „szépirodalmi” alkorpuszban három dokut találunk meg kétszer (egyforma a dokuazonosító számuk), a következőket:

doc#1874, szépirodalom	- Mire megjöttem nagymamáméktól, már aludt, de máskor is korán szok <lefeledni >, mert reggel ötre jár takarítani...
doc#1874, szépirodalom	- Nekünk sem szok <felelni >!
doc#1936, szépirodalom	Addig szok <maradni > a vénjányánál.
doc#1936, szépirodalom	Ott szok <árulni >.
doc#613, szépirodalom	Úgy szok a kezük <remegni >, hogy a végén nem bírják.
doc#613, szépirodalom	Sokat szok <verni >?

A találat információs (kékkel szedett) részére kattintva meg tudjuk nézni, melyek azok a munkák, amelyekből két-két találat van a 20 között:

doc#1874	Rakovszky Zsuzsa A Hold a hetedik házban
doc#1936	Szabó Magda Disznótor-Kigyómarás
doc#613	Parti Nagy Lajos Ibusár -Mauzóleum

Ha a statisztikába egy-egy dokuból csak egyetlen találatot akarunk figyelembe venni, akkor a 20-ból levonunk 3-mat, marad 17 találatunk.

### ***Az adatok minőségi elemzése***

Ha elolvassuk a találatok kibővített szövegét (a találati szóra kattintva, majd a fönti és a lenti + jelre is rákattintva), könnyen rájövünk, hogy a találatok többsége **párbeszéd**ekből való, tehát **nem az íróknak, hanem a szereplőknek a stílusára jellemzőek**.

Nézzünk meg néhányat! Ahol lehet, helykímélés céljából csökkentjük a kontextus méretét.

<p>– Nem tudok én semmit. Csak amit a népek beszélnek. – Ők meg mit beszélnek? – Tudod, hogy van az. Némelynek csak jár a szája összevissza, tanulsága meg semmi. – Én még nem hallottam. – Aztat nem csodálom. – Azután miért? – Mert az már csak mindig úgy <u>szok lenni</u>, hogy kin a nyelvüket köszörülgetik, az nem tud őróla. – De mostan már mondjad – így a Rozi élesedő szóval. – Ugyan mit fecsegnek az én Jánosomról? – Hogy például látták őket együtt Erzsivel az utcán. (Gyurkó László)</p>
<p>– Hát ez a nagy gyerek is megy majd aratni? Mi jót rágsz olyan étvággyal, kisöcsém? – Csak ennek a hőmérőnek az a tokja. Azt kedveli, mert hogy olyan fényes. Dehogy jön ez velünk. Beteg ez szegénykém. – Mi a baja? – Tegnap már harmincnolca volt, hálistennek, mert tegnapelőtt harmincnolc öt. Nem győzöm mérni. – Fájdítja valamijét? – Fejit, gyomrát, mindenit, ahogy már <u>szok lenni</u>. – Tán meghült. – Az. És az ágyban is csak forog. Nem tudom, mi érhetne. Már vizet is öntettem neki. (Illyés Gyula)</p>
<p>Lajos? Megint itt vagy a kurlival? Idén már jártál itt. – Az tavalyelőtt volt, Géza. Én kétévenként jövök. Volt valami a télen? – Semmi, Lajos. Olyan, mint máskor: a politikusok ugatnak, a népek éhesek. – A kutat senki sem piszkálta? Az éjjeliőr eltűnődött: – Hát, ahogy így mondod, eszembe jut; épp tegnap éjszaka láttam itt egy öreg, szakállas férfit. – Nem a kukás Béla volt a „Lordok Házá”-ból? – Nem, ezen látszott, hogy <u>szok mosdani</u>. A haja körül egy olyan furcsa fény lebegett, mintha egy elemlámpa égne a fejében. Tudod, tanultuk hittanban, hogy hívják, na mondd már... (Moldova György)</p>
<p>Úgy van, kérem, nyomott hangulatú hely, nagyon nyomott. Emlékezet? Uram, vannak súlyok, amikre ha muszáj, se lehet emlékezni. Ha meg nagyon nyom, érti, ott belül, hát inkább szálljon le a bicikliről és tolja. Igen, pont tíz perc kényelmes tempóban, bár ezt ő nem tapasztalatból mondja. S hogy na ja, éppenséggel van, aki ki <u>szok járni</u>, kíváncsiságból, vagy ki tudja. Egy ilyen örökös özvegyasszony-forma, az jön, vagyis megy, úgy mondják rendszeresen. Lerakja a biciklijét, a virágját, és áll a gazban, nem tudom, de két órát is egyfolytában. Beszélük, hogy biztosan van ott halottja. De hát kinek nincs? (Parti Nagy Lajos)</p>
<p>Tulajdonképpen micsoda foglalkozása van magának a városban? – kérdezte tőlem. – Kari munkás maga? Mestersége van? Vagy csak ilyen írnok? – Ahogy jobban fizetnek! – felelek. – És hasznomat veszük! – Úgy? – mondja Kasosné. – De legszívesebben tekereg! He? – Hát ahol úgy veszük jobban hasznomat! Kasosné néz a kocsisnékra. Azok, látom, csak kékülnek el a kacaghatnéktől, hogy én csak pipázok, és Kasosné táncolna mérgében, mint a forgószél. Nagyon jól látom, hogy Kasosné mindent ezekkel a kocsisnékkal <u>szok elpletykázni</u>. Már szemről értik egymást. (Tersánszky Józsi Jenő)</p>
<p>– Samu bácsi! hát nem itten alszik Belluska? – halljuk a kétségbeesett Cincárt. – Itt, hogyne – feleli nagyapja. – De az éjjel benn aludt nagynénijénél a városban. Azzal nagyon köszöni nekünk a szép zenét. Kár – teszi hozzá –, hogy a hallgató részeket nem hallotta, mert kissé nagyothall a bal fülére, s épp azon <u>szok feküdni</u>. Cincár, Cincár! Ha meg nem öl ez a kudarc, meg kell öljön röhejünk. (Tersánszky Józsi Jenő)</p>

Van egy olyan találat, amelynek nem igazán beszéltnyelvies a stílusa:



Ez nem az, ahol Jucika lakik. Udvarára példának okáért be szabad jönni zenészeknek, és hangversenyt adni ott. Hogy nem mindig fület gyönyörködtetőek ezek az udvari előadások? ... Hát nem! De azért akad köztük tűrhető is. Egy héten kétszer, pontosan szerda és szombat délután, két fiatalember szok megjeleni az udvar mélyén. Az egyik hegedűs, a másik harmonikás. A hegedűsnek van fuvolája is. Azon is játszik. A harmonikás pedig énekel. A harmonika egy fekete, öreg jószág, kicsikét tüdőbajos. Megesik, hogy bizonyos kombinációban hang helyett csak egy éles, semleges zihálás jön ki belőle. De mindent kipótol a harmonikás éneke. Remek tenorhangja van ennek a harmonikásnak. És aztán valami ellenállhatatlanul szívrebegettő zengést tud adni neki, ha akar. A ház méltányolja nagyon ennek a két udvari zenésznek a produkcióját. (Tersánszky Józsi Jenő)

Érezhető viszont, hogy a részlet, a választékos fogalmazás ellenére, párbeszédre utal vissza, az magyarázhatja a *szok* + főnévi igenév használatát.

Van továbbá egy találat, amely Petri György verséből való. A verset most az internetről idézzük, eredeti formájában, mivel a találati felület – amint az fentebb kiderült – nem engedi az egész szövegek megjelenítését:

#### **Ha minden kötél szakad**

„Nem tudtam mire menni Stanley-vel.”  
Mondta a regényhősnőm. Az a baj, hogy én se.  
Mondom én most meglehetősen utólag és megkésztetten, sőt megkéstelten (előfodrul <sic! sicc> ilyesmi, mint most a vékonybelem). Nem sok időm van ... hová is, ja igen *előre*, mert *hátra*, az kérem szépen:  
roskad rám döggivel. Az van. Múltam. Rovott.  
Mint mindenkinek. Ember voltam (végtére is? valahol? valaha?). Kinek mi köze hozzá?  
Nehéz volt az élet. Én is nehezítettem.  
Másokét, persze. Már ahogy ez lenni szok.  
Hát eképpen lettem ily tündöklő zsebpiszok (a barátaimnak: „pipamocsok”). Zárójel zárva.  
Mint a boltok vasárnap délután.  
Az összes lehetőség már kizárva:  
anyám még él, így nem lehetek árva.  
Sem. Úgy tetszik, afféle sem-sem vagyok, félelmemben el-eldalolgotok;  
voltaképp azt akartam csak mondani:  
hogy a legkevésbé sem mindegy, hogy a kötél akasztott ember nyakán,  
vagy kötélzáncos lába alatt szakad.

Látjuk, hogy a vers telis-tele van **stilisztikailag jelölt** formákkal, sőt egyéni szójátékkal is (*előfodrul*), ráadásul a *szok* használatát még az is elősegíti, hogy **rímhelyzetben** van (*szok* – *piszok*). Ilyen körülmények közt az mondhatjuk, hogy a *szok* nagyon illik a vers stílusához, de ebből nagy hiba volna levonni azt a következtetést, hogy a vizsgált nyelvi forma része volna a költői nyelvnek.

A célszót tartalmazó szövegrészek elolvasása után megállapíthatjuk, hogy **egyetlen találatról sem mondható el, hogy a szépirodalmi stílus része volna**, bár mindegyik jelenlétére kielégítő magyarázattal szolgálhatunk.

A vizsgálatból levonhatjuk azt a tanulságot, hogy a „szépirodalmi” alkorpusz esetében egyáltalán **nem mindegy**, hogy a szerző **saját szövegéről** van szó, vagy pedig **párbeszédreletről**, melynek stílusa és nyelvhasználata **nem az íróra, hanem az irodalmi mű szereplőjére jellemző**.

## ÷ B) A „sajtónyelvi” alkorpusz találatainak vizsgálata

A „sajtónyelvi” alkorpuszban tisztítás előtt 8 találatunk van, a következők:

korpusz: MNSZ2	találatok száma: 8
doc#2288,sajtó	Hordót is <b>szok ' mosni</b> .
doc#2202,sajtó	Jelenleg legfontosabb feladatok megvédeni a hidakat, <b>megakadályozni</b> a jégtorla <b>szok</b> kialakulását a folyókon, szükség esetén robbantásokkal megnyitni az utat a jégablak levonulása előtt.
doc#1071,sajtó	az orrát is betörte -nem volt rá tanú! -, de a Kieren Fallon egy zseni, letöltötte lovászként és ösztöndíjjal tanulva Kaliforniában a fél év büntetését, hazatért Angliába - merthogy ő északír -, és felhívta magára a figyelmet, és hamarosan a legnagyobb lovakat kapta meg a dubai társaság ellentáborától, Henry Ceciltől és ilyesmí (ő szintén tréner), és már úton van a harmadik évi zsokéchampionátus felé; na szóval felhevültem, mert Fiery (Tüzes) nyert, úgy méghozzá, hogy rögtön az élre állt (Fallonnal), aztán ilyenkor <b>szok</b> az <b>lenni</b> , hogy az utolsó kanyarban bájj.
doc#1071,sajtó	Simó György (ex-Narancs, most: Matáv) és Jaksity György (Concorde Rt.) halk és fegyelmezett, a Szerbhorvát azért bekérdez, hogy mi van a korrupcióval, én is rákérdezek a médiaügyekre, elindulna egy életes duma, de ahogy <b>szok ' ez lenni</b> , a műsorvezető időhiányra hivatkozva lefújja a meccset.
doc#1071,sajtó	Lenyomnak egy belső vizsgálatot, melynek az eredménye az imént jelölt százalékarányban <b>szok</b> negatív <b>lenni</b> .
doc#2291,sajtó	Pediglen havonta többször is meg <b>szok ' neki jőni</b> a vérmes vélemény, sőt, a " Best of OV 2009. " videó is ott vigyorg.
doc#2288,sajtó	Anyukám néha <b>szok</b> süteményt <b>sütni</b> , kirakja az asztalra, és aztán megeszi az egészet.
doc#1071,sajtó	és amiért tájhazánkban, az Észak-Balkánon is szokott, <b>szok</b> és <b>szokni</b> fog.

Látjuk, hogy a találatok mind különböznek egymástól, vagyis az elsődleges tisztítás után is 8 marad a találatok száma.

Viszont a másodlagos tisztítás során ki kell hagyni a második találatot, mivel ott elírás folytán jött létre a *szok* betűsor, nincs köze a *szok* igéhez, s véletlenül áll előtte a *megakadályozni* ige, a kettő közt nincs kapcsolat. A másodlagos tisztítás következtében tehát az adatok száma 7-re csökkent.

Ha harmadlagos tisztítást is akarunk végezni, egy-egy dokuból csak egy-egy találatot veszünk számba. Mivel összesen 8 találatunk van, átrendezés nélkül is könnyen észrevevesszük, hogy a doc#1071-es jelzetű dokuból 4 találatunk van, a doc#2288-as jelzetű dokuból pedig kettő. Így a harmadlagos tisztítás során a statisztikából 3+1 találatot törölünk, s marad 3 találatunk.

### **Az adatok minőségi elemzése**

Érdeemes észrevenni, hogy két példában is **hiányjel** van az ige után: *szok'*, ami jelzi, hogy a szerző **tudatában van** annak, hogy **stilisztikailag jelölt elemet használ**. Még így, „Mondat” típusú megjelenítés mellett is látszik, hogy ezek nem igazán sajtónyelvi mondatok, s ez még egyértelműbbé válik, ha a főnévi igenevekre rákattintva megnézzük a szöveggörnyezetet is.

Például a legrövidebb mondat, a *Hordót is szok' mosni* kontextusa ilyen:

Én a Misire nem mondhatok semmit. Jóvan, szépen beszélek vele, mindent meg is tudok vele csináltatni, amit kell. – A Misi már itt énnálam negyedik éve dolgozik. A férjem akkor lett beteg, akkor halt meg. Hát lehet, hogy ha szegény élne, még akkor is kellene. Misikém kapálni szokott, aztán most majd a kukoricát morzsoljuk le, majd amikor olyan idő lesz. Azt is leszedi nekem, a kukoricát ősszel, a szárát levágja. Télen meg havat lapátol, fát vág. Hordót is szok' mosni. Mindenhez ért! Meg permetez. Ritka ember, aki így tud permetezni, mint ő! Isten bizony! Az nagyon helyes, hogy mondja, ő nem be nem kever semmit. Be kell keverni a vegyszert, és ő úgy csinálja. Nekem elég jól megmaradt a szőlőm azóta is. Bár mondják a szomszédok, hogy láttuk, hogy permetez a Misi, megy ott mindenfelé! Mondom az a jó, ha mindehova megyen! Menjen a földre is, hadd veressen!

Bizonyára **interjúrészeletről** van szó vagy valamilyen **vendégszövegről**, vagyis ez az idézet **nem képviseli a sajtónyelvet**.

A többi találatról hasonlókat mondhatunk, tehát nagy valószínűséggel megállapíthatjuk, hogy a *szok* + főnévi igenév szerkezet **nem része a sajtónyelvnek**. (Biztosat az adatok alapján persze csak a „sajtónyelvi” alkorpuszról mondhatunk, nem a sajtónyelvről általában.)

### ÷ C) **A „tudományos” alkorpusz találatainak vizsgálata**

A „tudományos” alkorpusz taláataival könnyen „elbánunk”, mivel csak kettő van belőlük, ráadásul ezek is ugyanabból a dokuból valók:

korpusz: MNSZ2	
találatok száma: 2	
<a href="#">doc#2539,tudományos</a>	Vele néhány Mario játékban lehet <b>játszani</b> .
<a href="#">doc#2539,tudományos</a>	Mario a hátán <b>szok</b> rajta <b>lovagolni</b> .

Mivel ránézésre úgy tűnik, hogy két különböző találatról van szó, az elsődleges tisztítást a konkordancia valószínűleg „megússza”, a találatok száma 2 marad.

Az első találatban („mondat” megjelenítésnél) nem látjuk a *szok* szót, ezért mindenképpen meg kell néznünk a tágabb kontextust ahhoz, hogy kiderüljön, releváns-e a találat:

Toad A hercegnő inasa. Alacsony termetű, kék mellényt, fehér nadrágot visel. Feje fehér alapon piros pöttyös. Neve jelentése: Varangy. Toadette Toad húga. Ő nem jelenik meg a Mario játékokban, éppen ezért nem is tudunk túl sokat róla. Ám azt viszont ennek ellenére tudjuk, hogy nagyon aranyos. Rózsaszín ruhája nem téveszthető össze. Szintén kicsike, pont úgy mint a bátyja. Yoshi Vele néhány Mario játékban lehet játszani. Mario a hátán szok rajta lovagolni. Zöld dinoszaurusz, ki a barna cipője és narancssárgás tüskéi miatt nagyon helyes. Wario Wario az erősebb karakterek közé tartozik. Ruhája sárga-lila, sapkája is szintén sárga és a W betű rajta szintén megtalálható, cipője zöld, orra rózsaszín, ami nem hétköznapi, bajsa is cikcakkos.

Látjuk, hogy **a két főnévi igenevet a program ugyanahhoz a *szok* szóhoz sorolta**, mivel mindkettő öt szó távolságon belül van, csak más irányban. A részletet elolvasva látjuk, hogy a *játszani* főnévi igenevnek semmi köze a *szok* igéhez, a *játszani* az előtte álló *lehet* igével alkot szó szerkezetet.

Ez arra is példa, hogy **ha úgy állítjuk be a kontextust, hogy a két szó nagyobb távolságra is lehet egymástól, növeljük az ilyen irreleváns találatok veszélyét**. Ha két szó távolság lett volna

beállítva, ez a találat nem jelent volna meg. Viszont ha egy szó távolság lett volna beállítva, akkor a releváns *szok lovagolni* se jön elő, mivel a szerkezet két tagja közé ékelődik a *rajta*.

A másodlagos tisztítás során tehát az első, irreleváns találatot töröljük, marad egyetlen releváns találat.

Mivel csak egyetlen találatunk van, harmadlagos tisztítás nem végezhető.

### ***Az adat minőségi elemzése***

Ahhoz, hogy az adatot értékelni tudjuk, meg kell tudnunk, milyen dokuból való. Rákattintva a kék, információs részre ez a kép fogad:

```
doc.file    sci_hu_wiki_025
doc.region  magyarországi
doc.style   tudományos
div.type    article
```

A kódok ismerete nélkül is világos, hogy a találat a Wikipédiából, a *Super Mario Bros* című videojáték leírásából való ([https://hu.wikipedia.org/wiki/Super\\_Mario\\_Bros](https://hu.wikipedia.org/wiki/Super_Mario_Bros)). A szöveg gyermekes stílusából és a helyesírási hibákból jól látszik, hogy szerzője nem „tudós”. Ezt a gyanúkat igazolja a Wikipédia figyelmeztetése: „Ezt a szócikket tartalmilag és formailag is át kellene dolgozni, hogy megfelelő minőségű legyen.”

Ez az átdolgozás részben már meg is történt, mert a Wikipédiában 2018. szeptember 28-án a szövegben ez a forma javítva volt múlt idejűre. A fentebb az MNSZ2-ből idézett szöveg a fel-tüntetett napon így festett a Wikipédiában:

#### **Toad**

A hercegnő inasa. Alacsony termetű, kék mellényt, fehér nadrágot visel. Feje fehér alapon piros pöttyös. Neve jelentése: *Varangy*.

#### **Toadette**

Toad húga. Ő nem jelenik meg a Mario játékokban, éppen ezért nem is tudunk túl sokat róla. Ám azt viszont ennek ellenére tudjuk, hogy nagyon aranyos. Rózsaszín ruhája nem téveszthető össze. Szintén kicsike, pont úgy mint a bátyja.

#### **Yoshi**

Vele néhány Mario játékban lehet játszani. Mario a hátán sokat szokott lovagolni. Zöld dinoszaurusz, barna cipője van és narancssárgás tüskéi miatt nagyon helyes.

#### **Wario**

Wario az erősebb karakterek közé tartozik. Ruhája sárga-lila, sapkája is szintén sárga és a W betű rajta szintén megtalálható, cipője zöld, orra rózsaszín, ami nem hétköznapi, bajsza is cikcakkos.

Természetesen a Wikipédiában bekövetkezett változások nem tükröződnek az MNSZ2-ben, ahova a Wikipédia-anyag egy bizonyos időpontban került bele.

A változtatás figyelemreméltó tanulsága, hogy a szövegben nagyon sok minden szorult volna javításra, ám a szerkesztők csak a Yoshiról szóló részben eszközöltek javításokat. Nyilván a *szok* miatt nyúltak a szócikkhez, de ha már ott voltak, a következő mondatba még belejaví-

tottak, a többit viszont javítatlanul hagyták. Mindez arra utal, hogy a szerkesztők szemében a *szok* ilyen használata szemet szűrő, súlyos hiba volt, amely feltétlenül javításra szorult.

Meg kell még jegyezni, hogy e jegyzet készítésének idején (2023. október 24-én) a szócikkben már ez a javított adat sem szerepel. A játék szereplőiről írt rövid jellemzések egy részét ugyanis a szócikk legújabb változatából teljesen törölték, köztük Yoshit; neve csak felsorolásképpen jelenik meg a szócikk végén. Könnyen lehet, hogy a szerkesztők a szöveg gyöngye minősége miatt döntöttek a törlés mellett.

Teljesen nyilvánvaló, hogy az idézett szöveg **nem képviseli a tudományos stílust**, bátran figyelmen kívül lehet hagyni. Az elemzés után elmondható tehát, hogy az MNSZ2 egyetlen releváns találatot sem tartalmaz erre a formára, s az is valószínű, hogy mindez a tudományos nyelvire általában is igaz.

### ÷ D) A „hivatalos” alkorpusz találatainak vizsgálata

A „hivatalos” alkorpusz találatainak száma is 2.

korpusz: MNSZ2	
találatok száma: 2	
doc#901,hivatalos	És ha elmondhatjuk, hogy mondjuk a lakás betörések száma, amely területen igazából soha nem tudtunk eredményeket elérni, az elmúlt évben mintegy 16 százalékkal csökkent, vagy elmondhatom azt, hogy például a gépkocsilopás vonatkozásában nemzetközi összehasonlításban általában a budapesti példa <b>szok</b> ott <b>előjönni</b> ...
doc#901,hivatalos	Én általában a tényeket <b>szok</b> tam <b>elmondani</b> .

Első ránézésre kiderül, hogy mindkettő elírásból keletkezett, nem a *szok* ige jelen idejéről van szó. Ezért az elsődleges tisztítás után nem marad adat.

Elemzésre nincs szükség, elemzés nélkül is levonható az a következtetés, hogy a vizsgált forma a hivatali nyelvire sem lehet jellemző, ha egyszer a „hivatali” alkorpuszban, egyetlenegy releváns találat sincsen.

### ÷ E) A „beszélt nyelvi” alkorpusz találatainak vizsgálata

A „beszélt nyelvi” alkorpusz a célszóra 5 találatot adott, a következőket:

korpusz: MNSZ2	
találatok száma: 5	
doc#2632,beszéltnyelvi	Persze, meg kölcsön, így adnak nekem, a szomszédom <b>szok</b> nekem <b>adni</b> kétezetet, ezer forintot, hát úgy, hát így az ember megspóroltat, amit tud.
doc#2632,beszéltnyelvi	Házunktól szegődünk, van 7-8-10 ember, van, amikor 15 is, összeszedődünk, <b>szok lenni</b> itt nálam is gyülekezet, és megkezdjük az imával, megkezdődik ígével a Bibliából, szólunk az Isten szavát, ők hallgatnak, én szolgálok nekijé, énekelünk, dicsőítünk az Urat.
doc#2638,beszéltnyelvi	Férfi.: -.. és ott ahol van 7-8-10 ember, van amikor 15 is összeszedődünk, <b>szok lenni</b> itt nálam is gyülekezet és megkezdjük az imával, megkezdődik ígével, a bibliából szólunk az isten szavát, ők hallgatnak én szolgálok nekijek, énekelünk, dicséritjük az urat, olyan 1 órát, 1,5 tart, és akkor ki mindenki megyen békével haza.
doc#2613,beszéltnyelvi	Vannak itt külön felakasztva rácsra ilyen kis függő karikák, csengő van a végébe, és ezen <b>szok tornászni</b> .
doc#2631,beszéltnyelvi	Majustól <b>szok lenni</b> átineuro/?

Első ránézésre feltűnik a 2. és a 3. találat hasonlósága. Ha megnézzük őket (a főnévi igenévre rákattintva) nagyobb kontextusban, az alábbiakat látjuk:

Hát dolgozok ott kint az utcán, sepregetek. Valahogy megélünk! Itt a Jancsi pásztornál nagyon rend van, éltünk a szeretettel. Szeretünk a pástorunkat nagyon! Éltünk a szeretettel.

Most járunk több mint egy éve már ide, megleltünk a helyünket, hála az Úrnak, hiszen szolgálatot. Mv.: - Mit szolgál? Vezető, vider vagyok. Házi csoportom is volt. Mv.: - Az mit jelent, hogy házi csoport? Házunktól szegődünk, van 7-8-10 ember, van, amikor 15 is, összeszedődünk, szok **lenni** itt nálam is gyülekezet, és megkezdjük az imával, megkezdődik igével a Bibliából, szólunk az Isten szavát, ők hallgatnak, én szolgálom nekije, énekelünk, dicsőítünk az Urat. Egy órát, másfelet tart, s akkor ki megyen, ki megyen békével haza. Mv.: - Megváltozott az élet itt a táborba, amióta ez a sok egyház jött? Nagyon. Ó, persze, nagyon is megváltozott, mert itten régebben nem így éltek a romák, ittak, verekedgettek, bajok voltak meg minden, de ahogy bejött az Istennek

- Most járunk több mint egy évi már ide, de nagyon jól érezzük magunkat, megleltük a helyünket, hála az úrnak, ide járunk, viszünk szolgálatot. Rip.: - Maga mit szolgál? Férfi: - Nekem van házi csoportom, vezető, lider vagyok. Rip.: - Lider? Férfi.: - Igen, vezető. Rip.: - Mit jelent, hogy házi csoport? Idejönnek az emberek. Férfi.: -. és ott ahol van 7-8-10 ember, van amikor 15 is összeszedődünk, szok **lenni** itt nálam is gyülekezet és megkezdjük az imával, megkezdődik igével, a bibliából szólunk az isten szavát, ők hallgatnak én szolgálom nekijek, énekelünk, dicséretjük az urat, olyan 1 órát, 1,5 tart, és akkor ki mindenki megyen békével haza. Rip.: - Megváltozott az élet itt a táborban, amióta ez a sok egyház jött? Férfi.: - Persze, nagyon is megváltozott, mert itten régebben nem így éltek a romák, ittak, verekedgettek, bajok voltak, meg minden, de hogy bejött az istennek a szava

A két idézet összevetéséből kiderül, hogy **ugyanannak a szövegnek két eltérő átírata került bele az adatbázisba**. Mind a kettőn jól látszik, hogy az átíró nem ismeri a közeget, és félrehallja az adott közösségre jellemző szavakat: az egyik a *líder* helyett *vidert* ért, a másik *dicsőítjük* helyett *dicséretjük* formát hoz.

Egyéb hibák is vannak, amelyekből levonható a következtetés, **hogy a „beszélt nyelvi”** alkorpuszból származó adatokat (is) **nagy körültekintéssel kell kezelni** (főleg amikor nagyon alacsony a számuk).

Mivel azonos beszélgetés két átíratáról van szó, az elsődleges tisztítás során a találatok számát 5-ről 4-re csökkentjük.

A maradék 4 adat releváns, így másodlagos tisztításra nincs szükség.

Mivel az egyik doku (doc#2632) kétszer fordul elő, az esetleges harmadlagos tisztítás során az adatok számra 4-ről 3-ra csökkenne.

### ***Az adatok minőségi elemzése***

Először is feltűnő, hogy az adatok száma **nagyon alacsony**, pedig leginkább a beszélt nyelvben várhatnánk a *szok* + főnévi igenév szerkezet használatát.

A találatok alacsony számának egyik oka az lehet, hogy rádiós műsorokról van szó. Feltételezhető, hogy azon beszélők közt is, akik ténylegesen használják ezt a szerkezetet, lehetnek olyanok, akik tudják, hogy a szerkezet megbélyegzett, s ezért formálisabb beszédhelyzetben **ügyelnek rá, hogy ne csússzon ki a szájukon**.

Az MNSZ2 munkatársai sajnos **nem közlik** a felhasználóval, hogy a rádiós szövegek átírása **milyen elvek** szerint történt, sőt azt sem tudjuk meg, hogy az átiratok **eleve az MNSZ2 számára készültek-e**, vagy eredetileg valamilyen **más célt** szolgáltak.

Pedig ezeknek a kérdéseknek a tisztázása az adatok értelmezése szempontjából nagyon fontos lenne! Ha például nem eleve az MNSZ2 számára készültek az átiratok, vagy az átírók nem voltak tisztában a szövegek céljával, **előfordulhat**, hogy **maguk az átírók javították munkájuk részeként a „durva nyelvhelyességi hibákat”, s ebből és más hasonlókból ezért van olyan kevés az alkorpuszban.**

Egyáltalán nem mindegy, hogy az alacsony szám annak tudható-e be, hogy az interjúalanyok erős önkontrollt gyakoroltak, vagy annak, hogy az átírók figyelme olykor lankadt, és emiatt itt-ott elfelejtették kijavítani a „durva nyelvhelyességi hibákat”.

Azt a gyanúkat, hogy a szövegekben a pusztán átíráson túl javítások is történtek, az a tény is erősíti, hogy az átiratokban feltűnő a megakadásjelenségek nagyfokú hiánya. Ha az átírók fel voltak jogosítva arra, hogy a megakadásjelenségeket eltüntessék az átiratokból, nem olyan nehéz elképzelni, hogy ugyanazt tették a „durva nyelvhelyességi hibákkal” is.

Az alacsony elemszám további oka lehet, hogy a rádiós műsorokban **nem kerül sor** a helytelennek tartott nyelvi formának a fórumokban és más hasonló műfajokban oly gyakori tudatos, **játékos-ironikus** használatára (ez nagyon erőteljesen jelentkezik az alább említett „személyes” alkorpuszban).

Ha elolvassuk a találatokat a legbővebb kontextussal, megbizonyosodhatunk róla, hogy ezek valóban **riportokból** származnak, vagyis nem a rádiónál dolgozó személyek beszélnek, hanem **egyszerű emberek**, akiket a riporterek megszólaltatnak.

Az egyik példát (kétféle átiratban) már közöltük, úgyhogy csak a maradék hármat nézzük meg:

Ő családtag? Igen, igen, igen. R.: - Gyönyörű szép árnyalatok vannak rajta, erőteljes színek. Az ember ha kiengedi, akkor rászáll a vállára, fejére, s akkor bohóckodásba is belekezd sokszor. A rácsra fel van függesztve, és akkor leesik a földre, alábújik, hömbörgeti, és rugdalja jobbra-balra. Amikor csörög a telefon kérdezzük mindig, hogy ki az. Ő is mondja, hogy ki az? Vannak itt külön felakasztva rácsra ilyen kis függő karikák, csengő van a végébe, és ezen szok tornászni. Van mikor majdnem leszakítja, annyira bele van a játékba bolondulva. Tegnap már a nejem mondta, hogy egy új szót megtanult, szeretlek, adj egy puszit. Én nagyon szeretem, el lehet fogadni társnak is. - Na, nem mondasz semmit? R.: - Gátlásos. Igen, kicsit meg van illetődve. Mivel egyedül van otthon, tehát nincs annyi madár között, mint most itt, szerintem azért nem mondja, mert most a hangokra figyel. Lehet, hogy otthon meg valami hangot ki fog mondani.

Hát hogyha nyár van csak, kökényt, csipkét meg a gomba, ennyi, átveszik a gombát, oszt ' akkor akármennyit szedjünk, akkor is jó, mert még ha munka sincs, ugye egy hónapba nem tudnak még 5-6 napot se dolgozni, azt a csepp szoci! Hát kapjuk a családít, rá kell adni villanyra, rá kell adni a vízre! Mv.: - És ez a szagos gomba ez most már kellene, hogy legyen? Hát mán kellene, de hát nem veszik, csak így mink megfőznénk. Májustól szok lenni átineuro/?/, az az értékesebb. Mv.: - Hogy veszik? Hát átveszik 3-400 forintért, hol ötszázért, hát oszt ' menjünk a boltba! Mv.: - Mennyit tudnak összeszedni, tehát mondjuk hogyha elmennek gombászni, abból maguknak mennyi pénz van? Hát akkor 4-5 kilót szedjünk, egy fő,

a fiaim is külön vannak, én is megszedem azt a négy kilót, azt a hármat, kétezret, hol ezerötöt, hol ezer forintot kapok.

Hát hogyha így megkapjuk a pénzt, bevásárolunk hát a hűtőbe, apródonként ilyen kisebb darabokat, elrakosgatom, azt ' akkor úgy. R.: - Honnan tudnak még pénzt szerezni? Így hogyha elmegyek így a parasztnál dolgozni, azt ' akkor úgy. R.: - Napszámba? Csak így elhínik, hát amit kell, megcsináltatom nekik. R.: - Ja itt a faluba? Persze, meg kölcsön, így adnak nekem, a szomszédom szok nekem adni kétezret, ezer forintot, hát úgy, hát így az ember megspóroltat, amit tud. R.: - Broda Lajosné, Fáj polgármestere, ebben a faluban kiket, hány embert érint az, hogy januártól egy családban egy ember kaphat RÁT-ot, azaz rendelkezésre állási támogatást, hányan veszítettek így jövedelmet, tehát akik emiatt most kevesebb pénzből élnek?

Ezek a példák – alacsony számuk miatt – természetesen nem bizonyítanak semmit, arra viszont jók, hogy megerősítsék eddigi ismereteinket, miszerint a szerkezet használata az **egyszerű, iskolázatlan emberekre** lehet jellemző.

### ÷ F) A „személyes” alkorpusz találatainak vizsgálata

Utolsónak a „személyes” alkorpusz adatait nézzük meg, az egyszerűség kedvéért most egybevéve ennek fórumos és közösségi részét.

Annak érdekében, hogy megkapjuk mind a két alkorpuszra vonatkozó adatokat, a lekérdezés lefuttatása előtt ki kell pipálni a mindkét alkorpusz melletti jelölőnégyzetet, így:

Viszajelzés  
Próbálja ki a Mazsola eszközt igék és argumentumaik közvetlen vizsgálatához!

doc.style

- beszéltnyelvi
- hivatalos
- sajtó
- személyes-fórum
- személyes-közösségi
- szépirodalom
- tudományos

Mindent

div.column

div.type

Konkordancia készítése Törlés

A keresés 413 találatot ad, tisztítás nélkül. A tisztításhoz nem érdemes 350-es véletlen mintát kérni, mert az ezzel kapcsolatos számítások (a 350-es mintán nyert eredményeknek a 413-as teljes adatszámra való kivetítése) több időt vennének el, mint amennyi idő alatt a 350 fölötti 63 találat átnézhető. Ezért a teljes anyag, mind a 413 találat átnézésével szűrjük ki az ismétlődő dokukat, miután a bal oldali kontextus szerint rendeztük a találatokat.



Az alábbi kép az első néhány találatot mutatja:

korpusz: MNSZ2	
találatok száma: 413	
doc#974,személyes-fórum	Ő alá <b>szok ' írni</b> , hogy Soltika+ Ficikutyá-aszörösgyerek (.cvnvnjfv88888+ -) aköbön. éljensáprlis.. semmintnemértó... délenkékazég...!!!
doc#987,személyes-fórum	Ettől még lehetne jó, magam már régóta szeretnék befejezni egy kisregényt, nem ez a stílusom, de ílyet is <b>szok ' írni</b> .
doc#974,személyes-fórum	Az a fránya membrán is ki <b>szok ' repedni</b> !
doc#973,személyes-fórum	Nem, csak a végén mindig késelés <b>szok ' lenni</b> .
doc#2699,személyes-közösségi	A padlástéri tűz nem kis eséllyel áramlopás miatt <b>szok ' keletkezni</b> amúgy, de persze a párástok vegyének mindenből újat nekik.
doc#2775,személyes-közösségi	A probléma ott <b>szok ' lenni</b> , hogy ha mégsem tudnánk, kik a hutuk/tuszik, akkor mindig kitermelődnek - helyesebben kitermelünk magunknak - olyanokat, akik megmondják nekünk, hogy kik a hutuk/tuszik.
doc#2775,személyes-közösségi	A probléma ott <b>szok ' lenni</b> , hogy ha mégsem tudnánk, kik a hutuk/tuszik, akkor mindig kitermelődnek - helyesebben kitermelünk magunknak - olyanokat, akik megmondják nekünk, hogy kik a hutuk/tuszik.
doc#999,személyes-fórum	Én mennék, ha odaérek, remélem kitartotok, mer ' lapzártázom, és az sokáig <b>szok ' tartani</b> :-)
doc#1012,személyes-fórum	A kisebbik ma hatalmasat alakított: feleségem hatalmas pocakjával az ágyon <b>szok ' heverni</b> , és gyönyörködik, hogy a pici hogy mozog benne, rá-ráteszi a kezét.
doc#1005,személyes-fórum	Szerintem a Csülök most valami alkotóháza ' csücsül... és isz, ha van esze ( <b>lenni szok</b> neki)... de ha rólam is ír az írói munkásságába bele, hogy a felügyelő úr aztán mindenfélét kérdezzessen, akkor rágurítok egy hordót!
doc#1005,személyes-fórum	Ugrálni nem ugrálok, <b>úszni</b> a hógyerek <b>szok</b> , én csak hazacipelem.
doc#993,személyes-fórum	<b>Énekelni</b> magyarul nem <b>szok</b> , mer ' erős az akcentusa...
doc#986,személyes-fórum	demeter aszondta ő <b>szok</b> 100-al <b>húzni</b> felfelé az érdi emelkedőn... melleleg már 95 is sok:))
doc#2732,személyes-közösségi	Ahogy <b>mondani szok</b> ....
doc#976,személyes-fórum	Na, visszajöttem a Pincéből, ahol egy kicsit pihentem, meg ittam is, ahogy <b>lenni szok</b> , de most sem értek egy mukkot sem.
doc#1008,személyes-fórum	Ha Söröshordó, akkor benne vagyok, mer'hogy megszomjaztam, ahogy <b>lenni szok</b> nálam...
doc#974,személyes-fórum	Ezáltal teljesen olyan lesz minden, ahogy <b>lenni szok</b> , senki sem tud és nem ért semmit, én pedig rendesen be tudok rúgni:))) Na, áll az ajánlat?
doc#978,személyes-fórum	Mint minden más, ez is majd helyrerázódik, izé, helyreáll, vagy helyredől, na, ahogy <b>lenni szok</b> ....
doc#991,személyes-fórum	Szögedében mindenki Lyól berúgott, oszt ' annyiban maradtunk... - utána a társaság társulatilag támolygott tovább szana is, meg széjel is, továbbá összességében vissza, ide ni, ahogy <b>lenni szok</b> kluboséknál, így aztán csak ezek a marhára bizalmas jelentések maradtak meg... ahogy <b>lenni szok</b> ...
doc#979,személyes-fórum	Az euroastra csak a magyar ferdtést közli, azt is lopták valahonnan, ahogy <b>lenni szok</b> , de az AltaVistával meg lehet találni az angol nyelvű "szabályokat " is, tessék, lehet válogatni: Netiquet 1, Internet Etiquette... van ott 666 találat, már ebből is látni, hogy Number of the Beast:)))))
doc#991,személyes-fórum	Szögedében mindenki Lyól berúgott, oszt ' annyiban maradtunk... - utána a társaság társulatilag támolygott tovább szana is, meg széjel is, továbbá összességében vissza, ide ni, ahogy <b>lenni szok</b> kluboséknál, így aztán csak ezek a marhára bizalmas jelentések maradtak meg... ahogy <b>lenni szok</b> ...
doc#1012,személyes-fórum	Na, pedig haza akartam <b>menni</b> .
doc#2916,személyes-közösségi	A ház nem <b>szok</b> 1 nap alatt <b>leamortizálódni</b> , talán, ha kijártak volna évekkal ezelőtt és segítettek volna a házat karbantartani.
doc#2946,személyes-közösségi	Ha mégis kedvem szottyanna valamiért alkoholt <b>inni</b> (ahogy sokan másoknak <b>szok</b> szottyanni), akkor bárban amúgy sem tenném, mert 1: nem járok bárokba 2: ott már most sem lehet rágyújtani mellé 3: díszkóba sem járok 4: ott sem lehet már rágyújtani mellé 4: kimondottan jól esik a hír, hogy csak otthon lehet piázni (hogy előadásokon is? akkor iratkozzunk be egyetemre, ott is előadnak, nem?:), és azt is csak úgy, ha másokat nem zavarunk vele..
doc#2752,személyes-közösségi	Az a lényeg, ne szaporodjanak ivarosan, ami <b>teljesülni szok</b> .
doc#974,személyes-fórum	Na, má ' megin ' televagyok ezzel a duplántriplánlátással, meg olyan deliriumom is van, ami <b>lenni szok</b> , csupa ördög szaladgál itt körülöttem, még szerencse, hogy nem az én korszóm után kapkodnak:))
doc#991,személyes-fórum	Nem pocsek az a cikk, csak olyan, mint amilyen <b>lenni szok</b> mindig, ha internetről van szó.
doc#979,személyes-fórum	Amivel <b>pisilni szok</b> !
doc#990,személyes-fórum	Viszont a wizard várban nem nagyon <b>szok</b> animate-dead <b>lenni</b> (csak resurrection).
doc#1013,személyes-fórum	Sajna nyáron a tengerparton biza fény van meg kékség, spec ez engem nem <b>szok</b> annyira <b>zavarni</b> , söt...))) puzsol: stewe

Az elsődleges tisztítás során 102 ismétlődést találtunk, vagyis az elsődleges tisztítás után a „személyes” alkorpuszban 311 találat marad. Az elsődleges tisztítás fontosságát itt jól mutatja, hogy a találatoknak nem kevesebb, mint **egynegyede** bizonyult irrelevánsnak.

A másodlagos tisztítás során 17 találatról derült ki, hogy irreleváns. Nagy részük azért, mert **a szok és a főnévi igenév véletlenül kerültek egymás mellé**, nem alkotnak szerkezetet.

Több esetben a **szok** önmagában releváns találat, mert egy másik ige is található mellette, amellyel ténylegesen szerkezetet alkot, de ez **máshol szerepel a találatok között**, így ez nem ok arra, hogy ezeket a találatokat ne távolítsuk el a másodlagos tisztítás során. Ennek szemléltetéséhez kevés kontextus is elegendő:

Izrael állam, mint USA-csatlós, közel-keleti előretolt bázis, világos, hogy hatalmas katonai és egyéb támogatást kap. Ez szövetségeseknél így **szok** lenni. Kár ebbe **belemagyarázni** a világirányítást. Izrael szomszédpolitikáján meg sokszor tényleg nincs mit mentgetni. De - csak hasonlatként - attól még Kadhafi sem világirányító, hogy antidemokrata.

Szerinted egy tegnapi blogba ma este is a rezsitémát csócsáljuk? 2013-09-23 Spániel@ Knósszosz: ezt Mesterházi fogja **mondani** , **ő szok** ilyen szép combosokat ígérni.

Ott bor <u>szok</u> lenni. Ki kell <u>inni</u> a maradékot? Kell a hely az újnak? Akkor butulni kell sokat...:)
A Báró úr pedig a Cliff Richard valójában, csak épp annyira tanult meg magyarul, amennyire ékezetek nélkül lehet <u>írni</u> . Énekelni magyarul nem <u>szok</u> , mer ' erős az akcentusa... mekklátja
Valótlanságokat írsz, hosszan, vagyolyanokat amiket elfelejtesz jól ismerni. De ettől még az általad felsoroltak nem lesznek igazak Tévhitbe lehet <u>élni</u> , de nem <u>szok</u> mentálisan jó vége lenni a dolognak.
Na, vili, hogy aki eztet a „rendet” már ismeri, az ilyenkor időben arrébb tántorog, vagy kúsz... Utána meg ismét eltelik egy kis idő és noválják re az egészet és akkor megin' lehet nyugodtan <u>inni</u> ... Ez mindig így <u>szok</u> lenni. Ezért nyugi!
Én ilyenkor a Felügyelő urat tudom <u>sajnálni</u> - olyan szomorú <u>szok</u> lenni...:))) Megaszondták neki, hogy találjon, de nincsen mit, mer'hogy a tapasztalt szeszcsempész sztrovacsekok mindenet rég megittak és még beszélni sem tudnak, legfeljebb lehelnek, de attól a Felügyelő úr csak szédül... pláne, ha egy kis hagyma is van benne... vörös...
Az egyik bácsi reggel szokott jönni-ő is csóró nagyon-megveszi a kenyeret, ezt-azt ami aznapra kell a mamának, na ő szokott hozni álrtaában <u>hozni</u> . 1 kis fröcsit <u>szok</u> inni-ilyenkor kettőt.:)

Látjuk, hogy a program az aláhúzott főnévi igenévvvel „hozta össze” a *szok* igét, miközben a mondatban van egy másik főnévi igenév is, amellyel tényleges szerkezetet alkot. Ez a tényleges szerkezet is szerepel a találatok közt (máshol), úgyhogy a fentieket bátran törölhetjük.

Más esetekben a *szok* nem ige, hanem másfajta szórész, az alábbi idézetekben például egy trágár szó eufemisztikus írásmódjának köszönhetően keveredett a *szok* esetei közé:

Ott a 70 éven felüliek lemeszárlását nem tiltotta a törvény. Talán valami hasonlót kelle- ne <u>bevezetni</u> a vén f@ <u>szok</u> esetére is.:)
A György Teichter Ekkora ballabas lövést, még ezek a szerencsétlen balf@ <u>szok</u> se mer- nek <u>meglépni</u> !:)

Máshol a szavakkal való játék folytán állt elő a *szok* mint szórész:

Én csak ráleheltem erre a szoftra (vagy a Kekire:))) és onnantól kezdve, minden összeomlik, nem, összeromlik, ez sem, ja, igen, összeömlik! Na! Most már mindenki csak úgy egybe az egészet fog <u>írni</u> , ahogy én <u>szok</u> . tam-tam ta-tam, vagy ilyesmi. Ne igazsd magad ezért, in- kább igyál 1 Sört! Egészségedre!
---

Néhány esetben kifejezetten elírásról van szó:

Osztt mégse megy rajta valami (példul két program akad). Most mire fogjam? Hogy tudom ezt <u>bizonyítani</u> ? Vagy fizessek <u>szok</u> száz rugót egy „ szakértő „ cégnek, aki majd semmit sem állapít meg?
Ez alapvető szempont minden fon to sabb kapcsolatban. Házasság előtti tanácsadásokon nyersen fel <u>szok</u> tam <u>tenni</u> a kérdést: „Elégge megbízol ebben a férfiban (vagy nőben) ah hoz, hogy a kezébe tedd életed hátralévő részét?”
Szép ábrándként hálózta be Romániát autópályákkal Victor Pon- ta, kár, hogy ez is csak olyanfajta álom, melyből kábultan <u>szok</u> tunk <u>felébredni</u> .

Van olyan eset is, hogy a beszélő szándékát (a szöveg értelmét) nehéz megállapítani, az azonban biztos, hogy nem az általunk vizsgált szerkezetről van szó:

Vah úr Most mi van? - kérdezem ilyenkor. (milyenkor is?:) de ez most busztopik, úgyhogy itt semmi sírhatnak, azt tessék <u>eltenni</u> másnapra én is akkor <u>szok</u> ! beep merre kolbászolsz sofőr
---

uraság? csak nem a szervizeket járod?

A hinduról nekem főleg nem ezek jutnak eszembe, inkább a guru szóról. Persze a néhai Ami-ga számítógép kivételkezelése is Guru Meditation üzenettel szok volt járni, ebből lett a PC Guru magazine.

A 17 irreleváns találat leszámítása után a találatok számra 294-re apadt.

A harmadlagos tisztítás elvégzésétől most eltekintünk.

### **Az adatok minőségi elemzése**

A releváns adatok olvasása során feltűnő, hogy a *szok* + főnévi igenév szerkezet használata nagyon sokszor (legtöbbször?) **stilisztikailag szándékosan jelölt: stílusjáték** eredménye, ill. tudatosan használt **stílus eszköz** a beszélt nyelvben is; továbbá **ironikus** éle is lehet. Néhány példa:

Elvtársak! Mosmá igen rutinossan ki szok lenni írva a nickem a gyengébbek kedviért! 1 kis házikó alaku palatáblára, és ottan lóg a fán! Rá köll gyöjnek az autó típus felösmerése nem erős oldala a tasztaltársaságnak. De még a nyakkendőről is ismerhető fel vagyok ottan a domb oldalon + meg a dús, göndör szemembe lógó hajjam is + ismerhető! + a tattook, + a mokka és kaczagás ami töllünk fel száll. + a hajj süt Ó vass! + a ..... + a ..... + a ..... SÖ TÖ BÖ. Haligejza hali; ~)

Na, itt van ez a borangyalka, ő szok nekem világítani, amikor megyek le a Pincébe piáér', így aztán nem is ér semmi baj. Eztet úgy is mondhatnám, hogy ő az én őrző-védőm, és a bóóót ott van a do-logban, hogy ő nem is isz, így aztán több marad nekem is, másnak is:))) Egészség!

Ne hallgass erre a neduddgira, mer'hogy ő az, csak tagaggya, aztán amikor meg akarják verni, akkor meg kimagyaráz, ahogy mindig szok! Meg is van ez írva a Könyvben! Ottan leveleket szok írni, amiket senki sem ért. Olvass!

Pudli sztrovacsek, te olvasni is szoktál, vagy csak írsz? A másik topikban teljesen világosan (szerintem, és a duplatripláságom ellenére állítom) elbőfögtem, hogy van ugyebár egy Gazda, akinél mi mindannyian vendégek vagyunk. Ez idáig érthető? Nem volt nagyon zavaros a megfogalmazásom, ahogy mindig lenni szok? Na jó, akkor menjünk tovább. A Gazda vigyáz a kocsmájának a tisztaságára és arra, hogy a vendégek ott jól érezzék magukat, ezért néha takarít, sőt, takarítószemélyzetet, kidobóembereket alkalmaz, azaz ő dönt arról, hogy milyen legyen az ő kocsmája. Ez is világos? Esetleg hozzak egy Kőbányait (a címére rá szok lenni írva, hogy világos, csak olvasni meg inni kell)?

Még ott is, ahol nincsenek a szövegekörnyezetben olyan más feltűnő elemek, amelyek stílusjátékra vagy iróniára utalnak, is gyaníthatjuk, hogy erről van szó. Ezért korántsem biztos, hogy például az alábbi találatok közül mindegyik a szerző „normális” nyelvhasználatát tükrözi.

Hakkinen azóta igazán nagymenő, amióta mercedessel megy. ez persze nem azt jelenti, hogy nem tud vezetni, de előtte nem igazán volt említésre méltó eredménye. az igazság az, hogy nekem a „zú-zós” pilóták tetszenek, mint Coulthard, vagy Irvine. ezen a versenyen viszont meglepett a finn srác mert eredményesen versenyzett, de ha megfigyelted, akkor csak a pálya egy bizonyos pontján előzött a kocsjában lévő lóerőfölényre taktikázva. ellentétben mondjuk Irvine-nal, aki ugyan messze nem volt olyan eredményes, de sokkal „fantáziadúsabban” szok versenyezni. üdv: homár

no, az én általam használt dolgok: a sok víz (brrrrrrrrrrrrrrr) tényleg jó. Ha a fejed szok fájni másnap, akkor az utolsó pohár piával egy Quarelin. Meg reggel még egy. Ha a gyomroddal szok baj lenni, (akkor persze egy jó kis hányás a legjobb), akkor minden benzinkútnál lehet kapni egy Samarin nevű

pezsgőport, gyomorhaj ellen rulez (fejfájásra nem jó) ezt is jobb még este bevenni. Másnap reggel sok-sok C-vitamin (nem lehet túladagolni), meg magnézium pezsgőtabletta, mert az kell az idegrendszernek de az alkohol kimossa állítólag. (Lehet, hogy este jobb, de akkor kinek jut eszébe?) De híjja ide valaki pl Mr Spockot, hátha tud valami okosat.

... aztán a szülők végül is autóval mentek. én utánuk bringával, tekertem, mint az állat a mutter biciklijével. ennek olyan áttétele van, hogy izomból a legerősebb fokozatban is max. 25-tel tudok menni, de azt sokáig nem bírom tartani. ezen volt világitás, a faterén a hátsó lámpa nem működik. odaértem leizzadva, kipirulva, meg-szom-jaz-va. az eső rögtön rákezdett, én be a házba, 1 pálinka, 1 sör, minden süti végigkótol. gyorsan 4 óra lett, elrohantunk a templomba ételszentelésre. pinkán délután szok lenni, így a környékeliek odajárnak, akik nem akarnak résztvenni az esti misén. állítólag ez az új, fiatal pap maratoni miséket tart, amire panaszkodnak a hívek hahaha.

Nekem a dohányboltban mondták a „hirt”, azzal, hogy nem sokára megnyílnak a nemzeti italboltok is... dupla áron a pia, és csak az államnak csöngetve... :) Engem mondjuk nem érint – minden pletyka ellenére, én most „féláron” sem tudom, és nem is akarom megvenni – ... Ha mégis kedvem szotytyanna valamiért alkoholt inni (ahogy sokan másoknak szok szottyanni), akkor bárban amúgy sem tenném, mert 1: nem járok bárokba 2: ott már most sem lehet rágyújtani mellé 3: diszkóba sem járok 4: ott sem lehet már rágyújtani mellé 4: kimondottan jól esik a hír, hogy csak otthon lehet piázni (hogy előadásokon is?? akkor iratkozzunk be egyetemre, ott is előadnak, nem?:), és azt is csak úgy, ha másokat nem zavarunk vele..

A források egyébként elég **egyoldalúak**, úgy néz ki, **hogy ugyanaz a néhány fórumozó „felelős” a találatok jelentős részéért**, akik ráadásul egyértelműen **ironikusan** használják a kifejezést. Emiatt hasznos lenne az MNSZ2 adatait a Google segítségével feltárt adatokkal kiegészíteni.

Ezen az egyoldalúságon épp **a harmadlagos tisztítással** lehetne segíteni. Ettől azért tekintünk el mégis, mert az infó szerinti rendezés funkciója az MNSZ2-ben nem működik, s ilyen elemszám mellett szövegfájllal dolgozni nagyon időigényes volna.

## ÷ G) Összefoglalás

Ez a rövid elemzés így is elegendő volt ahhoz, hogy igazolást nyerjen az a tény, hogy a **szok + főnévi igenév kapcsolata egyedül a beszélt nyelvhez közel álló „személyes” alkorpuszra jellemző**. Hogy azon belül pontosan milyen stílusértékű, azt a „személyes” alkorpusz találatainak, azok kontextusának stilisztikai elemzése mutatná meg.

Ha a fenti adatokat táblázatba foglaljuk, ezt kapjuk:

<i>alkorpusz</i>	<i>találatok száma tisztítás nélkül</i>	<i>találatok száma az elsődleges tisztítás után</i>	<i>találatok száma a másodlagos tisztítás után</i>	<i>találatok száma a harmadlagos tisztítás után</i>
szépirodalmi	23	21	20	17
sajtónyelvi	8	8	7	3
tudományos	2	2	1	–
hivatalos	2	2	0	–
beszélt nyelvi	5	4	4	3
személyes 1, 2	413	311	294	?
<b>összesen</b>	<b>453</b>	<b>348</b>	<b>326</b>	<b>?</b>

Az ilyen táblázatos forma arra csábít, hogy az egyes alkorpuszokra vonatkozó számadatokat összehasonlítsuk egymással, ám ennek a kísértésnek mindenképpen ellen kell állni. Igaz, az olyan hatalmas különbségek esetében, mint amilyenek itt vannak a „személyes” alkorpusz és az összes többi közt, ránézésre is meg lehet állapítani (ahogy fentebb már meg is állapítottuk), hogy a vizsgált szerkezet használata a „személyes” alkorpuszra jellemző, más esetekben, amikor a gyakorisági arányok nem ilyen radikálisan térnek el egymástól, a számadatok közvetlen összehasonlítása nagyon **félrevezető** volna, mivel **az egyes alkorpuszok mérete nem azonos**.

Képzeljünk el egy analóg példát! Egy egyetemi csoportba 10 hallgató jár, ebből kettő férfi, nyolc pedig nő. A kutató meg akarja vizsgálni, kikre jellemzőbb az akciófilmek kedvelése, a férfiakra vagy a nőkre. A kutatás során kiderül, hogy a csoportban mindkét férfi kedveli az akciófilmeket, a nyolc nő közül pedig kettő kedveli. Vagyis 2 nő és 2 férfi kedveli a csoportban az akciófilmeket. Mégsem vonná le ebből egy épeszű kutató sem azt a következtetést, hogy a csoportban a férfiakra és a nőkre egyformán jellemző az akciófilmek kedvelése, hiszen a csoportban csak két férfi van, de nyolc nő. A férfiak 100%-a kedveli a csoportban az akciófilmeket, de a nőknek csak 25%-a. Az adatoknál tehát nem az abszolút, hanem az ún. relatív gyakoriság számít.

Ehhez hasonló a helyzet az egyes alkorpuszokkal is, mivelhogy azok mérete jelentősen eltér egymástól.

A minőségi elemzés során minden csoportban kimutattuk, hogy **még a releváns adatok sem bizonyítják azt, hogy a vizsgált szerkezet használatos volna az adott stílusrétegben**. Emiatt, de azért is, mert az elemszám – a „személyes” alkorpuszt kivéve – nagyon alacsony, ebben az esetben **nincs értelme relatív gyakoriságot számolni**. Ezért a relatív gyakoriság kiszámítását – ami feltétele annak, hogy az egyes alkorpuszok adatait össze tudjuk egymással hasonlítani – *A stílusréteg szerinti mennyiségi vizsgálata* című fejezetben mutatjuk be, **más példák**on (a *kaszárnya vs. laktanya*, valamint a *követően vs. után* példáján).

# & STÍLUSRÉTEG SZERINTI MENNYISÉGI VIZSGÁLAT

## ÷ 1. Általános kérdések

A mennyiségi vizsgálat célja, hogy **számadatok segítségével** tudjuk valószínűsíteni a vizsgált nyelvi forma **stílusértékét**. Abból indulunk ki, hogy az MNSZ2 stílusrétegek szerinti alkorpuszai jellemzően **eltérő stílusrétegekhez** tartozó szövegeket tartalmaznak:

1. A „szépirodalmi” alkorpuszról tudjuk, hogy az többnyire szépirodalmi művekből áll, melyeknek tematikája rendkívül sokszínű. Feltételezzük, hogy ezek többségében a **választékos** stílus dominál.

2. A „sajtónyelvi” alkorpusz különféle sajtótermékekből vett szövegeket tartalmaz, melyek feltehetőleg többnyire **kiegyenlített, enyhén formális** vagy **enyhén választékos** stílusúak, emellett kisebb számban **bizalmas stílusú** elemeket, sőt szövegeket is tartalmazhatnak.

3. A „tudományos” alkorpusz javarészt tudományos és szakmai témájú írásokat tartalmaz, melyek valószínűsíthetően **enyhén** vagy **erősebben formális** stílusúak, még ha vendégelemként informális, szlenghez tartozó és beszélt nyelvies elemek is előfordulhatnak bennük.

4. A „hivatalos” alkorpusz olyan szövegeket tartalmaz, melyekben ugyancsak a **kiegyenlített** és az **enyhén** vagy **erősebben formális** stílus dominál. Mivel a parlamenti felszólalások szóban hangzottak el, ezek mellett itt is számolunk kisebb számban legalább **enyhén bizalmas** stílusú, **beszélt nyelvies** megnyilvánulásokkal is.

5. A „beszélt nyelvi” alkorpuszról tudjuk ugyan, hogy nagy számban tartalmaz írott nyelvi, felolvasott szövegeket, de mivel riportok átiratai is vannak benne, feltételezhetjük, hogy akadhatnak bennük a **beszélt nyelvre** jellemző jelenségek is, annak ellenére, hogy a riportalanyok a rádióknak nyilatkozva inkább **kiegyenlített** vagy **enyhén formális** stílusban nyilatkoznak meg.

6. A „személyes-fórum” alkorpuszról feltételezhetjük, hogy zömében **erősen** vagy **enyhén informális**, sokszor **erős érzelmi töltésű** szövegeket tartalmaz.

7. A „személyes-közösségi” alkorpusz valószínűleg **nagyon vegyes** anyagot tartalmaz, mivel a közösségi médiában a legkülönbélebb stílusú szövegek jelennek meg, akár eredeti, akár megosztott szöveggént. Ezért itt egyaránt találkozhatunk **informális, formális, választékos** és **kiegyenlített** stílusú szövegekkel.

A vizsgált nyelvi formára vonatkozó adatokat **alkorpusz szerinti bontásban** nézzük meg, a számadatokból ún. **relatív gyakoriságot** számolunk, majd az egyes alkorpuszokra vonatkozó relatív gyakorisági adatokat összehasonlítjuk egymással. Vagyis megnézzük, hogy az egyes alkorpuszokban mennyire gyakori a célforma, s ismerve az alkorpuszok közötti, főntebb jelzett különbségeket, ebből vonunk le következtetéseket a célszó stílusértékére nézve.

Mivel – amint fenti áttekintésünkéből kiderül – az egyes alkorpuszok anyaga stílus szempontjából korántsem homogén, a legmegbízhatóbb eredményekhez úgy jutunk, ha **a mennyiségi és a minőségi vizsgálatot kombináljuk**.

Ha két vagy több **szinonimát** hasonlítunk össze, egy másik módszert is alkalmazhatunk: **összeadhatjuk** a két vagy több szó **előfordulásait** egy-egy alkorpuszban és a korpusz egészében is, majd kiszámoljuk, hogy ebből **hány százalékot tesz ki az egyik, hányat másik, ill. többedik forma**.

Ez a módszer azonban csak akkor alkalmazható megbízhatóan, ha a két szó nemcsak valamely jelentésében szinonima, hanem vagy eleve **mindkettő monoszém** (egyjelentésű), vagy ha többjelentésű, **mindegyik jelentésükben megegyeznek**.

Ha a két szónak **azonos és eltérő jelentése is van**, csak olyankor alkalmazhatjuk ezt a módszert, ha a találatok száma nem túl magas, és képesek vagyunk minden találat kontextusát elolvasva megállapítani, milyen jelentésben használatos az adott szó, és a statisztikai vizsgálatot csak az azonos jelentésű találatokra korlátozzuk.

## ÷ 2. A kaszárnya és a laktanya példája

Az értelmező kéziszótár „kaszárnya” szócikkében szinonimikus értelmezést találunk (‘laktanya’), a „laktanya” szócikkben pedig egyetlen jelentésként szerepel a következő értelmezés: ‘katonaságnak (vagy más egyenruhás alakulatnak) a lakóhelyéül szolgáló, zárt épület(csoport)’. Ezekből levonhatjuk azt a következtetést, a két szó denotatív jelentése teljesen azonos, legalábbis a szótárírók megítélése szerint. Ez lehetővé teszi, hogy a föntebb említett mindkét vizsgálatot elvégezzük.

A *kaszárnya* stílusminősítése az értelmező kéziszótárban „kiveszőben”, a *laktanyáé* pedig „hivatalos”. A „kiveszőben” nem vonatkozik a szlovákiai magyar használatra, a szlovákiai magyarban ugyanis a szótár szerint a *kaszárnya* általános használatú (vagyis alapvetően közömbös stílusértékű).

A *laktanya* szó „hivatalos” minősítése némileg megtévesztő, hiszen a fogalom megnevezésére nem áll rendelkezésünkre más, közömbös stílusértékű szinonima. Ebből következően a *laktanya* kell használatos legyen nem hivatalos (nem szaknyelvi) kontextusban is. Ezért a korrekt minősítés „hivatalos is” volna.

A *kaszárnya* és a *laktanya* gyakoriságának vizsgálatához az MNSZ2-ben a „részletes keresés” opciót választjuk, s beállítjuk, hogy „szótőre” keresünk rá (a *kaszárnya*, ill. a *laktanya* lexémák előfordulására vagyunk kíváncsiak, ezért minden ragos és jeles alak ugyanolyan releváns találat számunkra, mint a toldaléktalan forma).

**A számításokat a saját alkorpuszunkon végezzük, tehát a vizsgálatból kihagyjuk a „személyes-közösségi” alkorpuszt.**

Alább a *kaszárnya* szóra vonatkozó beállításokat látjuk. A *laktanya*ra vonatkozó beállítások ugyanilyenek voltak.

A *kaszárnya* szótő az egész korpuszban 1545, a *laktanya* szótő 8183 találatot adott. Mi azonban – ahogy jeleztük – nem az egész korpuszt fogjuk vizsgálni, hanem a saját alkorpuszunkat, amely nem tartalmazza az MNSZ2 „személyes-közösségi” alkorpuszát. Enélkül a *kaszárnya* szótő 1444, a *laktanya* szótő pedig 7474 találatot ad.

Ha ezek után megnézzük a találatok számát az egyes alkorpuszokban (a főntebb a *szok + főnévi igenév* vizsgálata kapcsán leírt módszerrel), ezeket az eredményeket kapjuk:

	<i>stílusréteg</i>	<i>alkorpuszok mérete</i>	<i>találatok száma tisztítás nélkül</i>	
			<i>kaszárnya</i>	<i>laktanya</i>
1.	szépirodalmi	101,7	723	1055
2.	sajtónyelvi	430,3	338	3732
3.	tudományos	144,1	253	1140
4.	hivatalos	118,3	20	554
5.	beszélt nyelvi	92,5	84	616
6.	személyes-fórum	57,7	26	377
Σ	<b>MNSZ2 közösségi nélkül</b>	<b>944,6</b>	<b>1444</b>	<b>7474</b>

Ezeket az adatokat fogjuk használni egyrészt arra, hogy kiszámoljuk a *kaszárnya* és a *laktanya* relatív előfordulási gyakoriságát az egyes alkorpuszokban, másrészt arra, hogy megnézzük, egy-egy alkorpuszon belül milyen a *kaszárnya* és a *laktanya* részaránya, a kettő összegét tekintve 100%-nak.

#### ÷ A) Az előfordulás relatív gyakorisága az egyes alkorpuszokban

Amint főntebb már elmondtuk, a **fenti táblázatban olvasható az adatok nem alkalmasak arra, hogy az egyes alkorpuszokra vonatkozóan összehasonlítsuk őket egymással**, mivel az egyes alkorpuszok nem egyforma méretűek, hanem nagyon is nagy különbségek vannak köztük: a legnagyobb, a „sajtónyelvi” alkorpusz hét és félszer több szövegszót tartalmaz, mint a legkisebb, a „személyes-fórum” alkorpusz.

Ebből következően a táblázatban feltüntetett adatokból **nem állapítható meg, hogy a vizsgált szavak mely alkorpuszban a leggyakoribbak, melyekben kevésbé gyakoriak, me-**



**lyekben a legritkébbak.** Ehhez ki kell számolni az ún. **relatív gyakoriságot**, vagyis azt, hogy az egyes alkorpuszokban **1 millió szövegszóra hány találat esik.**

A relatív gyakoriság kiszámításához a következő képletet használhatjuk:

$$r = (t \times 1\,000\,000) : a,$$

ahol:

**r:** relatív gyakoriság

**t:** a találatok száma

**a:** az alkorpusz mérete.

Tehát a találatok számát besorozzuk 1 millióval, majd elosztjuk az alkorpuszba tartozó szövegszavak számával.

Mivel ilyen nagy számokkal nehézkes dolgozni, kényelmesebb a számítás, ha inkább az alkorpusz méretét osztjuk el 1 millióval, azaz e szerint a képlet szerint járunk el:

$$r = t : (a : 1\,000\,000)$$

Ez azért is előnyös így, mert az alkorpuszok mérete a különböző szavak vizsgálata során nem változik, így ezt az osztást csak egyszer kell elvégeznünk, míg ha a találatok számát szorozzuk, azt minden egyes keresésnél újra meg újra el kell végezni.

Az egyes alkorpuszok méretét az alábbi táblázat tartalmazza. A táblázatban azokat az adatokat közöljük, amelyek az „új alkorpusz készítése” opció választásával válnak láthatóvá a keresőfelületen:

alkorpusz	az alkorpuszok tényleges mérete (a)		az alkorpuszok egymilliomod része (m)	
	pontosan	százezer szóra kerekítve	pontosan	százezer szóra kerekítve
szépirodalmi	101 726 086	101,7 millió	101,726086	101,7
sajtónyelvi	430 303 070	430,3 millió	430,303070	430,3
tudományos	144 128 464	144,1 millió	144,128464	144,1
hivatalos	118 311 718	118,3 millió	118,311718	118,3
beszélt nyelvi	92 537 513	92,5 millió	92,537513	92,5
személyes-fórum	57 735 231	57,7 millió	57,735231	57,7
<b>összesen</b>	<b>944 742 082</b>	<b>944,7 millió*</b>	<b>944,742082</b>	<b>944,7</b>

\*Összeadással 944,6 millió jön ki a kerekítés miatt.

Ha az alkorpusz méretének egymilliomod részét elnevezzük m-nek, a képlet ilyen egyszerű lesz:

$$r = t : m$$

Tehát a találatok számát a „szépirodalmi” alkorpusz esetében 101,7-tel kell elosztani, a „sajtónyelvi” esetében 430,3-mal, a „tudományos” esetében 144,1-gyel, a „hivatalos” esetében 118,3-mal, a „beszélt nyelvi” esetében 92,5-tel, a „személyes-fórum” esetében 57,7-tel. Ha

az egész anyagban számoljuk ki relatív gyakoriságot, a találatok számát 944,7-tal kell elosztani.

Visszatérve a példánkhoz, a *kaszárnyához* és a *laktanyához*: a relatív gyakoriságot a **tisztítás nélküli adatokból** fogjuk kiszámítani, mivel nincsenek adataink arra nézve, mekkora volna az egyes alkorpuszok mérete, ha nem volnának benne ismétlődő (és más okból irreleváns) találatok. Ezért **az ismétlődő találatok kiiktatásával az irreleváns találatok torzító hatását nem tudnánk eltüntetni, mivel – amint láttuk – a relatív gyakoriság kiszámítása a teljes alkorpusz alapján történik, azt pedig nem tudhatjuk, hogy ott mennyi az irreleváns találat.**

Nyilvánvaló, hogy az is torzítaná az eredményeket, ha két eltérő státuszú adatot használnák a számításhoz: a teljes alkorpuszba kényszerűségből beleszámítanánk az irreleváns találatokat is, a találatok számába viszont nem.

Azzal, hogy a **saját alkorpuszunkon** végezzük a számításokat, azaz kihagyjuk a legtöbb ismétlődést tartalmazó „személyes-közösségi” alkorpuszt, remélhetőleg elérjük, hogy a számítások eredményei csak csekély mértékben fognak eltérni azoktól az eredményektől, amelyeket akkor kapnánk, ha lehetőségünk volna az ismétlődő találatoktól megtisztított adatbázissal dolgozni.

A fenti képlet alkalmazásával a *kaszárnya* és a *laktanya* **relatív gyakoriságára** nézve az egyes alkorpuszokban ilyen eredmények jöttek ki:

	<i>stílusréteg</i>	<i>alkorpusz mérete</i>	<i>kaszárnya</i>		<i>laktanya</i>	
			<i>találatok száma</i>	<i>relatív gyakoriság</i>	<i>találatok száma</i>	<i>relatív gyakoriság</i>
1.	szépirodalmi	101,7	723	7,11	1055	10,37
2.	sajtónyelvi	430,3	338	0,79	3732	8,67
3.	tudományos	144,1	253	1,76	1140	7,91
4.	hivatalos	118,3	20	0,17	554	4,68
5.	beszélt nyelvi	92,5	84	0,91	616	6,66
6.	személyes-fórum	57,7	26	0,45	377	6,53
Σ	<b>MNSZ2 közösségi nélkül</b>	<b>944,7</b>	<b>1444</b>	<b>1,53</b>	<b>7474</b>	<b>7,91</b>

Az eredmények értékeléséhez előjáróban meg kell jegyezni, hogy az alkorpuszonkénti gyakorisági eltéréseknek **nemcsak stílusbeli okai** lehetnek (hogy ti. egy szó jobban beleillik az egyik stílusba, mint a másikba), hanem az is előfordul, hogy **az egyes szavak gyakorisági előfordulását a szó denotatív jelentése is befolyásolja.**

Például aligha hihető, hogy mind a *kaszárnya*, mind a *laktanya* azért volna kiugróan gyakori a „szépirodalmi” alkorpuszban, mert stílusértékénél fogva oda illene bele a legjobban: sokkal inkább az a helyzet, hogy **az irodalmi művek nagyobb számban foglalkoznak olyan témákkal, melyek kapcsán szóba kerülnek a katonák és a laktanyák.** És viszont: a „személyes” alkorpuszban nyilvánvalóan nem azért olyan alacsony a két szó relatív gyakorisága, mert az informálisabb stílusba egyikük se illene bele, hanem azért, mert történetesen **az egyik fórumnak se volt témája mondjuk a katonai élet.**

Másfelől azonban az, hogy a *kaszárnya* előfordulási gyakorisága a „hivatalos” alkorpuszban elenyésző, míg a *laktanyáé* nem teljesen jelentéktelen, nem tartalmi okok-

nak tudható be, hiszen a két szó ugyanazt a tartalmat közvetíti, hanem annak, hogy **a ma hivatalosan használt megnevezés a *laktanya***. Hasonlóképpen a „sajtónyelvi” alkorpuszban a *kaszárnya* azért ritka, a *laktanya* viszont meglehetősen gyakori, mert **a mai stilisztikailag jelöletlen forma a *laktanya***, a *kaszárnya* Magyarországon stilisztikailag jelölt, régies, népies (a határon túli régiókban más a helyzet).

Az, hogy a *kaszárnya* a „tudományos” alkorpuszban relatíve gyakoribb, mint a legtöbb alkorpuszban (a „szépirodalmi” kivételével), valószínűleg annak tudható be, hogy a „tudományos” alkorpuszba (aránytalanul) sok történeti témájú szöveg került bele, s ezekben – érthető módon – gyakran lehet szó hadi eseményekről.

A fentiekből azt a tanulságot vonhatjuk le, hogy **a statisztikai stílusvizsgálat önmagában csak az olyan nyelvi formáknál tekinthető viszonylag megbízhatónak, melyeknek a denotátuma az élet minden területén nagyjából hasonló gyakoriságú**, így az, hogy az adott fogalom milyen gyakran kerül szóba az élet egyes területén, nem okoz az adatokban jelentős torzulásokat. A probléma, hogy erre nézve nincs statisztikánk, tehát csak a **megérzésünkre** támaszkodhatunk, ami szubjektív elemet visz bele az elemzésbe.

Növelheti a megbízhatóságot a **magas elemszám is**, amennyiben abban az alkorpuszban is, amelyben a legkisebb a relatív gyakoriság, **ez a „legkisebb” is elég nagy**. Az, hogy egy stílusrétegben egy szó mennyire megszokott, talán nemcsak attól függ, hogy más stílusrétegekben mi a helyzet, sőt lehet, hogy ez inkább csak színező tényező. Egy szó lehet a hivatali nyelvre jellemző, de ha sokszor kerül szóba a hétköznapi életben is, attól még lehet közömbös stílusértékű.

A szó stílusértékét az is befolyásolja, **van-e szinonimája, s az milyen stílusértékű**. Ha **egyáltalán nincs szinonimája, és nem rendkívül ritka fogalmat jelöl, szükségszerűen közömbös stílusértékű**. Ha **van szinonimája**, akkor a helyzetet az is befolyásolhatja, hogy **a szinonima milyen gyakori, ill. mely stílusrétegekben milyen gyakori**. Erre épp a *kaszárnya* : *laktanya* kettősség jó példa. Érdemes ezért összehasonlítani a két szinonima egymáshoz viszonyított gyakoriságát a fent jelzett másik módszerrel. Erről szól a következő alfejezet.

## **÷ B) A két szó gyakorisági aránya egy-egy alkorpuszon belül**

Láttuk, hogy a relatív gyakoriság alapján történő vizsgálat hátránya, **hogy az előfordulási gyakoriságot befolyásoló nyelvi és a tartalmi tényezők nem választhatók el egymástól**.

Mivel ‘a fegyveres testületek szálláshelye’ jelentésre két közkeletű szinonima is létezik, lehetőség van a tartalmi tényező torzító hatásának kiküszöbölésére. Ehhez egy-egy alkorpuszon belül kell összehasonlítani a *kaszárnya* és a *laktanya* előfordulási gyakoriságát.

Ennél a vizsgálatnál, mivel egy-egy alkorpuszon belül végzünk összehasonlítást, nem szükséges relatív gyakoriságot számolni. Ehelyett alkorpuszonként összeadjuk a két szóra vonatkozó adatokat, s ezt tekintve 100%-nak kiszámítjuk, hogy milyen az egyes szinonimák részaránya.

	alkorpusz	gyakoriság			részarány		
		kaszárnya	laktanya	kaszárnya + laktanya	kaszárnya	laktanya	kaszárnya + laktanya
1.	szépirodalmi	723	1055	1778	40,66	59,34	100,00
2.	sajtónyelvi	338	3732	4070	8,30	91,70	100,00
3.	tudományos	253	1140	1393	18,16	81,84	100,00
4.	hivatalos	20	554	574	3,48	96,52	100,00
5.	beszélt nyelvi	84	616	700	12,00	88,00	100,00
6.	személyes-fórum	26	377	403	6,45	93,55	100,00
Σ	<b>MNSZ2 közösségi nélkül</b>	<b>1444</b>	<b>7474</b>	<b>8918</b>	<b>16,19</b>	<b>83,81</b>	<b>100,00</b>

A táblázatból kiderül, hogy az egész vizsgált korpuszra nézve az összes találatnak mindössze 16,19%-a *kaszárnya*, a többi, 83,81% *laktanya*. Ha a célszavak minden stílusrétegben nagyjából egyformán volnának használatosak, hozzávetőleg ugyanezt a – durvábban kerekítve – 16 : 84 arányt találnánk minden alkorpuszban.

A valóságban alkorpuszonként jelentős különbségeket találunk. Mivel mindkét szónak ugyanaz a jelentése, itt nem lehet szó arról, mint fentebb, hogy ti. a denotátum – a katonaság vagy más fegyveres testület szálláshelye – eltérő gyakorisággal kerülne szóba az élet más-más területein. Itt a különbségeknek pusztán nyelvi oka kell legyen.

Az adatok azt mutatják, hogy a *kaszárnya* ezzel a módszerrel számolva a legnagyobb gyakoriságot a „szépirodalmi” alkorpuszban mutatja, s ezzel párhuzamosan a *laktanya* a legkisebbit. Vagyis annak ellenére, hogy a fogalom tartalmi okokból is ebben az alkorpuszban kerül leggyakrabban szóba, e tényező kiküszöbölésével is azt látjuk, hogy a *kaszárnya* szóval ma relatíve ebben az alkorpuszban találkozhatni a leggyakrabban: a *kaszárnya* gyakorisága itt megközelíti a *laktanyáét* (nagyjából 40 : 60 az arány a *laktanya* javára, amitől jelentősen eltér az egész korpuszban meglévő 16 : 84).

Az eredményhez az is hozzájárulhatott, hogy a „szépirodalmi” alkorpusz kakukktójás az MNSZ2-ben, ugyanis – amint az egyes alkorpuszok jellemzésekor láttuk – a többi alkorpuszban található szövegeknél **jóval régebbieket** is tartalmaz. Elképzelhető, hogy egy fél évszázaddal ezelőtt a *kaszárnya* még gyakoribb volt minden nyelvhasználati szintéren, mint manapság, s a „szépirodalmi” alkorpusz ennek emlékét őrzi.

A „szépirodalmi” alkorpuszon kívül csak a „tudományos” alkorpuszban nagyobb még a *kaszárnya* gyakorisága, mint a 16%-os korpuszarány, s ezzel párhuzamosan ebben az alkorpuszban kisebb még a *laktanya* gyakorisága, mint a 84%-os korpuszarány. Itt azonban a korpuszaránytól való eltérés csekély.

Mivel itt kiküszöböltük a téma gyakoriságát mint befolyásoló tényezőt, arra gondolhatunk leginkább, hogy a tudományos irodalom egy része **olyan korban játszódó hadi eseményekkel foglalkozik**, amikor a szóban forgó létesítménynek még *kaszárnya* volt a neve. A szerzők a történeti forrásokban nyilván leggyakrabban – a reformkornál régebbi forrásokban pedig kizárólag – a *kaszárnya* szóval (vagy annak valamely ritkább szinonimájával) találkoztak, ez pedig oda vezetett, hogy a saját szövegükben is valamivel gyakrabban használták a *kaszárnyát*, mint azt tették volna, ha nem azt olvassák a forrásokban.

Az érdekesség kedvéért nézzük meg a „tudományos” alkorpusz találatainak egy részét, mondatkontextus szerinti megjelenítésben, rendezési kritérium megadása nélkül, az adatok keverésével:

Keresés	<p>korpusz: MNSZ2 találatok száma: 253</p> <p><b>tudományos</b> A 16. században a városrész déli fele helyet adott egy <b>kaszárnának</b>, a nagyobb építkezések azonban csak a 17. században kezdődtek.</p> <p><b>tudományos</b> Az idegenek elszállásolása a városi lakosságra hárult, így ezután több <b>kaszárnát</b> és tiszti lakást is építettek itt.</p> <p><b>tudományos</b> A fiatal házaspár 1875 táján a község felső végén az egykori <b>kaszárnákkal</b> szemben felépítették egy földszintes hattengelyes neoklasszicista kúriát, melyet ikerablakok tagoltak keretekben könyöklőparkánnyal.</p> <p><b>tudományos</b> Ja, lehet, hogy van, a katonai <b>kaszárnák</b>, amik most le vannak rombolva.</p> <p><b>tudományos</b> Előkerültek színes falfestmények is a tiszti házakban, lakosztályokban <b>kaszárnákban</b>.</p> <p><b>tudományos</b> Pilar Ternera a Buendiák egymást követő nemzedékeinek ágyasa; Amarantha, akinek hervadó asszonyisága megigézi a kamaszodó fiúkat: Aureliano és Pilar Ternera fiát, Aureliano José-t, aki nagynénje illatát és kutató újjainak édességét keresi a <b>kaszárnya</b> körül járó, szerelmüket kínálógtató fonnyadt nőknél, és a két likertestvért, Második José Arcadiót és Második Aureliánót, akik addig cserélgetik nevüket, amíg saját maguk is belezavarodnak, s nem tudják, melyikük kicsoda.</p> <p><b>tudományos</b> Várfalak, <b>kaszárnák</b> nedves helyiségeinek kiszáritásával bízták meg.</p> <p><b>tudományos</b> Hatszáz nagykőrösi megrohamozta a pénzügyőrségi <b>kaszárnát</b>, betörte az ajtókat és ablakokat, megverte a pénzügyőröket, leszedte a császári sasokat a hivatali épületekről és beverte a helyben lakó zsidók ablakait.</p> <p><b>tudományos</b> Akkoriban <b>kaszárnák</b> fogták közre.</p> <p><b>tudományos</b> Kívül a <b>kaszárnán</b>, kint a kapunál,</p> <p><b>tudományos</b> 1816-tól a második világháborúig <b>kaszárnaként</b> üzemelt.]</p> <p><b>tudományos</b> BENKÓCZY EMIL, A <b>kaszárnából</b> : Nyr. 38 (1909): 46</p> <p><b>tudományos</b> A Pázmány Péter Katolikus Egyetem Bölcsész- és Társadalomtudományi Kara 1994. szeptember 22-én költözött Pilliscsábara, a volt szovjet (korábban Magyar Királyi) <b>kaszárnya</b> területére.</p> <p><b>tudományos</b> Ez most a jelszó Zsongtőle a <b>kaszárnya</b> ...</p> <p><b>tudományos</b> Az 1760-as években épült az akkor monumentálisnak számító templom, illetve egy <b>kaszárnya</b> kőépülete.</p> <p><b>tudományos</b> Így a "börtön" -könyve legalább annyira a <b>kaszárnya</b>, a kórház, a modern gyár, az internátus, a kollégium etc. elemzését is jelenti.</p> <p><b>tudományos</b> Így a "börtön" -könyve legalább annyira a <b>kaszárnya</b>, a kórház, a modern gyár, az internátus, a kollégium etc. elemzését is jelenti.</p> <p><b>tudományos</b> A kastélytól keletre az előbbiekkal szimmetrikus elrendezésben állt a bábszínház, a tisztviselői lakóházak két tömbje és a <b>kaszárnya</b>.</p> <p><b>tudományos</b> A szerzetesrendek 1784-es megszüntetése után az épületegyüttes a kamara tulajdonába ment át, hadi kórházként, <b>kaszárnaként</b>, raktárként is szolgált, majd az elhanyagolt ingatlant 1911-ben Schmidt Miksa bécsi bűtorgyáros és műgyűjtő vásárolta meg, rendbehozta, egy részét bemutatóteremnek, melléképületeit műhelynek használta, végül 1935-ben végrendeletében múzeum céljára a fővárosra hagyta.</p> <p><b>tudományos</b> A francia megszállás idején 1809 és 1814 között a kolostort <b>kaszárnává</b> alakították, a Santa Maria degli Angeli-templomot eladták egy magánzemélynek aki raktárként használta.</p> <p><b>tudományos</b> <b>Kaszárnya</b> balladák (Barrack Room Ballads, 1892, versek)</p> <p><b>tudományos</b> Az 1838-as földrengés után a szász gazdák még kijavították, de a falak egy részét és a kaputornyot 1870-ben lebontották és anyagát a <b>kaszárnába</b> építették be.</p> <p><b>tudományos</b> A börtön és a <b>kaszárnya</b> e felosztásokra és az atomizálásra törekvésre csak a legtisztabb megnyilvánulási formákat jelentik, de a hivatal, az üzem, a kórház és az iskola alapvetően ugyanezen az elven alapulnak.</p> <p><b>tudományos</b> A plájakók számára építették 1745-50-ben a kétszintes <b>kaszárnát</b>, amelyet később a határőrség használt.</p> <p><b>tudományos</b> Elég furcsa, de a klub harmadosztályban népszerűbb volt, mint az előző szezonokban az élvonalban: rendszeresen több mint 6500 néző volt a <b>kaszárnya</b> után nevezett Achter de Kazerne stadionban.</p> <p><b>tudományos</b> Megvett egy rozoga egyemeletes házat a Zöldfa-utcában (a mostani belvárosi <b>kaszárnya</b> helyén), azt fényesen átalakította, márványoszlopokkal feldíszítette s oda vonult vissza, mint valami augur, fogadva az odaözönlő népet, mely mindig fut a csudás után.</p> <p><b>tudományos</b> KATONAI LÉTESÍTMÉNY F Létesítmény A <b>Kaszárnya</b> Katonai kórház Katonai tábor</p> <p><b>tudományos</b> Bourbon lovassági <b>kaszárnya</b></p>
Mentés	
Megjelenítés	
KWIC/mondat	
Rendezés	
bal   jobb találat	
találat	
info	
keverés	
Véletlen minta	
Szűrés	
Gyakoriságok	
szóalakok	
szótövek	
elemzések	
dokumentumok	
Kollokációk	
Eredeti konk.	
Visszaállítás	
Próbálja ki a Mazzola eszközt ígék és argumentumaik közvetlen vizsgálatához!	

Az összes többi alkorpuszban a **kaszárnya** használati aránya kisebb a korpuszaránynál, a **lak-tanyáé** pedig nagyobb. Ez azt jelenti, hogy helytálló lehet az értelmező kéziszótár „kiveszőben” minősítése, legalábbis a magyarországi magyar köznyelvi nyelvváltozatokra nézve.

Azt, hogy a népnyelv milyen mértékben őrzi még a **kaszárnát** Magyarországon, nem tudjuk, mert az MNSZ2 köznyelvi szövegtörzs.

## & SZAKIRODALOM

- Ács Judit–Borbély Gábor–Makrai Márton–Nemeskey Dávid–Recski Gábor–Kornai András 2018. Hibrid nyelvtechnológiák. *Magyar Tudomány* 179/7, 955–963.
- Andor József 2016. A nyelvi szövegek koherenciájának konstitúciójáról. Dobi Edit–Andor József szerk., *Tanulmányok a szövegkoherenciáról*. Debrecen: Debreceni Egyetem Magyar Nyelvtudományi Tanszék. 21–45.
- Andor József 2017. Gondolatok a szövegtípus és a szövegműfaj fogalmáról és egymáshoz való viszonyáról. (Hozzászólás.) *Magyar Nyelv* 113/3, 302–310.
- Antal Zsófia 2021. A *KorSzak* tanulói korpusz bemutatása és a *magyar* nemzetiséget jelölő melléknév vizsgálata. Dóla Mónika–Nádor Orsolya–Szűcs Tibor szerk., *Hungarológiai Évkönyv* 22. Pécs: Pécsi Tudományegyetem Bölcsész- és Társadalomtudományi Kara. 7–21.
- Babarczy Anna 2019. Gyermeknyelvi korpuszok és erőforrások. Sulyok Hedvig–Juhász Valéria–Erdei Tamás szerk., *Beszéd- és nyelvelemző szoftverek a versenyképességért és az esélyegyenlőségért. Hun CLARIN korpuszok és nyelvtechnológiai eszközök a bölcsész- és társadalomtudományokban*. Szeged: Szegedi Tudományegyetem Juhász Gyula Pedagógusképző Kar Magyar és Alkalmazott Nyelvészeti Tanszék. 64–75.
- Bajzát Tímea Borbála 2022. A (pre)modális tartományokkal összefüggő főnévi igeneves mintázatok klaszterezésének lehetősége a történetiség perspektívájában. Forgács Tamás–Németh Miklós–Sinkovics Balázs szerk., *A nyelvtörténeti kutatások újabb eredményei XI*. Szeged: Szegedi Tudományegyetem BTK Magyar Nyelvészeti Tanszék. 25–49.
- Bakaja Zoltán 2020. Az indiai vallási irodalomban használatos beszélőjelölő igéjének magyar fordítása. Robin Edina–Seidl-Péché Olívia szerk., *Fókuszban a fordított és a tolmácsoló szöveg: korpuszalapú fordítás Magyarországon*. Budapest: ELTE BTK Fordítástudományi Doktori Program–MANYE Fordítástudományi Szakosztály. 229–260.
- Balaskó Mária 2020. Korpuszalapú fordítástudomány, és ami mögötte van. Robin Edina–Seidl-Péché Olívia szerk., *Fókuszban a fordított és a tolmácsoló szöveg: korpuszalapú fordítás Magyarországon*. Budapest: ELTE BTK Fordítástudományi Doktori Program–MANYE Fordítástudományi Szakosztály. 55–77.
- Bartha Csilla–Holecz Margit 2017. Jelnyelvi-élőnyelvi szótárépítés a 21. században : A korpusznyelvészet új perspektívái. Benő Attila–Fazakas Noémi szerk., *Élőnyelvi kutatások és a dialektológia*. Kolozsvár: Erdélyi Múzeum Egyesület. 213–225.
- Bartha Csilla–Varjasi Szabolcs–Holecz Margit 2016. A magyar jelnyelvi korpusz létrehozásának és annotálásának kihívásai. Tanács Attila–Varga Viktor–Vincze Veronika szerk., *XII. Magyar Számítógépes Nyelvészeti Konferencia*. Szegedi Tudományegyetem TTIK, Informatikai Intézet. 207–219.
- Bodó Csanád–Kocsis Zsuzsanna–Varga Fruzsina Sára 2017. A Budapesti Egyetemi Kollégiumi Korpusz: elméleti és módszertani kérdések. Benő Attila–Fazakas Noémi szerk., *Élőnyelvi kutatások és a dialektológia*. Kolozsvár: Erdélyi Múzeum Egyesület. 169–177.
- Bolonyai Flóra–Sebők Miklós 2020. Kvantitatív szövegelemzés és szövegbányászat. Jakab András–Sebők Miklós szerk., *Empirikus jogi kutatások: Paradigmák, módszertan, alkalmazási területek*. Budapest: Osiris Kiadó. 361–380.
- Bottyán Gergely 2007. Főnévi igeneves szerkezetek a magyarban. Egy kutatási téma helye a (számítógépes) nyelvészetben. Balogh Margit szerk., *Diszciplínák határain innen és túl*. Budapest: MTA Társadalomkutató Központ. 219–228.

- Bottyán Gergely–Sass Bálint 2005. Személyragos főnévi igeneves bővítményt megengedő predikátumok kinyerése a Magyar Nemzeti Szövegtárból. Alexin Zoltán–Csendes Dóra szerk., *III. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged: Szegedi Tudományegyetem Informatikai Tanszékcsoport. 265–266.
- Csányi Eszter 2014. A gazdasági szleng használati sajátosságai, egy *pilot study* bemutatása. *Economica* 7/4, 115–124.
- Csenges Dóra–Hatvani Csaba–Alexin Zoltán–Csirik János–Gyimóthy Tibor–Prószéky Gábor–Várad Tamás 2003. Kézzel annotált magyar nyelvi korpusz: a Szeged Korpusz. Alexin Zoltán–Csenges Dóra szerk., *I. Magyar Számítógépes Nyelvészeti Konferencia előadásai*. Szeged: Szegedi Tudományegyetem Informatikai Tanszékcsoport. 238–245.
- Csiszár Orsolya 2008. A kontextus hatása a kollokációk megértésére. Várad Tamás szerk., *II. Alkalmazott Nyelvészeti Doktorandusz Konferencia*. Budapest: MTA Nyelvtudományi Intézet. 8–16.
- Csorba Gábrriel 2010. Funkcióigés szerkezetek fordítása a szemantikai prozódia tükrében. 15–30. Várad Tamás szerk., *IV. Alkalmazott Nyelvészeti Doktorandusz Konferencia*. Budapest: MTA Nyelvtudományi Intézet. 15–30.
- Csúry István 2022. Az empirikus koherenciavizsgálat alapjai. Bevezetés a korrelációs korpuszelemzésbe. Dobi Edit–Boda István Károly szerk., *A szövegkoherencia elméleti és gyakorlati megközelítései*. Debrecen: Debreceni Egyetem Magyar Nyelvtudományi Tanszék. 154–168.
- Dankó Szilvia 2020. *Szógyakoróság korpuszalapú vizsgálata: autentikus és célnyelvi magyar szórakoztató irodalmi szövegek összevetése*. Robin Edina–Seidl-Pécs Olívia szerk., *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordítás Magyarországon*. Budapest: ELTE BTK Fordítástudományi Doktori Program–MANYE Fordítástudományi Szakosztály. 177–198.
- Dankó Szilvia–Csizér Kata 2022. Statisztikai próbák a fordítástudományi kutatásokban. Klaudy Kinga–Robin Edina–Seidl-Pécs Olívia szerk., *Bevezetés a fordítás és a tolmácsolás kutatómódszertanába I. Általános rész*. Budapest: ELTE FTT–MANYE Fordítástudományi Szakosztály. 59–76.
- Dér Csilla Ilona 2022. *Háthogy esetleg netalántán azt gondolnátok rólam... A szóval és a hát diskurzusjelölővel álló hogy kötőszós mellékmondatokról. Jelentés és nyelvhasználat* 9/1, 43–57.
- Dér Csilla Ilona 2022. *Ó, hogy azt csak a' szerelmek 's a' lány szellők hajtsák!* Az indulatszókát követő, illetve a nélkülük álló *hogy* kötőszós inszubordinált mondatok megjelenéséről. Forgács Tamás–Németh Miklós–Sinkovics Balázs szerk., *A nyelvtörténeti kutatások újabb eredményei XI*. Szeged: Szegedi Tudományegyetem BTK Magyar Nyelvészeti Tanszék. 45–57.
- Domonkosi Ágnes–Imrényi András–Jánk István–Kalcsó Gyula–Ludányi Zsófia–Rási Szilvia–Takács Judit.H. Tomesz Tímea 2022. Problémaalapú anyanyelvi nevelés – használatalapú gyakorlatok. Jánk István–H. Tomesz Tímea–Domonkosi Ágnes szerk., *A digitális oktatás nyelvi dimenziói*. Eger: Líceum Kiadó. 13–34.
- Dodé Réka–Ludányi Zsófia–Falyuna Nóra–Kuna Ágnes 2018. Poétika és korpusz. Hogyan nyújthat segítséget a korpusznyelvészet poétikus szövegek vizsgálatához? Domonkosi Ágnes–Simon Gábor szerk., *Nyelv, poétika, kogníció: Elmélet és módszer a poétikai kutatásban*. Eger: Líceum Kiadó. 175–196.

- Dóla Mónika–Viszket Anita–Kleiber Judit 2017. A határozott névelő a birtokos szerkezetben. Dóla Mónika–Nádor Orsolya–Szűcs Tibor szerk., *Hungarológiai Évkönyv* 18/1. Pécs: Pécsi Tudományegyetem Bölcsészettudományi Kara. 38–69.
- 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.
- Dömötör Andrea–Yang Zijian Győző–Novák Attila 2020. Nesze semmi, fogd meg jól! Zéró kopulák automatikus felismerése neurális gépi fordítással. Berend Gábor–Gosztolya Gábor–Vincze Veronika szerk. XVI. *Magyar Számítógépes Nyelvészeti Konferencia*. Szegedi Tudományegyetem TTIK, Informatikai Intézet. 385–398.
- Endrédi István–Prószéky Gábor 2016. A Pázmány Korpusz. *Nyelvtudományi Közlemények* 112: 191–205.
- Farkasné Puklus Márta 2018. Nyelv és jog – miskolci kutatások. Robin Edina–Zachar Viktor szerk., *Fordítástudomány ma és holnap*. Budapest: L’Harmattan. 57–70.
- Fóris Ágota 2008. *Kutatásról nyelvészeknek. Bevezetés a tudományos kutatás módszertanába*. Budapest: Nemzeti Tankönyvkiadó Zrt.
- Fóris Ágota 2018. Fordítás és terminológia: a terminológia szerepe a fordítási folyamatban. Robin Edina–Zachar Viktor szerk., *Fordítástudomány ma és holnap*. Budapest: L’Harmattan. 71–90.
- Fóris Ágota 2021. Alkalmazott nyelvészet, korpuszok és adatbázisok. A nyelvtechnológia és a korpusznyelvészet a terminológiaoktatásban. Dodé Réka–Ludányi Zsófia szerk., *A korpusznyelvészettől a neurális hálókig. Köszöntő kötet Váradi Tamás 70. születésnapjára*. Budapest: Nyelvtudományi Kutatóközpont. 91–101.
- Götz Andrea 2020. Diskurzusjelölők és kötőelemek a magyar szinkrontolmácsolt és eredeti diskurzusban: feltáró korpuszkutatás. Robin Edina–Seidl-Pécs Olívia szerk., *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordítás Magyarországon*. Budapest: ELTE BTK Fordítástudományi Doktori Program–MANYE Fordítástudományi Szakosztály. 329–350.
- Gráczy Tekla–Krepsz Valéria–Huszár Anna 2021. Egyéni variabilitás a beszédben. Markó Alexandra szerk., *Tanulmányok a beszédtudomány alkalmazásainak köréből*. Budapest: ELTE Eötvös Kiadó. 201–227.
- Haase Zsófia 2016. A komplex anaforák koherenciateremtő szerepéről. Dobi Edit–Andor József szerk., *Tanulmányok a szövegkoherenciáról*. Debrecen: Debreceni Egyetem Magyar Nyelvtudományi Tanszék. 178–197.
- Halácsy Péter–Kornai András–Németh László–Rung András–Szakadát István–Trón Viktor 2003. A Szószablya projekt. Alexin Zoltán–Csendes Dóra szerk., I. *Magyar Számítógépes Nyelvészeti Konferencia*. Szeged: Szegedi Tudományegyetem Informatikai Tanszékcsoport. 298.
- Halácsy Péter–Kornai András–Németh László–Sass Bálint–Varga Dániel–Váradi Tamás–Vonyó Attila 2005. A Hunglish korpusz és szótár. Alexin Zoltán–Csendes Dóra szerk., III. *Magyar Számítógépes Nyelvészeti Konferencia*. Szeged: Szegedi Tudományegyetem Informatikai Tanszékcsoport. 134–142.
- Héja Enikő 2021. Az EFNILEX első szakasza. Dodé Réka–Ludányi Zsófia szerk., *A korpusznyelvésztől a neurális hálókig. Köszöntő kötet Váradi Tamás 70. születésnapjára*. Budapest: Nyelvtudományi Kutatóközpont. 51–57.
- Holecz Margit–Bartha Csilla 2017. A Magyar Jelnyelvi Korpusz. Benő Attila–Fazakas Noémi szerk., *Élőnyelvi kutatások és a dialektológia*. Kolozsvár: Erdélyi Múzeum Egyesület. 241–251.



- Hollós Zita 2012. Az ördög a részletekben, mármint a szelekcióban rejlik – egy korpuszalapú kollokációs szótár, a SZÓKAPTÁR margójára. Pintér Tibor–Pődör Dóra– P. Márkus Katalin szerk., *Szavak pásztora. Írások Magay Tamás tiszteletére*. Szeged: Grimm Kiadó. 222–241.
- Horváth József 2000. A JPU korpusz kialakítása és tartalma. *Modern Nyelvoktatás* 6/2–3, 42–49.
- Horváth Péter 2022. Zárt fogalmi tartományok a magyar kanonikus költészetben – egy korpusznyelvészeti megközelítés. *Alkalmazott Nyelvészeti Közlemények* 15/2, 161–180.
- Horváth Viktória–Bóna Judit–Dér Csilla Ilona–Gyarmathy Dorottya–Hámori Ágnes–Huszár Anna–Kerpsz Valéria–Veidl Zsófia 2021. Fonetika és társalgáskutatás: a beszélőváltások kérdésköre. Markó Alexandra szerk., *Tanulmányok a beszédtudomány alkalmazásainak köréből*. Budapest: ELTE Eötvös Kiadó. 251–278.
- Hunyadi László–Szekrényes István 2021. Hangok, hangulatok, gesztusok: magyar nyelvű dialógusok multimodális vizsgálata. Dodé Réka–Ludányi Zsófia szerk., *A korpusznyelvészettől a neurális hálókig. Köszöntő kötet Váradi Tamás 70. születésnapjára*. Budapest: Nyelvtudományi Kutatóközpont. 81–89.
- Ittész Nóra 2012. Nyelvhasználat és szótári ábrázolás viszonya a korpuszalapú szótárban. *Magyar Nyelvőr* 136/4, 390–394.
- Ittész Nóra 2014. Lexikalizálódás, szemantika, grammatika, etimológia. Szempontok a magyar nyelv nagyszótára címszavainak meghatározásában. Laczkó Krisztina–Tátrai Szilárd szerk., *Elmélet és módszer. Nyelvészeti tanulmányok*. Budapest: Eötvös Collegium. 147–164.
- Juhász Valéria 2019. Kvalitatív és kvantitatív szövegelemzés szoftverrel. Sulyok Hedvig–Juhász Valéria–Erdei Tamás szerk., *Beszéd- és nyelvelemző szoftverek a versenyképességért és az esélyegyenlőségért. Hun CLARIN korpuszok és nyelvtechnológiai eszközök a bölcsészeti- és társadalomtudományokban*. Szeged: Szegedi Tudományegyetem Juhász Gyula Pedagógusképző Kar Magyar és Alkalmazott Nyelvészeti Tanszék. 53–63.
- Kalcsó Gyula 2007. Az elativusi, delativusi és ablativusi ragok vizsgálata a magyar nyelvű nyomtatott írásbeliség első fel századában. *Magyar Nyelv* 103/3, 282–300.
- Kalcsó Gyula 2008. A digitális korpuszokon alapuló nyelvtörténeti vizsgálatok korlátai. Büky László–Forgács Tamás–Sinkovics Balázs szerk., *A nyelvtörténeti kutatások újabb eredményei*. Szeged: Szegedi Tudományegyetem Magyar Nyelvészeti Tanszék. 95–111.
- Kalcsó Gyula 2014. A -stul/-stül eredetmagyarázatához. Zimányi Árpád szerk., *Acta Academiae Paedagogicae Agriensis. Sectio Linguistica Hungarica*. Eger: Eszterházy Károly Főiskola. 149–155.
- Kappel Péter, Modrián–Horváth Bernadett–Andreas Nolda–Vargáné Drewnowska Ewa 2020. Német-magyar nyelvtanulói korpusz (Dulko). Berend Gábor–Gosztolya Gábor–Vincze Veronika szerk. XVI. *Magyar Számítógépes Nyelvészeti Konferencia*. Szegedi Tudományegyetem TTIK, Informatikai Intézet. 369–383.
- Károly Krisztina 2003. Korpusznyelvészet és fordításkutatás. *Fordítástudomány* 5/2, 18–26.
- Károly Krisztina 2022. A nyelvi közvetítés empirikus kutatásának módszerei. Klaudy Kinga–Robin Edina–Seidl-Pécs Olívia szerk., *Bevezetés a fordítás és a tolmácsolás kutatómódszertanába I. Általános rész*. Budapest: ELTE FTT–MANYE Fordítástudományi Szakosztály. 27–58.
- Kis Balázs 2005. Automatikus terminológikeresés számítógéppel – kísérlet. *Fordítástudomány* 7/1, 84–96.

- Kis Balázs–Pohl Gábor 2005. Az automatikus terminológiakivonatolás módszerei és eredményei. Alexin Zoltán–Csendes Dóra szerk., *III. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged: Szegedi Tudományegyetem Informatikai Tanszékcsoport. 221–229.
- Kiss Gábor 2004. A piros, vörös és más színnevek használata a Magyar Nemzeti Szövegtár alapján. Geccsó Tamás szerk., *Variabilitás és nyelvhasználat*. Budapest: Tinta Könyvkiadó. 160–165.
- Kiss Gabriella–Kiss Margit–Pajzs Júlia 2004. A Nagyszótár történeti korpuszának elemzéséről. *Magyar Nyelv* 100/2, 185–191.
- Kiss Hermina 2016. Az érzelmek beszédre gyakorolt hatása, azaz a spontán beszéd szintaxisának érzelmekkel való kapcsolata a HuComTech Korpuszban. Tanács Attila–Varga Viktor–Vincze Veronika szerk., *XII. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged: Szegedi Tudományegyetem TTIK, Informatikai Intézet. 183–192.
- Kiss Katalin 2008. A morfológiai termékenység korpusz-alapú vizsgálata. Várad Tamás szerk., *II. Alkalmazott Nyelvészeti Doktorandusz Konferencia*. Budapest: MTA Nyelvtudományi Intézet. 56–65.
- Kiss Margit 2019. Stilometriai elemzés lehetőségei magyar történeti szövegtörzshen. *Digitális Bölcsészet* 2, 15–33.
- Kiss Márton 2009. Fokozó értelmű szókapcsolatok detektálása magyar szövegtörzshenben. Tanács Attila–Szauter Dóra–Vincze Veronika szerk., *VI. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged: Szegedi Tudományegyetem Informatikai Tanszékcsoport. 139–150.
- Klaudy Kinga 2001. Mit tehet a fordítástudomány a magyar nyelv „korszerűsítéséért”? *Magyar Nyelvőr* 125/2, 145–152.
- Klaudy Kinga 2005. Párhuzamos korpuszok felhasználása a fordításkutatásban. Lanstyák István–Vančoné Kremmer Ildikó szerk., *Nyelvészetről – változatosan. Segédkönyv egyetemisták és a nyelvészet iránt érdeklődők számára*. Dunaszerdahely: Gramma Nyelvi Iroda. 153–185.
- Klaudy Kinga 2018. Eredeti magyar szakszöveg vs. fordított magyar szakszöveg. Bódi Zoltán–Sz. Hegedűs Rita–Szöllősy-Sebestyén András szerk., *Magyar szaknyelvek a Kárpát-medencében. Segédkönyvek a nyelvészet tanulmányozásához*. Budapest: Tinta Kiadó–Magyar Nyelvstratégiai Intézet–Magyar Alkalmazott Nyelvészek és Nyelvtanárok Egyesülete. 1–12.
- Klaudy Kinga 2020. Visszatekintés egy korai korpuszelemzésre. Robin Edina–Seidl-Pécs Olívia szerk., *Fókuszban a fordított és a tolmácsolás szöveg: korpuszalapú fordítás Magyarországon*. Budapest: ELTE BTK Fordítástudományi Doktori Program–MANYE Fordítástudományi Szakosztály. 25–54.
- Klaudy Kinga 2021. Fordításelméleti vizsgálódások korpusznyelvészeti elemzése. *Modern Nyelvoktatás* 27/3–4, 31–42.
- Klaudy Kinga 2022. A tudományos párbeszéd játékszabályai. Klaudy Kinga–Robin Edina–Seidl-Pécs Olívia szerk., *Bevezetés a fordítás és a tolmácsolás kutatómódszertanába I. Általános rész*. Budapest: ELTE FTT–MANYE Fordítástudományi Szakosztály. 181–207.
- Kocsis Ágnes 2022. *Informatikai eszközök a nyelvészeti kutatásban. Segédanyag az informatikai eszközök használatához kognitív nyelvészeknek*. Budapest: ELTE BTK Nyelvtudományi Doktori Iskola.
- Kohn János 1999. Párhuzamos szövegek számítógéppel segített elemzése a fordításoktatásban. 1–2. rész. *Fordítástudomány* 1/1, 67–78; 2/1, 5–16.

- Kovács Éva 2014. Az angol mint lingua franca sajátosságairól. Zimányi Árpád szerk., *Acta Academiae Paedagogicae Agriensis. Sectio Linguistica Hungarica*. Eger: Eszterházy Károly Főiskola. 55–64.
- Kovácsné Dudás Andrea 2007. Szállodai honlapok elemzése korpusznyelvészeti eszközökkel. *Fordítástudomány* 9/1, 97–108.
- Kovács Tímea 2020. Korpusznyelvészet és nyelvi közvetítés. Fordításoktatási szakmai nap. *Fordítástudomány* 22/2, 107:117.
- Kövecses Zoltán 2022. A nemzet fogalma – másként. *Magyar Nyelvőr* 146/2, 137–164.
- Kresztyankó Annamária 2017. Etnicitáskonstrukciók a fiatalok nyelvi interakcióiban. Benő Attila–Fazakas Noémi szerk., *Élőnyelvi kutatások és a dialektológia*. Kolozsvár: Erdélyi Múzeum Egyesület. 178–185.
- Kugler Nóra 2008. Az episztemikus modalitást és evidencialitást jelölő módosítószók funkciói és a hozzájuk kapcsolódó műveletek. Tolcsvai Nagy Gábor–Ladányi Mária szerk. *Általános Nyelvészeti Tanulmányok XXII. Tanulmányok a funkcionális nyelvészet köréből*. Budapest: Akadémiai Kiadó. 269–307.
- Kugler Nóra–Tolcsvai Nagy Gábor szerk. 2000. *Nyelvi fogalmak kisszótára*. Budapest: Korona.
- Kurtán Zsuzsa 2020. A jogi-közigazgatási szaknyelv rétegzettsége és a közérthetőség. Bocz Zsuzsanna–Besznyák Rita szerk., *Porta Lingua 2020. Szaknyelvoktatás és -kutatás nemzetközi kontextusban. Cikkek, tanulmányok a hazai szaknyelvoktatásról*. Budapest: Szaknyelvoktatók és -Kutatók Országos Egyesülete. 123–134.
- Ladányi Mária 2008. Produktivitás és analógia a szóképzésben: a magyar főnévképzés néhány esete. Tolcsvai Nagy Gábor–Ladányi Mária szerk. *Általános Nyelvészeti Tanulmányok XXII. Tanulmányok a funkcionális nyelvészet köréből*. Budapest: Akadémiai Kiadó. 349–430.
- Ladányi Mária 2014. Idegen képzők a magyar szóképzésben: rendszeresség, produktivitás, versengés. Laczkó Krisztina–Tátrai Szilárd szerk., *Elmélet és módszer. Nyelvészeti tanulmányok*. Budapest: Eötvös Collegium. 237–262.
- Laki László János–Prószéky Gábor 2010. Statisztikai és hibrid módszerek párhuzamos korpuszok feldolgozására. Tanács Attila–Vincze Veronika szerk., VII. *Magyar Számítógépes Nyelvészeti Konferencia*. Szeged: Szegedi Tudományegyetem. 69–79.
- Ligeti-Nagy Noémi 2015. Szöveggörpuszok pontosabb annotációja gépi elemzéshez. Benő Attila–Fazakas Emese–Zsemlyei Borbála szerk., *Többnyelvűség és kommunikáció Kelet-Közép-Európában*. Kolozsvár: Babeş–Bolyai Tudományegyetem. 421–429.
- Lanstyák István 2009. A *platni* botránya (Egy új stílusminősítési rendszer felé). *Fórum Társadalomtudományi Szemle*, 11/2, 25–40.
- Lanstyák István 2018. Nyelvi originalizmus a szavak hangalakjának nyelvhelyességi megítélésében. Misad Katalin–Csehy Zoltán szerk., *Nova Posoniensia* VIII. Pozsony: Senczi Molnár Albert Egyesület. 7–34.
- Lanstyák István 2019. Nyelvi originalizmus a jelentés helyességének megítélésében. Misad Katalin–Csehy Zoltán szerk., *Nova Posoniensia* IX. Pozsony: Senczi Molnár Albert Egyesület. 30–58.
- Lanstyák István 2019. A funkciógés szerkezetek néhány általános kérdéséről. *Fórum Társadalomtudományi Szemle* 21/4, 61–91.
- Lanstyák István 2020. A funkciógés szerkezetek megítélése a nyelvhelyesség és a nyelvi helyénvalóság szempontjából. *Fórum Társadalomtudományi Szemle* 22/1, 53–78.
- Lanstyák István 2023. A nyelvi aformalizmus és a formális stílus. *Fórum Társadalomtudományi Szemle* 25/1, 3–36.

- Lanstyák István 2023. A nyelvi kontaminizmus és a magyar nyelv művelés. *Magyar Nyelvjárások* 61, 517–540.
- Lanstyák István 2023. Marginalizmusok néhány közkeletű protestáns magyar bibliafordításban. Misad Katalin–Csehy Zoltán szerk., *Nova Posoniensia XIII*. Comenius Egyetem–Szenci Molnár Albert Egyesület. 7–42. (Megjelenés alatt.)
- Makrai Márton 2021. Az EFNILEX és egy fiatal kutató. Hat év magyar szóbeágyazásokkal. Dodé Réka–Ludányi Zsófia szerk., *A korpusznyelvészettől a neurális hálókig. Köszöntő kötet Várad Tamás 70. születésnapjára*. Budapest: Nyelvtudományi Kutatóközpont. 103–107.
- Malaczkov Szilvia 2020. *A vonatkozó mellékmondat használata a feliratokban: TED-előadások magyar nyelvű összehasonlítható korpuszának vizsgálata*. Robin Edina–Seidl-Pécs Olívia szerk., *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordítás Magyarországon*. Budapest: ELTE BTK Fordítástudományi Doktori Program–MANYE Fordítástudományi Szakosztály. 261–290.
- Mány Dániel 2020. *Idegen szavak félautomatikus elemzése autentikus és fordított orvosi szövegekben: fordítási stratégiák angolról franciára és magyarra fordított beteg tájékoztatók tükrében*. Robin Edina–Seidl-Pécs Olívia szerk., *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordítás Magyarországon*. Budapest: ELTE BTK Fordítástudományi Doktori Program–MANYE Fordítástudományi Szakosztály. 123–175.
- Mittelholcz Iván 2019. Bevezetés az e-magyar programcsomag használatába. Sulyok Hedvig–Juhász Valéria–Erdei Tamás szerk., *Beszéd- és nyelvelemző szoftverek a versenyképességért és az esélyegyenlőségért. Hun CLARIN korpuszok és nyelvtechnológiai eszközök a bölcsészeti- és társadalomtudományokban*. Szeged: Szegedi Tudományegyetem Juhász Gyula Pedagógusképző Kar Magyar és Alkalmazott Nyelvészeti Tanszék. 44–52.
- Nagy Annamária Lilla 2020. *A szógyakoróság gépi vizsgálata műszaki szövegeken: terminusok és műszaki formulaszerű elemek*. Robin Edina–Seidl-Pécs Olívia szerk., *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordítás Magyarországon*. Budapest: ELTE BTK Fordítástudományi Doktori Program–MANYE Fordítástudományi Szakosztály. 199–227.
- Nagy C. Katalin 2022. „Aki tud megköszönném”. Magyar közvetett kérések vizsgálata korpuszpragmatikai megközelítésben. *Jelentés és Nyelvhasználat* 9/1, 15–42.
- Nagy T. István–Vincze Veronika 2016. Angol–magyar többszavas kifejezések szótárának automatikus építése párhuzamos korpuszok segítségével. Tanács Attila–Varga Viktor–Vincze Veronika szerk., *XII. Magyar Számítógépes Nyelvészeti Konferencia*. Szegedi Tudományegyetem TTIK, Informatikai Intézet. 298–304.
- Nemeskey Dávid Márk 2020. Egy emBERT próbáló feladat. Berend Gábor–Gosztolya Gábor–Vincze Veronika szerk. *XVI. Magyar Számítógépes Nyelvészeti Konferencia*. Szegedi Tudományegyetem TTIK, Informatikai Intézet. 409–418.
- Oravecz Csaba 2021. A Magyar nemzeti szövegtár. Dodé Réka–Ludányi Zsófia szerk., *A korpusznyelvészettől a neurális hálókig. Köszöntő kötet Várad Tamás 70. születésnapjára*. Budapest: Nyelvtudományi Kutatóközpont. 35–42.
- Pajzs Júlia 1997. Milyen szótár készíthető a nagyszótári korpuszból? *Szavak – nevek – szótárak. Írások Kiss Lajos 75. születésnapjára*. Budapest: Nyelvtudományi Intézet. 289–297.
- Pajzs Júlia 1999. Homonimák és homográfok egyértelműsítése a nagyszótári korpuszban. Gecső Tamás szerk., *Poliszémia, homonímia. Segédkönyvek a nyelvészet tanulmányozásához*. Budapest: Tinta Könyvkiadó. 245–248.

- Pajzs Júlia é. n. *Nyelvtechnológia a lexikográfia szolgálatában*.  
[http://real.mtak.hu/20147/1/Pajzs\\_Julia\\_NyT\\_Lex\\_u\\_132156.500844.pdf](http://real.mtak.hu/20147/1/Pajzs_Julia_NyT_Lex_u_132156.500844.pdf)
- Pápai Vilma 2020. Explicitáció – a fordítás univerzális jellemzője? Robin Edina–Seidl–Pécs Olívia szerk., *Fókuszban a fordított és a tolmácsoló szöveg: korpuszalapú fordítás Magyarországon*. Budapest: ELTE BTK Fordítástudományi Doktori Program–MANYE Fordítástudományi Szakosztály. 95–122.
- Pécs Olívia 2007. A lexikai kohézió vizsgálata a fordításkutatásban – új távlatok a korpusznyelvészeti módszerek köszönhetően. *Fordítástudomány* 9/1, 79–96.
- Pécs Olívia 2009. Miért érdekes a lexikai kohéziós eszközhasználat vizsgálata a fordításkutatás számára? Menyhárt József–Presinszky Károly szerk., *Fordítás – kétnyelvűség. Nyelvészeti tanulmányok a fordítás elméletéről és gyakorlatáról*. Nyitra: Konstantin Filozófus Egyetem Közép-európai Tanulmányok Kara. 27–38.
- Pécs Olívia 2010. Korpusznyelvészeti módszer a fordításkutatásban. Zimányi Árpád szerk., *A tudomány nyelve – A nyelv tudománya. Alkalmazott nyelvészeti kutatások a magyar nyelv évében*. Székesfehérvár–Eger: MANYE–Eszterházy Károly Főiskola. 208–216.
- Petykó Márton 2012. A blog műfaji jellemzőinek korpuszalapú vizsgálata. *Magyar Nyelvőr* 136/1, 45–72.
- Pintér Tibor 2003. Amit a modern nemzeti korpuszokról tudni kell. *Fórum Társadalomtudományi Szemle* 5/3, 71–85.
- Pintér Tibor 2003. Korpuszok – a múltban és napjainkban. *Katedra* 10/9, 24–25.
- Pintér Tibor 2007. „Határtalan” magyar nyelv – az első, határon túli magyar nyelvváltozatokat tartalmazó strukturált magyar nyelvi korpuszról. *Fórum Társadalomtudományi Szemle* 9/1, 165–182.
- Pintér Tibor 2008. Gondolatok a Kárpát-medencei magyar nyelvi korpusz bővítéséről. A magyar nyelv „határtalanításának” egyik újabb eredménye 1. *Magyar Nyelv* 104/1, 81–88.
- Pintér Tibor 2008. Gondolatok a Kárpát-medencei magyar nyelvi korpusz bővítéséről. A magyar nyelv „határtalanításának” egyik újabb eredménye 2. *Magyar Nyelv* 104/1, 211–217.
- Pintér Tibor 2011. A Kárpát-medencei magyar nyelvi korpusz bővítéséről. Benő Attila–Péntek János szerk., *A Termini Magyar Nyelvi Kutatóhálózat tíz éve: Tanulmányok, beszámolók, kutatási programok*. 277–299.
- Pintér Tibor 2013. Nyelvtechnológiai javak a magyar nyelvhasználatért. Fedinec Csilla–Ilyés Zoltán–Simon Attila–Vizi Balázs szerk., *A közép-európaiság dicsérete és kritikája*. Pozsony: Kalligram Kiadó. 577–591.
- M. Pintér Tibor 2017. Adatbányászat a kétnyelvűség mögött. Benő Attila–Gúti Erika–Juhász Dezső–Sztókai Szilvia–Terbe Erika–Trócsányi András szerk., *Tudományköziség és magyarságtudomány a nyelvi dimenziók tükrében*. Budapest: Termini Egyesület. 158–167.
- M. Pintér Tibor 2017. Számszerűsített kétnyelvűség. Adatbányászati technikák haszna a nyelvészetben. Misad Katalin szerk., *Kétnyelvűség térben és időben*. 125–157.
- M. Pintér Tibor 2021. Online segédletek a magyar nyelvű bibliafordítások olvasásához. *Modern Nyelvoktatás* 27/3–4, 43–57.
- M. Pintér Tibor 2022. Online bibliaolvasók szerepe és terminológiai megoldások keresése a bibliai konkordanciakészítésben. *Alkalmazott Nyelvtudomány* 22/1, 90–103.
- M. Pintér Tibor–P. Márkus Katalin 2021. Korpuszok a bibliafordításban mint a lexikológiai vizsgálatok eszközei. Fabiny Tibor–M. Pintér Tibor szerk., *πῶς ἀναγινώσκεις; Hogyan olvasod? Felekezeteket összekötő Egyesített Bibliaolvasó felé*. 72–89.

- M. Pintér Tibor–P. Márkus Katalin 2022. The Role of Online Bible Readers in Biblical Concordance Making. *Hungarian Studies Yearbook* 4/1, 183–196.
- Porkoláb Ádám 2015. A nyelvi nacionalizmus és fetisizmus ideológiájának korpusznyelvészeti vizsgálata magyar és angol nyelvű blogbejegyzésekben. Nádor Orsolya–Szűcs Tibor szerk., *Hungarológiai Évkönyv* 16/1, 24–38.
- Porkoláb Ádám 2015. Az internetes kommunikáció különbségei blogokban és Facebook-oldalokon. *Argumentum* 11, 250–264.
- Porkoláb Ádám 2015. Az iskolában elsajátított írott nyelvi norma presztízse az internetes kommunikációban. *Iskolakultúra* 25/11, 75–89.
- H. Prikler Renáta–Huszár Zoltán 2020. Populista aspektusok a politikai beszédekben. Az „én” és a „mi” személyes névmások használata 20. századi diktátorok retorikájában. *Tudásmenedzsment* 21/1–2, 145–163.
- Prószéky Gábor 2000. A magyar morfológia számítógépes kezelése. Kiefer Ferenc szerk., *Strukturális magyar nyelvtan 3. Morfológia*. Budapest: Akadémiai Kiadó. 1024–1065.
- Prószéky Gábor 2001. Nyelvtechnológiai alapkutatások – mik azok? Bartha Magdolna–Stephanides Éva szerk., *A nyelv szerepe az információs társadalomban. A X. Magyar Alkalmazott Nyelvészeti Kongresszus előadásainak válogatott gyűjteménye*. Székesfehérvár: Kodolányi János Főiskola. 13–23.
- Prószéky Gábor 2012. A nyelvtechnológia és a magyar nyelvtudomány. *Magyar Nyelv* 108: 1–18.
- Prószéky Gábor 2012. A magyarországi számítógépes nyelvészet történeti áttekintése. Prószéky Gábor–Várady Tamás szerk., *Általános Nyelvészeti Tanulmányok XXIV. Nyelvtechnológiai kutatások*. Budapest: Akadémiai Kiadó. 17–45.
- Prószéky Gábor 2017. A számítógép, az elektronikus kommunikáció és az internet hatása. Tolcsvai Nagy Gábor szerk., *A magyar nyelv jelene és jövője*. Budapest: Gondolat Kiadó. 321–335.
- Prószéky Gábor 2021. A gépi fordítás hetvenéves története. Dodé Réka–Ludányi Zsófia szerk., *A korpusznyelvészettől a neurális hálókig. Köszöntő kötet Váradi Tamás 70. születésnapjára*. Budapest: Nyelvtudományi Kutatóközpont. 141–156.
- Prószéky Gábor é. n. A magyar nyelv és a magyar nyelvtechnológia viszonyáról. [https://mta.hu/data/dokumentumok/i\\_osztaly/1\\_Eloadasok\\_tara/Magyar\\_nyelv\\_es\\_kutatasa\\_20020502/Proszeky\\_Nyelvtechnologia\\_20020502.pdf](https://mta.hu/data/dokumentumok/i_osztaly/1_Eloadasok_tara/Magyar_nyelv_es_kutatasa_20020502/Proszeky_Nyelvtechnologia_20020502.pdf)
- Prószéky Gábor–Olaszy Gábor–Váradi Tamás 2006. Nyelvtechnológia. Kiefer Ferenc szerk., *Magyar nyelv*. Budapest: Akadémiai Kiadó. 1038–1072.
- Rebrus Péter–Kornai András–Varga Dániel 2012. Egy általános célú morfológiai annotáció. Prószéky Gábor–Várady Tamás szerk., *Általános Nyelvészeti Tanulmányok XXIV. Nyelvtechnológiai kutatások*. Budapest: Akadémiai Kiadó. 47–80.
- Reményi Andrea Ágnes é. n. Tervezési megfontolások a Magyar Nemzeti Szövegtár számára. <https://adoc.pub/queue/remenyi-andrea-agnes-tervezesi-megfontolasok-a-magyar-nemzet.html> (Utolsó letöltés: 2023. november 1.)
- Robin Edina–Dankó Szilvia–Götz Andrea–Nagy A. Lilla–Pataky Éva–Szegh Henriette–Zolczer Péter 2016. Fordítástudomány és korpuszkutatás: bemutatkozik a Pannonia Korpusz. *Fordítástudomány* 18/2, 5–26.
- Robin Edina–Seidl-Pécs Olívia 2020. Korpuszalapú fordításkutatás Magyarországon. Robin Edina–Seidl-Pécs Olívia szerk., *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordítás Magyarországon*. Budapest: ELTE BTK Fordítástudományi Doktori Program–MANYE Fordítástudományi Szakosztály. 9–24.

- Sass Bálint 2009. Korpusznyelvészeti eszköz a magyar igék bővítményszerkezetének vizsgálatára. Sinkovics Balázs szerk., *LingDok 8 – Nyelvész-doktoranduszok dolgozatai*. Szeged: JATEPress. 143–155.
- Sass Bálint 2016. Nyelvészeti szövegkeresők, Nemzeti Korpuszportál. *Magyar Tudomány* 177/7, 798–808
- Sass Bálint 2017. Keresés a korpuszban: a kibővített Magyar történeti szövegtár új keresőfelülete. Forgács Tamás–Németh Miklós–Sinkovics Balázs szerk., *A nyelvtörténeti kutatások újabb eredményei IX*. Szeged: Szegedi Tudományegyetem Magyar Nyelvészeti Tanszék. 267–277.
- Sass Bálint 2017. Keresés korpuszban: a kibővített Magyar történeti szövegtár új keresőfelülete. Forgács Tamás–Németh Miklós–Sinkovics Balázs szerk., *A nyelvtörténeti kutatások újabb eredményei IX*. Szeged: Szegedi Tudományegyetem Magyar Nyelvészeti Tanszék. 267–277.
- Sass Bálint 2019. Keresés korpuszban 2: így kerestek ti. Sulyok Hedvig–Juhász Valéria–Erdei Tamás szerk., *Beszéd- és nyelvelemző szoftverek a versenyképességért és az esélyegyenlőségért. Hun CLARIN korpuszok és nyelvtechnológiai eszközök a bölcsész- és társadalomtudományokban*. Szeged: Szegedi Tudományegyetem Juhász Gyula Pedagógusképző Kar Magyar és Alkalmazott Nyelvészeti Tanszék. 21–30.
- Sass Bálint 2020. A duplakocka modell és az igei szerkezeteket kinyerő „ugrik és marad” módszer nyelvfüggetlensége, valamint néhány megjegyzés az UD annotáció univerzalizálásáról. Berend Gábor–Gosztolya Gábor–Vincze Veronika szerk. *XVI. Magyar Számítógépes Nyelvészeti Konferencia*. Szegedi Tudományegyetem TTIK, Informatikai Intézet. 399–407.
- Sebők Szilárd–Lanstyák István 2013. A szófeltárástól a szócikkírásig. Misad Katalin–Csehy Zoltán szerk., *Nova Posoniensia III*. Pozsony: Szenczi Molnár Albert Egyesület–Kalligram Kiadó. 15–66.
- Seidl-Pécs Olivía 2018. Melyek a (szak)fordító és a fordításkutató munkáját segítő legfontosabb nyelvi korpuszok? Robin Edina–Zachar Viktor szerk., *Fordítástudomány ma és holnap*. Budapest: L’Harmattan. 175–191.
- Seidl-Pécs Olivía 2020. Korpuszalapú fordításkutatás: lehetőségek és nehézségek. Fókuszban a korpuszépítés és a korpuszalapú elemzés. Robin Edina–Seidl-Pécs Olivía szerk., *Fókuszban a fordított és a tolmácsolás szöveg: korpuszalapú fordítás Magyarországon*. Budapest: ELTE BTK Fordítástudományi Doktori Program–MANYE Fordítástudományi Szakosztály. 79–94.
- Seresi Márta 2022. Kutatási adatok kezelése a fordítástudományban. Klaudy Kinga–Robin Edina–Seidl-Pécs Olivía szerk., *Bevezetés a fordítás és a tolmácsolás kutatómódszertanába I. Általános rész*. Budapest: ELTE FTT–MANYE Fordítástudományi Szakosztály. 93–107.
- Siklósi Borbála–Novák Attila 2016. Közeli rokonunk, az autó. Tanács Attila–Varga Viktor–Vincze Veronika szerk., *XII. Magyar Számítógépes Nyelvészeti Konferencia*. Szegedi Tudományegyetem TTIK, Informatikai Intézet. 27–36.
- Simon Eszter 2019. Magyar nyelvű történeti korpuszok. Sulyok Hedvig–Juhász Valéria–Erdei Tamás szerk., *Beszéd- és nyelvelemző szoftverek a versenyképességért és az esélyegyenlőségért. Hun CLARIN korpuszok és nyelvtechnológiai eszközök a bölcsész- és társadalomtudományokban*. Szeged: Szegedi Tudományegyetem Juhász Gyula Pedagógusképző Kar Magyar és Alkalmazott Nyelvészeti Tanszék. 31–43.
- Simon Eszter–Váradi Tamás 2018. Magyar nyelvtechnológiai infrastruktúra a társadalomtudományok szolgálatában. *Magyar Tudomány* 179/7, 1017–1026.

- Simon Eszter–Sass Bálint 2012. Nyelvtechnológia és kulturális örökség, avagy korpuszépítés ómagyar kódexekből. Prószéky Gábor–Várady Tamás szerk., *Általános Nyelvészeti Tanulmányok XXIV. Nyelvtechnológiai kutatások*. Budapest: Akadémiai Kiadó. 243–264.
- Simon Eszter–Mittelholcz Iván–Ferenczi Zsannett 2019. Automatikus szótárépítés kisebbségi finnugor nyelvekre. Pletl Rita–Kovács Gabriella szerk., *Soknyelvűség és többnyelvűség Európában*. Kolozsvár: Scientia Kiadó. 53–64.
- Simon Gábor 2018. Az igei jelentés metaforizációjának mintázatai. Nyelvtan- és korpuszvezérelt esettanulmányok. *Jelentés és Nyelvhasználat* 5, 1–36.
- Simon Gábor 2020. Metafora a korpuszban: módszertani lehetőségek. *Literatura* 46/3, 297–311.
- Simon Gábor 2022. A számolás joga, avagy korpusznyelvészet és irodalomtudomány. *Tempevölgy: Kultúra, művészet, tudomány* 13/3, 84–92.  
[http://real.mtak.hu/148697/1/SG\\_A%20sz%C3%A1mol%C3%A1s%20joga.pdf](http://real.mtak.hu/148697/1/SG_A%20sz%C3%A1mol%C3%A1s%20joga.pdf)
- Simon Gábor–Szlávi Eszter–Bajzát Tímea–Ballagó Júlia–Havasi Zsuzsanna–Roskó Mira 2019. Metaforaazonosítás magyar nyelvű szövegekben: egy módszer adaptálásáról. *Magyar Nyelvőr* 143/2, 223–247.
- Simon László 2016. A digitális korszak vívmányainak hasznosulása a lexikográfiában: a nagyszótári projekt informatikai fejlesztéséről. *Magyar Tudomány* 177/7, 809–815.
- Szabó Gergely 2017. Obszcenitások és megítélésük hétköznapi társalgásokban és metadiskurzusokban. Benő Attila–Fazakas Noémi szerk., *Élőnyelvi kutatások és a dialektológia*. Kolozsvár: Erdélyi Múzeum Egyesület. 186–194.
- Szabó Martina Katalin 2011. Jelentés az első digitalizált magyar–orosz párhuzamos korpusz (HunOr) építéséről. Problémák és lehetőségek. Parapatics Andrea–Csernák-Szuhánszky Debóra szerk., *Félúton konferencia 7*. Budapest: Eötvös Loránd Tudományegyetem BTK Nyelvtudományi Doktori Iskola. 1–12.
- Szabó Martina Katalin–Schmalcz András–Nagy T. István–Vincze Veronika 2011. A HunOr magyar–orosz párhuzamos korpusz. Tanács Attila – Vincze Veronika szerk., *VIII. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged: JATEPress. 341–347.
- Szabó Martina Katalin–Ring Orsolya–Vincze Veronika 2020. 1956 és 1989 között keletkezett propagandaszövegek nyelvi sajátosságai. Berend Gábor–Gosztolya Gábor–Vincze Veronika szerk. *XVI. Magyar Számítógépes Nyelvészeti Konferencia*. Szegedi Tudományegyetem TTIK, Informatikai Intézet. 357–368.
- Szabó Martina Katalin–Vincze Veronika–Hangya Viktor 2016. Aspektusszintű annotáció és szentimentet módosító elemek egy magyar nyelvű szentimentkorpuszban. Tanács Attila–Varga Viktor–Vincze Veronika szerk., *XII. Magyar Számítógépes Nyelvészeti Konferencia*. Szegedi Tudományegyetem TTIK, Informatikai Intézet. 174–182.
- Szabó Tamás Péter 2008. Szabályok értelmezése és bemutatása normatív nyelvhasználatot leíró kiadványokban. Sinkovics Balázs szerk., *Lingdok 7. Nyelvész-doktoranduszok dolgozatai*. Szeged: Szegedi Tudományegyetem Nyelvtudományi Doktori Iskola. 211–229.
- Szabó Tamás Péter 2009. Nyelvhasználói igények – szolgáltatói szemléletű szótárak. Kuna Ágnes–Veszelszki Ágnes szerk., *3. Félúton konferencia: Az Eötvös Loránd Tudományegyetem nyelvész-doktoranduszainak konferenciája*. Budapest: ELTE BTK Nyelvtudományi Doktori Iskola. 287–308.
- Szabó Veronika 2020. Korpusznyelvészeti adalékok az infinitívuszt tartalmazó szerkezetek tanításához. *THL<sub>2</sub>: A magyar nyelv és kultúra tanításának szakfolyóirata* 16/1–2, 7–36.



- Szegh Henriett 2020. *Harmadik kód a tolmácsolásban: létezik tolmácsolási szöveg? Intermodális korpuszok és esettanulmány.* Robin Edina–Seidl-Pécs Olívia szerk., *Fókuszban a fordított és a tolmácsolás szöveg: korpuszalapú fordítás Magyarországon.* Budapest: ELTE BTK Fordítástudományi Doktori Program–MANYE Fordítástudományi Szakosztály. 291–328.
- Szirmai Monika 2005. *Bevezetés a korpusznyelvészetbe. A korpusznyelvészet alkalmazása az anyanyelv és az idegen nyelv tanulásában és tanításában.* Budapest: Tinta Könyvkiadó.
- Szita Szilvia 2017. Ön most eljött vagy megjött? Korpuszhasználat a magyar mint idegen nyelv tanításában. Dóla Mónika–Nádor Orsolya–Szűcs Tibor szerk., *Hungarológiai Évkönyv 18/1.* Pécs: Pécsi Tudományegyetem Bölcsészettudományi Kara. 81–93.
- Szita Szilvia 2021. A *MagyarOK* nyílt korpusz használatáról. Dóla Mónika–Nádor Orsolya–Szűcs Tibor szerk., *Hungarológiai Évkönyv 22.* Pécs: Pécsi Tudományegyetem Bölcsész- és Társadalomtudományi Kara. 72–88.
- Tatár Zoltán–Varga Zoltán–Főző Eszter 2021. Beszélőazonosítás a kriminalisztikában. Markó Alexandra szerk., *Tanulmányok a beszédtudomány alkalmazásainak köréből.* Budapest: ELTE Eötvös Kiadó. 229–250.
- Tóth Ágoston 2013. Az ember, a korpusz és a számítógép. Magyar nyelvű szóhasználati mérések humán és disztribúciós szemantikai kísérletben. *Argumentum 9*, 301–310.
- Tóth Etelka 2011. Korpusznyelvészet és a tankönyvek. *Eruditio – Educatio 6/2*, 80–93.
- Várad Tamás 1997. Adatok a nagyszótári korpusz összetételéről. Kiss Gábor–Zaicz Gábor szerk., *Írások Kiss Lajos 75. születésnapjára.* Budapest: MTA Nyelvtudományi Intézete. 437–446.
- Várad Tamás 2000. Modern nyelvi technológiák a magyar nyelvért. Kiefer Ferenc–Gósy Mária szerk., *Helyzetkép a magyar nyelvtudományról.* Budapest: MTA Nyelvtudományi Intézet. 146–156.
- Várad Tamás 2000. Szótár, korpusz – magyar nemzeti szövegtár. Gecső Tamás szerk., *Lexikális jelentés, aktuális jelentés. Segédkönyvek a nyelvészet tanulmányozásához IV.* Budapest: Tinta Kiadó. 263–270.
- Várad Tamás 2001. Kontrasztív szemantikai kutatások párhuzamos korpusz segítségével. Gecső Tamás szerk., *Kontrasztív szemantikai kutatások.* Budapest: Tinta Könyvkiadó. 268–275.
- Várad Tamás 2001. A nyelvhasználat empirikus vizsgálatáról. Andor József–Szűcs Tibor–Terts István szerk., *Színes eszmék nem alszanak... Szépe György 70. születésnapjára.* Pécs: Lingua Franca Csoport. 1285–1291.
- Várad Tamás 2020. Mesterséges intelligencia a korpusznyelvészetben. Kovács Tímea–Adorján Mária szerk., *Korpusznyelvészet és nyelvi közvetítés.* Budapest: Károli Gáspár Református Egyetem–L'Harmattan Kiadó. 141–161.
- Várad Tamás é. n. *Magyar korpusznyelvészet és nyelvtechnológia.*  
[https://mta.hu/data/dokumentumok/i\\_osztaly/1\\_Eloadasok\\_tara/Magyar\\_nyelv\\_es\\_kutatasa\\_20020502/Varadi\\_Tamas\\_korref\\_20020502.pdf](https://mta.hu/data/dokumentumok/i_osztaly/1_Eloadasok_tara/Magyar_nyelv_es_kutatasa_20020502/Varadi_Tamas_korref_20020502.pdf)
- Várad Tamás–Héja Enikő 2010. Szaknyelv és korpusznyelvészet. Dobos Csilla szerk., *Szaknyelvi kommunikáció.* Miskolc–Budapest: Miskolci Egyetem–Tinta Könyvkiadó. 107–125.
- Várad Tamás–Oravecz Csaba 2014. A Magyar nemzeti szövegtár egymilliárd szavas új változata. *Magyar Tudomány 175/9*, 1054–1061.
- Vincze Veronika 2019. Bevezetés a korpuszok és nyelvi adatbázisok világába. Sulyok Hedvig–Juhász Valéria–Erdei Tamás szerk., *Beszéd- és nyelvelemző szoftverek a versenyképes-*

- ségért és az esélyegyenlőségért. Hun CLARIN korpuszok és nyelvtechnológiai eszközök a bölcsészet- és társadalomtudományokban. Szeged: Szegedi Tudományegyetem Juhász Gyula Pedagógusképző Kar Magyar és Alkalmazott Nyelvészeti Tanszék. 5–20.*
- Vincze Veronika–Felvégi Zsuzsanna–R. Tóth Krisztina 2010. Félig kompozicionális szerkezetek a SzegedParalell angol–magyar párhuzamos korpuszban. Tanács Attila–Vincze Veronika szerk., *VII. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged: Szegedi Tudományegyetem Informatikai Tanszékcsoport. 91–101.
- Vincze Veronika 2019. Bevezetés a korpuszok és nyelvi adatbázisok világába. Sulyok Hedvig–Juhász Valéria–Erdei Tamás szerk., *Beszéd- és nyelvelemző szoftverek a versenyképességért és az esélyegyenlőségért. Hun CLARIN korpuszok és nyelvtechnológiai eszközök a bölcsészet- és társadalomtudományokban. Szeged: Szegedi Tudományegyetem Juhász Gyula Pedagógusképző Kar Magyar és Alkalmazott Nyelvészeti Tanszék. 5–20.*
- Yang Zijian Győző–Laki László 2016. Gépi fordítás minőségbecslésének optimalizálása kétnyelvű szótár és WordNet segítségével. Tanács Attila–Varga Viktor–Vincze Veronika szerk., *XII. Magyar Számítógépes Nyelvészeti Konferencia*. Szegedi Tudományegyetem TTIK, Informatikai Intézet. 37–46.

## & A SZOROZAT KÖTETEI

2022

- Nyelvi váltakozások az igék körében. Nyelvhelyességi és nyelvi helyénvalósági kérdések.* Pozsony: Szenci Molnár Albert Egyesület. ISBN 978-80-974359-0-5. 208 p.  
[https://fphil.uniba.sk/fileadmin/fif/katedry\\_pracoviska/kmjl/2023\\_LI\\_01\\_Igek\\_v3.pdf](https://fphil.uniba.sk/fileadmin/fif/katedry_pracoviska/kmjl/2023_LI_01_Igek_v3.pdf)
- Nyelvi váltakozások az igenevek körében. Nyelvhelyességi és nyelvi helyénvalósági kérdések.* Pozsony: Szenci Molnár Albert Egyesület. ISBN 978-80-974359-1-2. 97 p.  
[https://fphil.uniba.sk/fileadmin/fif/katedry\\_pracoviska/kmjl/2023\\_LI\\_02\\_Igenevek\\_v03.pdf](https://fphil.uniba.sk/fileadmin/fif/katedry_pracoviska/kmjl/2023_LI_02_Igenevek_v03.pdf)
- Nyelvi váltakozások az igekötős igék körében. Nyelvhelyességi és nyelvi helyénvalósági kérdések.* Pozsony: Szenci Molnár Albert Egyesület. ISBN 978-80-974359-2-9. 180 p.  
[https://fphil.uniba.sk/fileadmin/fif/katedry\\_pracoviska/kmjl/2022\\_LI\\_03\\_Igekotos\\_igek\\_v5.pdf](https://fphil.uniba.sk/fileadmin/fif/katedry_pracoviska/kmjl/2022_LI_03_Igekotos_igek_v5.pdf)
- Nyelvi váltakozások a névszók körében. Nyelvhelyességi és nyelvi helyénvalósági kérdések.* Pozsony: Szenci Molnár Albert Egyesület. ISBN 978-80-974359-3-6. 252 p.  
[https://fphil.uniba.sk/fileadmin/fif/katedry\\_pracoviska/kmjl/2023\\_LI\\_04\\_Nevszok\\_v02.pdf](https://fphil.uniba.sk/fileadmin/fif/katedry_pracoviska/kmjl/2023_LI_04_Nevszok_v02.pdf)
- Nyelvi váltakozások a névutók körében. Nyelvhelyességi és nyelvi helyénvalósági kérdések.* Pozsony: Szenci Molnár Albert Egyesület. ISBN 978-80-974359-4-3. 246 p.  
[https://fphil.uniba.sk/fileadmin/fif/katedry\\_pracoviska/kmjl/2023\\_LI\\_05\\_Nevutok\\_v02.pdf](https://fphil.uniba.sk/fileadmin/fif/katedry_pracoviska/kmjl/2023_LI_05_Nevutok_v02.pdf)

2023

- Váltakozások a magyar nyelvben nyelvhelyességi és nyelvi helyénvalósági szempontból.* Pozsony: Szenci Molnár Albert Egyesület. ISBN 978-80-974359-9-8. 94 p.  
[https://fphil.uniba.sk/fileadmin/fif/katedry\\_pracoviska/kmjl/2023\\_LI\\_Valtozasok\\_a\\_magyar\\_nyelvben\\_nyelvhelyessegi\\_es\\_nyelvi.pdf](https://fphil.uniba.sk/fileadmin/fif/katedry_pracoviska/kmjl/2023_LI_Valtozasok_a_magyar_nyelvben_nyelvhelyessegi_es_nyelvi.pdf)
- A magyar nyelvi váltakozások dialektális háttere.* Pozsony: Szenci Molnár Albert Egyesület. ISBN 978-80-69002-05-0. 88 p.  
[https://fphil.uniba.sk/fileadmin/fif/katedry\\_pracoviska/kmjl/Nyeal\\_Dialvalt\\_1.pdf](https://fphil.uniba.sk/fileadmin/fif/katedry_pracoviska/kmjl/Nyeal_Dialvalt_1.pdf)
- A magyar nyelvi váltakozások kronolektális háttere.* Pozsony: Szenci Molnár Albert Egyesület. ISBN 978-80-69002-06-7. 86 p.  
[https://fphil.uniba.sk/fileadmin/fif/katedry\\_pracoviska/kmjl/Nyeal\\_Kronvalt.pdf](https://fphil.uniba.sk/fileadmin/fif/katedry_pracoviska/kmjl/Nyeal_Kronvalt.pdf)
- A magyar nyelvi váltakozások stílusbeli háttere.* Pozsony: Szenci Molnár Albert Egyesület. ISBN 978-80-69002-08-1. 122 p.  
[https://fphil.uniba.sk/fileadmin/fif/katedry\\_pracoviska/kmjl/Nyeal\\_Stilvalt.pdf](https://fphil.uniba.sk/fileadmin/fif/katedry_pracoviska/kmjl/Nyeal_Stilvalt.pdf)
- A magyar nyelvi váltakozások érzelmi háttere.* Pozsony: Szenci Molnár Albert Egyesület. ISBN 978-80-69002-07-4. 115 p.  
[https://fphil.uniba.sk/fileadmin/fif/katedry\\_pracoviska/kmjl/Nyeal\\_Expvalt.pdf](https://fphil.uniba.sk/fileadmin/fif/katedry_pracoviska/kmjl/Nyeal_Expvalt.pdf)
- Magyar nyelvi váltakozások vizsgálata a Magyar nemzeti szövegtár segítségével.* Pozsony: Szenci Molnár Albert Egyesület. ISBN 978-80-69002-09-8. 125 p.

*A személytelenség kifejezésének eszközei. Nyelvhelyességi és nyelvi helyénvalósági kérdések.*  
Pozsony: Szenci Molnár Albert Egyesület. ISBN 978-80-69002-00-5. 148 p.

ISBN 978-80-69002-09-8