

KERESÉS KORPUSZBAN 3: BRÜSSZELRŐL JELENTJÜK

Prószéky Gábornak, 70. születésnapjára

SASS BÁLINT | HUN-REN Nyelvtudományi Kutatóközpont
sass.balint@nytud.hun-ren.hu | DOI: 10.18135/PG70.2024.13

Kedves Gábor! Többek között a Te rendszeres felkéréseid nyomán jó pár éve tartok előadásokat és publikálok (Sass, 2017, 2022) arról, hogy szerintem hogyan lehet nyelviszeti kutatás céljára korpuszból megfelelően kinyerni a szükséges adatokat, és ezekre építve hogyan lehet egy szép és módszertanilag helyes korpuszvezérelt kutatást elvégezni. Kerek születésnapod alkalmából arra gondoltam, hogy itt az ideje, hogy megtudjuk, hogy nekem magamnak ez hogyan megy. :) *Isten éltessen sokáig!*

Inspirált ennek a tanulmánynak a megírására az is, hogy nemrég vált hozzáférhetővé a Magyar Országgyűlésben 2014 és 2023 között elhangzott beszédeket és hozzászólásokat tartalmazó ParlaMint 4.0 korpusz (Ligeti-Nagy et al., 2023). Ennek a korpusznak a gazdag annotációs rétegei nagyon sokféle kutatásra és összevetésre adnak lehetőséget. Számomra a legizgalmasabbnak a kormányoldali és az ellenzéki kommunikáció jellegzetességeinek vizsgálata és összehasonlítása tűnt.

Egy aspektust, vagyis igazából egyetlen szót választottam ki, és ez az utóbbi években sokak által nagyon sokszor emlegetett, sok ráakódott jelentésréteggel bíró városnév: a *Brüsszel*. Hogyan jelenik meg a kormányoldali és az ellenzéki beszédben ez a szó, hogyan és mire használja a két oldal, mik a hasonlóságok és mik a különbségek a két oldal vonatkozásában?

A kutatási hipotéziseim felállítása után az adatkezelés technikai részéről szólok, itt bemutatom az általam elkészített és a jelen kutatás elvégzéséhez is használt új, általános célú adatfeldolgozó eszközt; majd a vizsgálatok, az eredmények és a tanulságok bemutatása következik.

1. Hipotézisek

A vizsgálandó hipotéziseimet a sajtóban megjelentek, a közbeszéd és saját benyomásaim alapján fogalmazom meg.

H₁: A kormányoldal számára inkább aktív *Brüsszel*: egy olyan ágens, aminek (akinek?) szándékai vannak, csinál valamit; az ellenzék számára inkább passzív szereplő: egy város, ahol valami történik.

H₂: A kormányoldal kommunikációs paneleket, előregyártott frázisokat használ, amiket aztán sokan és sokszor ismételnék. Az ellenzék nem vagy sokkal kevésbé él ezzel az eszközzel.

H₃: A kormányoldal számára *Brüsszel* inkább negatív, az ellenzék számára inkább pozitív. A kormányoldal kommunikációjában *Brüsszel* általában ellenünk tesz valamit; az ellenzék kommunikációjában pedig inkább értünk.

2. Adatfeldolgozás: csql

A ParlaMint 4.0 korpusz szabadon elérhető a NoSketchEngine (NoSkE) korpuszkezelő rendszerben (Rychlý, 2007) a https://www.clarin.si/ske/#concordance?corpname=parlamint40_hu címen. A felület a lekérdezés támogatásán túl (pl. CQL, szűrés) számos hasznos kiegészítő funkciót biztosít (pl. gyakoriságlista-készítés, kollokációkeresés), mégis gyakran az a leginkább célravezető, ha az adatokat letöltés után táblázatos formában saját magunk kezelhetjük és dolgozhatjuk fel egyéni igényeink szerint.

Ilyenkor viszont sok esetben abba a nehézségbe ütközünk, hogy a NoSkE letöltési lehetőségei meglehetősen korlátozottak. A felkínált négyféle fájlformátum lényegében ugyanazt az adathalmazt tartalmazza, és ugyanazokat a problémákat rejti. Most a **.txt** formátumon mutatom ezt be. A konkordancia találatok sorozata, egy találat négy mezőből áll: egy fejlécmező metaadatokkal, valamint pontosan három tartalmi mező: bal kontextus, találat, jobb kontextus. A tartalmi mezők *szóközzel* elválasztva tartalmazzák az egyes tokeneket, egy tokennek a reprezentációja pedig a hozzá tartozó lekérdezett jellemzők értékeit *perjellel* (/) elválasztva. Ebből egyenesen következik a probléma: gondot okoz, ha az adattartalomban szóköz vagy perjel szerepel, és ez természetesen elő is fordul. A dizájn hibája az, hogy nem sikerült extrémális elemet találni a reprezentáció formai részének megjelenítésére. Nyilvánvaló megoldás lehetne az XML használata, de ezt a NoSkE nem valósítja meg, az **.xml**-ként hirdetett letöltési formátumában is a most leírt szóköz-perjel reprezentáció jelenik meg az XML mélyén.

A létrehozott **csql** nevű eszköz egyik funkciója éppen az, hogy kezeli a fenti problémás formátumot, azaz egyértelműen beazonosíthatóvá teszi az egyes tokeneket és az egyes jellemzőket (1. ábra). A szóközők kezeléséhez szükséges az eszköz rendelkezésére bocsájtani egy olyan, a NoSkE-ből letöltött kiegészítő fájlt, amely ugyanazt a konkordanciát tartalmazza, amit vizsgálni akarunk, de kizárólag a szóalakok vannak benne megjelenítve; a perjelek kezeléséhez pedig paraméterben meg kell adnunk, hogy melyek azok a mezők, amelyek perjeleket tartalmazhatnak.

```
a) nyolc/nyolc/NUM/Case=Nom NumType=Card Number=Sing
   Brüsszel/percben/Brüsszel/perc/NOUN/Case=Ine
   Number=Sing előadni/előad/VERB/VerbForm=Inf Voice=Act
b) ['nyolc', 'nyolc', 'NUM', 'Case=Nom++NumType=Card++Number=Sing']
   ['Brüsszel/percben', 'Brüsszel/perc', 'NOUN', 'Case=Ine++Number=Sing']
   ['előadni', 'előad', 'VERB', 'VerbForm=Inf++Voice=Act']
```

1. ábra. Példa a szóköz-perjel reprezentáció hátrányaira. Az ábra *a)* részén a *nyolc Brüsszel/percben előadni* szöveg reprezentációja látható 4 jellemző (szóalak, szótő, szófaj, morfológia) megjelenítésével. Látjuk, hogy szóalakban előfordul perjel, morfológiában pedig szóköz. Az ábra *b)* részén látjuk a **csql** által kitisztázott viszonyokat.

A `csql` másik funkciója az, hogy lehetővé teszi, hogy a meglévő három tartalmi mezőt tetszőleges továbbiakkal egészítsük ki. Azaz szabad kezet ad abban, hogy az adathalmazunk milyen (származtatott) jellemzőit szeretnénk vizsgálni, kiegészítőleg meghatározhatjuk, hogy mi az, ami számunkra releváns az adott kutatás szempontjából. Ezt egy Pythonban implementált függvény segítségével valósíthatjuk meg. Feljegyezhetjük külön mezőbe adott token adott jellemzőjét: például a jobb kontextus első szavának szófaját, vagy a találat utolsó szóalakját. Szabadon származtathatunk további jellemzőket a meglévőkből. Külön mezőt szentelhetünk például annak, hogy a találati szó nagybetűs-e, hogy a találat hány tokenből áll, mi a jobbra eső legközelebbi ige, hány főnév jelenik meg a kontextusban vagy a ParlaMint-korpuszokban meglévő függőségi annotációt felhasználva akár azt, hogy hány testvérbővítménye van egy adott bővítménynek. A könnyű visszakereshetőséghez hasznos a sorszám, ezért azt alapértelmezetten mindig hozzávesszük a fejlémező metaadataihoz.

Az eszköz szabadon hozzáférhető a <https://github.com/sassbalint/csql> címen. Az alábbi vizsgálatok során a NoSkE felület funkcióit és a `csql` eszköz lehetőségeit egymást kiegészítve használtam.

3. Adathalmaz

A lekérdezés, amellyel a vizsgálati adatokhoz hozzájutottam, nagyon kézenfekvő volt: a

brüsszel

keresőkifejezést használtam egyszerű keresésben a ParlaMint 4.0 korpusz https://www.clarin.si/ske/#concordance?corpname=parlamint40_hu címen található keresőfelületén. Azaz minden olyan szóalakat lekértem, amely bármilyen módon tartalmazza a **brüsszel** karaktersorozatot, ez mind a vizsgálat tárgyát képezi most. Ez a lekérdezés minden szóba jöhető adatot teljes körűen lefed (Sass, 2022: principle #2). A metaadatok közé a beszélőt (`speech.speaker_id`) és a kormányoldal-ellenzék státuszt (`speech.date speech.party_status`) vettem be; a jellemzők közül a szóalakat, a szótövet, a szófaját, a morfológiai jegyeket, és az adott tokenhez tartozó függőségi fej szótövet jelenítettem meg.

Bízom benne, hogy jelentős mennyiségű adatot kapok. A 6853 találat a korpusz 36,3 millió szavas méretét tekintve – azaz a 189 találat/millió szó (röviden: 189/M) gyakoriság – talán nem tűnik soknak, de a nagybetűs szótövű szavak listájának elejét átnézve azt látjuk, hogy a helyek között Magyarország, az Európai Unió, Európa és Budapest után az ötödik helyen szerepel Brüsszel, azaz a városok között a második leggyakoribb. Néhány jellegzetes példa az 1. táblázatban látható.

azonosító	szóalak	előfordulási szám
1	<i>brüsszeli</i>	2101
2	<i>Brüsszel</i>	1654
3	<i>Brüsszelben</i>	1492
4	<i>Brüsszelből</i>	309
...		
5	<i>brüsszeleznek</i>	7
6	<i>Brüsszel-Strasbourg-Soros-féle</i>	6
7	<i>Brüsszel-ellenes</i>	5
8	<i>Brüsszel-Strasbourg-Soros-triumvirátus</i>	4
9	<i>brüsszelitáknál</i>	4
...		
10	<i>brüsszelitábbak</i>	3
...		
11	<i>Brüsszel/percben</i>	1
...		
12	<i>brüsszeliőrület-vonat</i>	1

1. táblázat. Példák a *brüsszel* egyszerű lekérdezésre kapott adathalmazból. A 6853 elemű találati listában 79 különféle szóalak fordul elő. Látjuk, hogy a szónak nagyon kiterjedt bokra alakult ki: a ragozott alakok (3, 4) mellett megjelennek a különféle összetételek (6, 8, 12) és a képzett alakok, mint például a *brüsszelez* (5) vagy a *brüsszelita* (9, 10), sőt egészen speciális képződmények is, amilyen a *Brüsszel/perc* (11) mértékegység. A három leggyakoribb alak együtt az összes előfordulás több mint háromnegyedét adja.

4. Vizsgálatok és eredmények

Vizsgálataim célja a kormányoldal és az ellenzék *Brüsszel*-használatának összevetése. E két kategória mellett a korpuszban létezik egy „nem besorolt” kategória is. Ezt a továbbiakban nem veszem figyelembe, bár ránézésre ebben a csoportban is a kormányoldali kötődésű emberek (Dömötör Csaba, Orbán Balázs, Schanda Tamás stb.) vannak túlsúlyban.

név	#	
Rétvári Bence	532	A 6853 találat eloszlása: 3727 (54,4%) kormányoldal, 1799 (26,2%) ellenzék (és 1327 [19,4%] nem besorolt). A kormányoldali szövegmenyiség 13,3 millió szó, az ellenzéki 16,9 millió szó, eszerint a <i>brüsszeles</i> kifejezések gyakorisága rendre 279/M, illetve 107/M. A kormányoldalon tehát majdnem háromszoros ezen kifejezések gyakorisága. A megszólalók gyakorisági listájából is jól látszik a kormányoldali túlsúly, az első 5 helyezett mindegyike a Fidesz–KDNP-hez kötődik, ahogy ez a 2. táblázatból látszik.
Dömötör Csaba	306	
Nacsa Lőrinc	230	
Orbán Viktor	215	
Vejkey Imre	192	

2. táblázat. A *Brüsszelt* legtöbbször használó megszólalók. Mindannyian kormányoldaliak.

A konkordancián lefuttatott kollokációvizsgálatban megjelennek a jól ismert szavak: *bürokrata, szankció, elhibázott, szankciós, politika, Soros, képvisel, bevándorlaspárti, háború, elit*. A főnevek legtöbbször a *brüsszeli* jelzővel szerepelnek, *Soros* általában a *Brüsszellel* koordinációban jelenik meg, a *háború* a *brüsszeli szankciók* kíséretében, a *képvisel* igével pedig legtöbbször a *Brüsszelt* vagy *Brüsszel érdekeit képviseli* szókapcsolatban találkozunk.

A *Brüsszelt* középen tartalmazó 5-gramok közül 53 előfordulással magasan a leggyakoribb az *elhibázott brüsszeli szankciós politika*. Az 5-gramok között *elhibázottat* és *szankciót* is bárhol tartalmazó találat 142 darab van, ez önmagában az összes előfordulás 2%-a (!), azaz minden ötvenedik *Brüsszel* ebben a frázisban jelenik meg. A kifejezés eloszlása szélsőséges: 116 db (8,7/M) kormányoldal vs. 8 db (0,5/M) ellenzék.

Az aktív/passzív szerep megragadására a legrimitivebb megközelítés egyszerűen annak vizsgálata, hogy a *Brüsszel* milyen raggal milyen gyakori. A ragtalan alanyi előfordulás ágéntív használatra utalhat, a tárgyi vagy helyhatározói előfordulás pedig páciensire. Az eredményeket a 3. táblázat mutatja be.

alak	kormányoldal	ellenzék	kormányoldal%	ellenzék%	%-arány
<i>Brüsszel</i>	910	426	24,4%	23,7%	~1
<i>Brüsszelben</i>	825	378	22,1%	21,0%	~1
<i>Brüsszelbe</i>	103	85	2,8%	4,7%	1,7
<i>Brüsszelt</i>	82	56	2,2%	3,1%	1,4
összesen	3727	1799			

3. táblázat. Különböző ragozott alakok gyakorisága. Ezen az absztrakt szinten inkább hasonlóságokat találunk. A *-ba/-be* és *-t* mutat eltérést.

Mit mutatnak meg a szélsőségesen az egyik oldalhoz kötődő kifejezések? Ebben a vizsgálatban a *Brüsszel* és a hozzá tartozó függőségi fej alkotta kétszavas egységeket vizsgáltam **csq1** és további elemzés segítségével. Először összegyűjtöttem ezeket a kétszavas egységeket, majd lekérdeztem a gyakoriságukat a kormányoldali, valamint az ellenzéki alkorpuszban, végül a kettő aránya szerint sorba rendeztem őket. Így a lista elején a kormányoldalra, a végén pedig az ellenzékre jellemző elemeket kaptam meg (4. táblázat).

sorszám	kifejezés	kormány	ellenzék	K/E	E/K
1	<i>brüsszeli elit</i>	40	–	∞	0,0
2	<i>Brüsszelből kap</i>	31	–	∞	0,0
3	<i>brüsszeli intézmény</i>	21	–	∞	0,0
4	<i>brüsszeli politikus</i>	30	2	15,0	0,1
5	<i>Brüsszelben dolgozik</i>	22	2	11,0	0,1
6	<i>brüsszeli politika</i>	128	12	10,7	0,1
7	<i>brüsszeli vezető</i>	19	2	9,5	0,1
8	<i>brüsszeli javaslat</i>	18	2	9,0	0,1
9	<i>Brüsszel tesz</i>	19	3	6,3	0,2
10	<i>brüsszeli döntéshozó</i>	22	4	5,5	0,2
11	<i>brüsszeli Bizottság</i>	31	6	5,2	0,2
12	<i>brüsszeli szankció</i>	139	27	5,1	0,2
13	<i>brüsszeli bürokrácia</i>	33	7	4,7	0,2
14	<i>brüsszeli bürokrata</i>	117	26	4,5	0,2
15	<i>brüsszeli döntés</i>	26	6	4,3	0,2
16	<i>brüsszeli terv</i>	13	3	4,3	0,2
17	<i>Brüsszelben van</i>	34	10	3,4	0,3
18	<i>Brüsszel akar</i>	10	3	3,3	0,3
19	<i>Brüsszelben képvisel</i>	39	12	3,3	0,3
20	<i>Brüsszelben tesz</i>	17	6	2,8	0,4
21	<i>Brüsszelből érkezik</i>	13	5	2,6	0,4
22	<i>Brüsszelbe küld</i>	13	5	2,6	0,4
23	<i>brüsszeli forrás</i>	18	8	2,3	0,4
24	<i>Brüsszelben mond</i>	16	10	1,6	0,6
25	<i>Brüsszelben dönt</i>	11	7	1,6	0,6
26	<i>Brüsszel ad</i>	12	9	1,3	0,8
27	<i>Brüsszel van</i>	10	8	1,3	0,8
28	<i>Brüsszelbe megy</i>	12	10	1,2	0,8
29	<i>Brüsszelnek ad</i>	9	8	1,1	0,9
30	<i>Brüsszel mond</i>	13	12	1,1	0,9
31	<i>Brüsszelben megszavaz</i>	10	11	0,9	1,1
32	<i>Brüsszelben aláír</i>	8	10	0,8	1,3
33	<i>Brüsszeltől vár</i>	7	13	0,5	1,9
34	<i>Brüsszelbe kimegy</i>	5	20	0,3	4,0
35	<i>Brüsszelt megállít</i>	4	17	0,2	4,3
36	<i>brüsszeli pénz</i>	4	24	0,2	6,0
37	<i>Brüsszelre mutogat</i>	–	31	0,0	∞

4. táblázat. A célszó és a hozzá tartozó függőségi fej alkotta (jelző+főnév vagy bővítmény+ige) kifejezések elemzése. A táblázat a legalább 20-szor előforduló kifejezéseket tartalmazza, a kormányoldalra jellemzőek fent, az ellenzékre jellemzőek lent helyezkednek el.

Az alábbiakban a zárójeles számokkal a 4. táblázat példaira utalok. A kormányoldalra jellemző kifejezések legtöbbször erős negatív kontextusban jelennek meg. A *brüsszeli elit* (1) *korrupt és bevándorlás párti*, *„lábön lőtte az európai gazdaságot”* és *„figyelmen kívül hagyta az állampolgárok döntését”*. A *brüsszeli intézmények* (3) *„példa nélküli lépésre készülnek”*, *„ránk akarnak erőltetni”* valamit és *„kétarcú politikát folytatnak”*. A *brüsszeli politikusok* (4) *baloldaliak*, *„nyíltan bevándorlás pártiak”* és *„vádolnak minket”*. Még a *Brüsszelben dolgozik* (5) is szinte mindig úgy jelenik meg a korpuszban, hogy a magyar ellenzék Magyarország érdeke ellen dolgozik Brüsszelben.

Az ellenzékre jellemző kifejezések viszont inkább semleges, érzelemmentes kontextusban fordulnak elő (*Brüsszelben aláír* [32], *Brüsszelbe kimegy* [34]), vagy a Fidesz–KDNP-re utalnak, a kormányoldali frázisokat idézik (*Állítsuk meg Brüsszelt!* [35]). A *Brüsszelre mutogat* viszont inkább pozitívan állítja be *Brüsszelt*, mint aki nem szolgált rá arra, hogy a bajokért őt hibáztassák. Nem véletlen, hogy ebben a kontextusban a *mutogat* csak harmadik személyben fordul elő, és a kormányoldal sohasem használja.

A kormányoldal jelentős számban használ olyan kifejezéseket, amelyek *Brüsszelenek* erőteljes szándékot tulajdonítanak, egyúttal kifejezik az ezzel a szándékkal való szembenállást, és ezt jellemzően érzelmileg fűtött szavakkal teszik. *Brüsszel* például *„a háború árát az európai emberekkel fizetteti meg”*, vagy *„nem tesz semmit”*, vagy *„szankciós lépéseket tesz”*, vagy *„hiába akar valamit elérni”*, vagy *„új adókat akar ránk kényszeríteni”*, vagy *„kényszerbetelepítést akar”*, de mi *„nem fogjuk beengedni őket, akárhogy is akarja Brüsszel”*.

5. Összefoglalás és tanulságok

Ebben a részben áttekintjük, hogy mit mondhatunk az 1. részben megfogalmazott hipotézisekről a 4. részben elvégzett vizsgálatok alapján.

A **H₁** hipotézisünk az volt, hogy *Brüsszel* a kormányoldal számára inkább aktív, szándékokkal bíró cselekvő ágens; az ellenzék számára pedig inkább passzív szereplő. A 3. táblázatban bemutatott egyszerű vizsgálat nem mutatta ezt ki, annyi látszott, hogy a passzívabb *-ba/-be* ragos, illetve tárgyragos forma kicsivel jellemzőbb az ellenzék megnyilvánulásaira. Amint láttuk, a kormányoldal sok olyan kifejezést használ, ami *Brüsszelenek* szándékot tulajdonít, ráadásul olyan szándékot, ami a kormányoldal szándékaival szemben áll, ami ellen „harcolni kell”, és ez okot ad az érzelemmel telített megfogalmazások használatára. Azt mondhatjuk, hogy a **H₁** kormányoldalra vonatkozó része igaznak bizonyult, és úgy tűnik, hogy a kormányoldal–ellenzék különbség nem az aktív–passzív tengelyen, hanem inkább az érzelmi telítettség skálán jelentkezik.

A **H₂** hipotézisünk az volt, hogy a kormányoldal hajlamos kommunikációs paneleket, előregyártott frázisokat használni, és ezeket számos alkalommal ismétli; az ellenzék nem él ezzel az eszközzel, viszont gyakran utal a kormányoldal paneleire kritikus hangon. Láttuk, hogy az *elhibázott*, a *brüsszeli* és a *szankció* szavak különféle kombinációiból képzett panel kiemelten gyakori a korpuszban: önmagában az előfordulások 2%-át adja, és 20-szor gyakoribb a kormányoldali alkorpuszban, mint az ellenzéki-

ben. Az ellenzéki megszólalók minden esetben a Fidesz–KDNP-re utalva gunyorosan és/vagy számonkérően ejtik ki a kifejezést (például: „*hát, egyetlenegy magyar ember szavazta meg ezeket az elhibázott brüsszeli szankciókat*”). Az egyetlen kivétel Toroczkai László, aki komolyan használja, egy alkalommal. A kifejezést 39 (!) különböző kormányoldali megszólaló használja, így az a sejtés is igaz, hogy nemcsak sokszor, hanem sokan is ismétlik ezt a panelt. Azt az ellenzéki oldalon népszerű vélekedést tehát, hogy a kormányoldal előre megfontolt szándékkal és összehangoltan folytatja ezt a fajta kommunikációt, a nyelvi vizsgálatunk alátámasztja. Az *Állítsuk meg Brüsszelt!* panel gyakrabban jelenik meg az ellenzéki kommunikációban a Magyar Országgyűlésben, ez összefügghet azzal, hogy a kormányoldal viszont plakátokon hirdette ezt az üzenetet az egész országban.

A **H₃** hipotézisünk szerint *Brüsszel* a kormányoldal számára inkább negatív, az ellenzék számára pedig inkább pozitív szereplő; a kormányoldal kommunikációjában *Brüsszel* általában ellenünk tesz valamit, az ellenzék kommunikációjában pedig inkább a mi érdekünkben. Fentebb láttuk, hogy a kormányoldalra jellemző kifejezések nagyon gyakran erős negatív kontextusban jelennek meg, és szinte mindig Magyarország-ellenes, a magyarok érdekeivel ellentétes élt kapnak. Az ellenzékre jellemző kifejezések ugyanakkor többnyire semlegesek, néhány kifejezés – mint például a *Brüsszelre mutogat* – pedig ha közvetetten is, de határozottan pozitív szereplőként hivatkozik *Brüsszelre*. A másik oldal *Brüsszel*hez való viszonyának percepciója megjelenik olyan alakulatokban is, mint az *Orbán–Brüsszel-háború*. A **H₃** esetében is igaz, hogy a kormányoldal tekintetében teljesen helytállóan bizonyult a hipotézis, az ellenzék tekintetében viszont csak részben. Azonban az így is egyértelmű, hogy *Brüsszel* megítélésében markánsan eltér egymástól a két oldal.

Irodalom

- Ligeti-Nagy, N., Dodé, R., Jelencsik-Mátyus, K., Varga, Zs., Héja, E., & Váradi, T. (2023). A new ParlaMint corpus for Hungarian – 30m tokens of annotated parliamentary data. In Berend G., Gosztolya G. & Vincze V. (szerk.) *XIX. Magyar Számítógépes Nyelvészeti Konferencia* (pp. 433–446). Szeged: SZTE TTIK, Informatikai Intézet.
- Rychlý, P. (2007). Manatee/Bonito – A Modular Corpus Manager. In *1st Workshop on Recent Advances in Slavonic Natural Language Processing* (pp. 65–70). Brno: Masaryk University. <https://nlp.fi.muni.cz/trac/noske>.
- Sass B. (2017). Keresés korpuszban: a kibővített Magyar történeti szövegtár új keresőfelülete. In Forgács T., Németh M. & Sinkovics B. (szerk.) *A nyelvtörténeti kutatások újabb eredményei IX.* (pp. 267–277). Szeged: SZTE Magyar Nyelvészeti Tanszék.
- Sass, B. (2022). Principles of corpus querying: A discussion note. *Acta Linguistica Academica*, 69(4) (pp. 599–614).