

Sketch Engine

Bevezetés

A számítógéppel olvasható és elemezhető korpuszok a nyelvtudomány több területén számos új kutatási területet és lehetőséget nyitottak meg. Közülük például a korpuszalapú fordítástudományi kutatások megjelenése Magyarországon az 1990-es évekre tehető (Klaudy 2005), elterjedésükről azonban a 2000-es évek elejétől beszélhetünk, ami a technológia fejlődésének köszönhető (Robin & Seidl-Péché, 2020). Az ilyen típusú vizsgálatok jelentősége abban rejlik, hogy a szövegmenyiség méretének köszönhetően statisztikailag is megalapozott következtetések vonhatók le a nyelvhasználatra vonatkozóan. Bár a kutatóknak több ingyenes, vagy részben ingyenes eszköz áll rendelkezésükre a korpuszok elemzésének elvégzésére, ezeket gyakran csak bizonyos korlátozásokkal használhatjuk – pl. az annotálás vagy a lemmatizálás hiánya miatt (Robin et al., 2016; Seidl-Péché & Robin, 2019).

Jelen ismertető a Sketch Engine (<https://www.sketchengine.eu/>) nevű eszközt mutatja be, amely lehetővé teszi a nyelvészeti korpuszkutatások során jellemzően használt elemzések elvégzését.

Sketch Engine

A Sketch Engine korpuszelemző eszköz, melynek segítségével különféle kereséseket és műveletek végezhetünk saját és az adatbázisban rendelkezésre bocsátott szöveggyűjteményeken. A kínált lehetőségek – amelyek között a lemmatizálás automatikus elvégzése is szerepel – hasznosak lehetnek nyelvészek, fordítók, oktatók és diákok számára is, hiszen a korpuszok elemzése során számos olyan információhoz juthatunk, amelyek a szótárakból és glosszáriumokból hiányoznak. A használathoz előfizetés szükséges, melynek díja az előfizetés típusától függ: a felsőoktatásban hallgatók és dolgozó kutatók – az ismertető megírásának pillanatában – egyénileg havi bruttó 10,66 euróért juthatnak licenchez, azonban lehetőség van éves fizetési ütemezésre, illetve csoportos előfizetésre is, melynek köszönhetően az egy főre eső havidíj mérséklődik. Amennyiben nem hallgatóként vagy kutatóként, hanem egyénileg szeretnénk előfizetni, a havi bruttó alapdíj 19,99 euró, azonban ez esetben is lehetőségünk van a fizetési gyakoriságtól függően alacsonyabb összegért hozzájutni a szolgáltatáshoz. Az említett díjak a rendszer használatán túl magukban foglalnak egymillió szónyi tárhelyet arra az esetre, ha saját korpuszokkal is dolgozni szeretnénk (amennyiben ettől nagyobb méretű a korpuszunk, úgy további tárhely vásárlása szükséges és lehetséges). Mielőtt azonban fizetnénk a használatért, a Sketch Engine összes funkcióját ingyen kipróbálhatjuk 30 napon át.

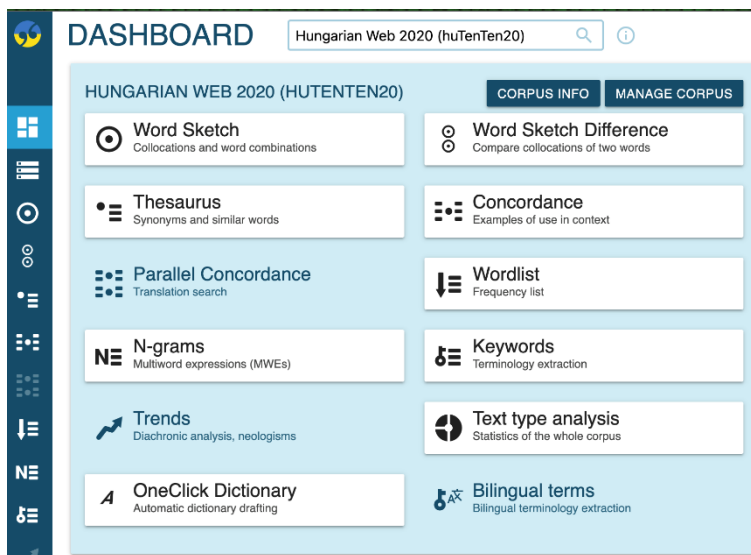
A Sketch Engine segítségével végezhető elemzések

A regisztrációhoz csupán néhány alapvető adat (név, e-mail-cím, munkahely, a használat célja, számlázási adatok) megadása szükséges, majd néhány percen belül megérkezik postafiókunkba a rendszer használatához szükséges jelszó. Bejelentkezés után a korpusz kiválasztása (*Select Corpus*) nézet jelenik meg. Itt a nyelv megadása után lehetőségünk van az adott nyelven elérhető bármely korpuszt kiválasztani. A rendszerben számos nyelv és az egyes nyelveken gyakran több korpusz is hozzáférhető. Az elérhető nyelvek, az azokhoz tartozó korpuszok és azok jellemzői (például annotált, lemmatizált) a rendszer súgójában található táblázatban tekinthetők meg (*Languages in Sketch Engine*¹). A magyar nyelvű szövegek tanulmányozását jelenleg 21 különböző korpusz segítségével végezhetjük, melyek között megtalálhatók általános (így internetről gyűjtött) és szűkebb területről származó (például európai uniós) gyűjtemények egyaránt.

A nyelv kiválasztása után a *Dashboard* felületre érkezünk, amely tulajdonképpen egyfajta vezérlőpult – innen érhetők el az egyes műveletek, amelyek a kiválasztott korpuszon elvégezhetők.

1. ábra

A Sketch Engine kezdőlapja



Szóvázat – a szavak nyelvtani és kollokációs viselkedése (Word Sketch)

A Word Sketch funkció a szavak kollokációjának vizsgálatára alkalmas, használatához csupán a keresett lemmát kell megadnunk, majd rövid időn belül a rendszer kilistázza az adott szó környezetével kapcsolatos adatokat. Ennek segítségével meg-

¹ <https://www.sketchengine.eu/corpora-and-languages/>

tudhatjuk például, hogy milyen igékkel, névszókkal, névelőkkel, névutókkal fordul elő leggyakrabban az adott kifejezés, de a megjelenő összesítés egyes listáit ki is bonthatjuk, így még több információhoz juthatunk az éppen vizsgált jelenséggel kapcsolatban. A keresést optimalizálhatjuk úgy is, ha a lemma megadása után az *advanced* fülön pontosítjuk a beállításokat (*például* szófaj megadása). A 2. ábrán az *egészség* szóra keresve megjelenő összefoglaló látható.

2. ábra

A szóvázlát (*Word Sketch*) művelet eredménye az *egészség* szóra keresve

modifiers of "egészség"	verbs with "egészség" as object	verbs with "egészség" as subject	"egészség" and/or ...
mentális mentális egészség	veszélyeztet egészséget veszélyeztet	a-z Összes cikk Egészség A-Z témában: Az	biztonság egészség és biztonság
lelki lelki egészség	örvend jó egészségre örvend	cenzuráz Cenzurázott egészség	szépség Szépség és egészség
gyakorolt egészségre gyakorolt	károsít károsítja az egészséget	helyreáll az egészség helyreáll	környezet egészségre és a környezetre
ártalmas az egészségre ártalmas	megőríz megőrizzük egészségünket	romlik az egészség romlik	gazda növekednek, a gazda szervezet egészségét is befolyásolhatják
káros az egészségre káros	kíván jó egészséget kívánok	felelősködik felelősködik az egészségre	jólét egészség és jólét
testi-lelki testi-lelki egészség	vigyáz vigyázom az egészségére	lefog rdi egészségre lefog	megőrzés egészségének megőrzésében
testi testi egészség	óv egészségét óvja	összefügg egészség szorosan összefügg	közérzet egészség és a jó közérzet
csont csontok egészségének	árt árt az egészségnek	javul egészség javul	boldogság egészség és boldogság
bőr bőr egészségének	kockáztat egészségét kockáztatja	megromlik az egészségre megromlik	épség épséget vagy egészséget
emberi emberi egészségre	javít javítja az egészséget	megtestesül a megtestesült egészség	élet egészségének és életének
gazda növekednek, a gazda szervezet egészségét is befolyásolhatják	befolyásol befolyásolja az egészséget	károsodik az egészségre károsodik	érdek egészségének és életének megővése érdekében
fog fogak egészségének	viszanyer viszanyerte egészségét	megér emnyit megír az egészségre	jólét egészsége és jóléte

Tezaurusz (*Thesaurus*)

A Tezaurusz szógyűjtemény a megadott lemmával egy szemantikai mezőbe tartozó kifejezéseket gyűjt össze. Ahhoz, hogy az összeállított lista megbízható legyen, nagy méretű korpuszokra van szükség. A fejlesztői ajánlások szerint a minimális méret kb. százezer szó, a megbízható működéshez azonban milliárdos nagyságrendű szószámot javasolnak – emiatt célszerű lehet e funkció esetében a beépített korpuszokkal dolgozni, ha a saját szöveggyűjteményünk mérete ettől elmarad.

Párhuzamos konkordancia (*Parallel Concordance*)

A párhuzamos konkordancia vizsgálatához párhuzamos korpuszokra van szükség. A művelet segítségével arra kaphatunk választ, hogy bizonyos szavakat vagy kifejezéseket hogyan fordítottak le más nyelvre. Az 3. ábrán látható, hogy a német *Parlament* szót tartalmazó szövegrészeket hogyan fordították magyarra.

3. ábra

A Parlament szó fordításainak listája a párhuzamos konkordancia segítségével

The screenshot shows the PARALLEL CONCORDANCE interface. The search query is "EUR-Lex 22016 parallel - German". The results are displayed in a table with columns for the source text (German) and the target text (Hungarian). The results are numbered 1 through 10, corresponding to the list items in the image.

Source (German)	Target (Hungarian)
(1) Sieht ein nach Teil V Kapitel 2 des Vertrags über die Europäische Union erlassener Beschluss die Aussetzung, Einschränkung oder vollständige Einstellung der Wirtschafts- und Finanzbeziehungen zu einem oder mehreren Drittländern vor, so erlässt der Rat die erforderlichen Maßnahmen mit qualifizierter Mehrheit auf gemeinsamen Vorschlag des Hohen Vertreters der Union für Außen- und Sicherheitspolitik und der Kommission.	(1) Ha az Európai Unióról szóló szerződés V. címének 2. fejezete szerint elfogadott valamely határozat az egy vagy több harmadik országgal fennálló gazdasági és pénzügyi kapcsolatok részbeni vagy teljes megakadályozására vagy korlátozására írja elő, a Tanács az Unió külügyi és biztonságpolitikai főképviselője és a Bizottság által előterjesztett közös javaslat alapján, mindkettőt többszempes elfogadja a szükséges intézkedéseket.
Er unterrichtet hierüber das Europäische Parlament.	Erőlti tájékoztatja az Európai Parlamentet.
Der Gerichtshof der Europäischen Union überwacht die Rechtmäßigkeit der Gesetzgebungsakte sowie der Handlungen des Rates, der Kommission und der Europäischen Zentralbank, soweit es sich nicht um Empfehlungen oder Stellungnahmen handelt, und der Handlungen des Europäischen Parlaments und des Europäischen Rates mit Rechtswirkung gegenüber Dritten.	Az Európai Unió Bírósága megvizsgálja a jogalkotási aktusok jogszerűségét, valamint a Tanács, a Bizottság és az Európai Központi Bank jogi aktaikat, kivéve az ajánlásokat és a véleményeket, továbbá az Európai Parlament és az Európai Tanács harmadik személyekre joghatással járó aktaikat.
Er überwacht ebenfalls die Rechtmäßigkeit der Handlungen der Einrichtungen oder sonstigen Stellen der Union mit Rechtswirkung gegenüber Dritten.	Az Európai Unió Bírósága megvizsgálja továbbá az Unió szervei vagy hivatalai által elfogadott, harmadik személyekre joghatással járó jogi aktusok jogszerűségét.
Zu diesem Zweck ist der Gerichtshof der Europäischen Union für Klagen zuständig, die ein Mitgliedstaat, das Europäische Parlament, der Rat oder die Kommission wegen Unzuständigkeit, Verletzung wesentlicher Formvorschriften, Verletzung der Verträge oder einer bei seiner Durchführung anzuwendenden Rechtsnorm oder wegen Ermessensmissbrauchs erhebt.	E célból az Európai Unió Bírósága hatáskörrel rendelkezik az olyan keresetek tekintetében, amelyeket valamely tagállam, az Európai Parlament, a Tanács vagy a Bizottság nyújt be hatáskör hiánya, lényeges eljárási szabályok megsértése, a Szerződéses vagy az alkalmazására vonatkozó bármely jogi rendelkezés megsértése vagy hatáskörrel való visszaélés miatt.
Das Europäische Parlament und der Rat beschließen gemäß dem ordentlichen Gesetzgebungsverfahren die auf dem Gebiet der sozialen Sicherheit für die Herstellung der Freizügigkeit der Arbeitnehmer notwendigen Maßnahmen; zu diesem Zweck führen sie insbesondere ein System ein, das zu- und abwandernden Arbeitnehmer und Selbstständigen sowie deren anspruchsberechtigten Angehörigen Folgendes sichert:	Az Európai Parlament és a Tanács rendes jogalkotási eljárás keretében a szociális biztonság területén elfogadja azokat az intézkedéseket, amelyek a munkavállalók szabad mozgásának biztosításához szükségesek; e célból olyan eszközrendszert hoz létre, amely a migráns munkavállalók és önról vállalkozók, valamint az azok jogaín jogosultak számára biztosítja
(2) Das Europäische Parlament und der Rat legen gemäß dem ordentlichen Gesetzgebungsverfahren und nach Anhörung des Wirtschafts- und Sozialausschusses die gemeinsame Organisation der Agrarmärkte nach Artikel 40 Absatz 1 sowie die anderen Bestimmungen fest, die für die Verwirklichung der Ziele der gemeinsamen Agrar- und Fischereipolitik notwendig sind.	(2) A 40. cikk (1) bekezdésében előírt közös piacszervezést, továbbá a közös agrárpolitika és a közös halászati politika célkitűzéseinek megvalósításához szükséges egyéb rendelkezéseket a Gazdasági és Szociális Bizottsággal folytatott konzultációt követően rendes jogalkotási eljárás keretében az Európai Parlament és a Tanács állapítja meg.

N-gramok (N-grams)

Az N-gramok, azaz a többszavas kifejezések funkció segítségével a nyelvben előforduló többszavas kifejezéseket gyakoriságát vizsgáljuk. A keresés elindításakor lehetőségünk van beállítani, hogy hány szóból álló kifejezésekre van szükségünk, de természetesen további keresési feltételeket is megadhatunk (pl. a keresési eredmény tartalmazzon bizonyos szavakat). A 4. ábrán olyan, három vagy négy tagból álló kifejezések láthatók, amelyek tartalmazzák a *néz* igét.

4. ábra

Három vagy négy tagból álló kifejezések, amelyek tartalmazzák a *néz* igét

The screenshot shows the N-GRAMS interface. The search query is "Hungarian Web 2020 (huTenTen20)". The results are displayed in a table with columns for the N-gram and its frequency. The results are numbered 1 through 30, corresponding to the list items in the image.

N-gram	Frequency	N-gram	Frequency	N-gram	Frequency
1 úgy néz ki	11,470	11 Így néz ki	821	21 Hogyan néz ki	515
2 néz ki a	3,875	12 és úgy néz	608	22 nagyon jól néz ki	512
3 jól néz ki	3,317	13 Hogy néz ki	588	23 de úgy néz ki	511
4 Úgy néz ki	2,511	14 jobban néz ki	549	24 néz ki és	496
5 Így néz ki	2,154	15 és úgy néz ki	533	25 hogy néz ki a	478
6 hogy néz ki	1,753	16 nagyon úgy néz	526	26 nem néz ki	477
7 is néz ki	1,308	17 nagyon úgy néz ki	522	27 nem úgy néz	476
8 néz ki egy	1,152	18 de úgy néz	520	28 úgy néz ki a	461
9 hogyan néz ki	1,093	19 néz ki mint	518	29 nem úgy néz ki	456
10 néz ki az	1,079	20 nagyon jól néz	518	30 néz ki ez	415

You are only allowed to access 1,000 items. [Get more](#)

Trendek (Trends)

A trendek opció a diakrón vizsgálatok során jelenthet segítséget, a használat gyakoriságának változásairól adnak felvilágosítást. A funkció a működéséből adódóan csak időbélyeggel ellátott korpuszokon használható. Lehetőség van egy-egy konkrét szó vagy kifejezés vizsgálatára, alapbeállítás mellett pedig azokat az elemeket listázza a rendszer, amelyek a vizsgált időszakban a legnagyobb – akár pozitív, akár negatív – változáson mentek keresztül.

Szóvázatok eltérései (Word Sketch Difference)

A Word Sketch Difference, azaz a szóvázatok eltérései funkció akkor segít, ha két lemma kollokációját szeretnénk összehasonlítani, ezáltal jobban meg szeretnénk ismerni a kettő közötti különbségeket. A keresési eredmények nagyjából a Word Sketch lehetőségnél megjelenőkkel azonosak azzal az eltéréssel, hogy az egyes szavak mellett két számérték jelenik meg. Az első számérték az adott találat és az első keresett szó, a második számérték az adott találat és a második keresett szó együttes előfordulásának gyakoriságát mutatja. Az 5. ábrán a *megy* és a *jár* igék eltérő kollokációi láthatók. Az ábrából az is kiderül, hogy a *megy* szó jelentősen gyakrabban fordul elő a *játék* szó mellett (*megy a játék*), míg a *jár* és a *köszönet* együttes előfordulása sokkal jellemzőbb, mint a *megy* és a *köszöneté*.

5. ábra

A *megy* és a *jár* ige kollokációinak összehasonlítása

The screenshot shows the 'WORD SKETCH DIFFERENCE' tool interface. At the top, it displays 'Hungarian Web 2020 (huTenTen20)' and search filters for 'megy' (4,454,495x) and 'jár' (2,982,335x). Below are three panels showing concordance data for 'megy/jár' and/or ...', 'objects of "megy/jár"', and 'subjects of "megy/jár"'. The data is presented in tables with columns for the word and its frequency.

"megy/jár" and/or ...				objects of "megy/jár"				subjects of "megy/jár"			
jön	5,323	219	...	rovás	14,727	0	...	játék	13,767	383	...
megy	4,398	283	...	csőd	14,690	0	...	változás	3,845	189	...
marad	1,264	82	...	nyugdíj	17,123	66	...	dolog	21,723	1,782	...
csinál	839	80	...	férj	31,784	330	...	min	3,925	373	...
megnéz	781	162	...	irány	16,036	624	...	munka	3,725	455	...
van	5,052	2,445	...	változás	16,933	2,114	...	s	6,989	3,006	...
lát	738	645	...	él	5,619	19,291	...	idő	3,440	9,064	...
szeret	527	704	...	iskola	11,937	43,137	...	pont	2,266	8,759	...
kel	306	316	...	út	16,238	63,797	...	hír	333	5,611	...
beszél	433	779	...	hely	7,380	31,015	...	büntetés	39	3,110	...
olvas	176	350	...	siker	718	47,623	...	kedvezmény	0	3,477	...
jár	283	2,050	...	következmény	0	35,545	...	köszönet	0	4,117	...

Konkordancia (Concordance)

A konkordanciavizsgálat segítségével szavak vagy kifejezések környezetét vizsgálhatjuk. A kereséshez meg kell adnunk egy lemmát és a rendszer kilistázza a megadott szó

(vagy kifejezés) konkrét előfordulásait a szövegben a környezettel együtt. Az 6. ábra a *kérdez* lemma előfordulásait, illetve azok forrásait mutatja be.

6. ábra

A *kérdez* lemma előfordulásai

The screenshot shows the CONCORDANCE tool interface. At the top, it displays 'Hungarian Web 2020 (huTenTen20)' and a search bar containing 'kérdez'. Below the search bar, there are navigation icons and a 'SUBSCRIBE' button. The main area shows a list of 14 search results, each with a source URL and a KWIC snippet. The KWIC snippets show the word 'kérdez' in context, often with a question mark or a colon. The results are sorted by frequency, with the first result being 'emzeti Közzolgálati Egyetem Hadtudományi és Honvédtisztképző Kar kibéharboris szakértjétől arról kérdeztük, hogy fenyegeti-e az internetes háború Magyarországot?</p></div>
<div data-bbox="93 398 272 418" data-label="Section-Header>
<h2>Szólista (Wordlist)</h2>
</div>
<div data-bbox="93 421 906 512" data-label="Text>
<p>A szólista (Wordlist) funkció segítségével különböző szempontok szerint összeállított gyakorisági listát hozhatunk létre. A keresési szempontok között megadhatunk szófajra vonatkozó kritériumokat, tartalmazandó karaktersorozatot, konkrét lemmákat stb. A 7. ábrán látható lekérdezésben azok a szavak jelennek meg, amelyek a <i>hoz</i> karaktersorozattal kezdődnek.</p>
</div>
<div data-bbox="93 528 153 545" data-label="Section-Header>
<h2>7. ábra</h2>
</div>
<div data-bbox="93 545 406 563" data-label="Text>
<p>A <i>hoz</i> karaktersorozattal kezdődő szavak</p>
</div>
<div data-bbox="93 568 895 832" data-label="Complex-Block>

<p>The screenshot shows the WORDLIST tool interface. At the top, it displays 'Hungarian Web 2020 (huTenTen20)' and a search bar containing 'hoz'. Below the search bar, there are navigation icons and a 'SUBSCRIBE' button. The main area shows a list of 30 words starting with 'hoz', sorted by frequency. The words and their frequencies are as follows:</p>
<table border='1'>
<thead>
<tr>
<th>Word</th>
<th>Frequency ? ↓</th>
<th>Word</th>
<th>Frequency ? ↓</th>
<th>Word</th>
<th>Frequency ? ↓</th>
</tr>
</thead>
<tbody>
<tr><td>1 hozzá</td><td>2,239,655 ...</td><td>11 hozzájuk</td><td>212,647 ...</td><td>21 hozzászólására</td><td>79,193 ...</td></tr>
<tr><td>2 hozott</td><td>527,890 ...</td><td>12 hozzászólások</td><td>191,767 ...</td><td>22 hoztam</td><td>78,239 ...</td></tr>
<tr><td>3 hozta</td><td>335,734 ...</td><td>13 hozták</td><td>149,911 ...</td><td>23 hozzávalók</td><td>77,155 ...</td></tr>
<tr><td>4 hozzátette</td><td>293,191 ...</td><td>14 hoztak</td><td>144,371 ...</td><td>24 hozzájárult</td><td>75,603 ...</td></tr>
<tr><td>5 hoz</td><td>292,815 ...</td><td>15 hozzájárul</td><td>130,108 ...</td><td>25 hozzátéve</td><td>63,570 ...</td></tr>
<tr><td>6 hozza</td><td>284,257 ...</td><td>16 hozzáférés</td><td>115,815 ...</td><td>26 hozhat</td><td>58,344 ...</td></tr>
<tr><td>7 hozzánk</td><td>275,498 ...</td><td>17 hozzád</td><td>111,915 ...</td><td>27 hozzáférést</td><td>55,989 ...</td></tr>
<tr><td>8 hozzám</td><td>274,547 ...</td><td>18 hozzák</td><td>104,311 ...</td><td>28 hozzáadott</td><td>54,625 ...</td></tr>
<tr><td>9 hozni</td><td>261,823 ...</td><td>19 hoznak</td><td>99,418 ...</td><td>29 hozzon</td><td>52,509 ...</td></tr>
<tr><td>10 hozzászólás</td><td>261,476 ...</td><td>20 hozzáadjuk</td><td>83,329 ...</td><td>30 hozzájárulás</td><td>52,177 ...</td></tr>
</tbody>
</table>
<p>You are only allowed to access 1,000 items. Get more</p>
</div>

Kulcsszavak (Keywords)

A kulcsszavak funkció tulajdonképpen automatikus terminológiakivonatolás, melynek során az elemző bizonyos algoritmusok segítségével igyekszik kiszűrni a terminusok kritériumainak megfelelő szavakat a korpuszból. A rendszer az egy és a több szóból álló terminusok kivonatolására is képes. A 8. ábrán látható lekérdezési eredmény a német nyelvű EUR-Lex korpuszból automatikusan kigyűjtött terminusokat tartalmazza.

8. ábra

A német nyelvű EUR-Lex korpuszból automatikusan kivonatolt terminusok

KEYWORDS

EUR-Lex 2/2016 parallel – German



SINGLE-WORDS ✓

MULTI-WORD TERMS ✓



reference corpus: German Web 2020 (deTenTen20) (items: 8,640,514)

Term	Term
1 europäische Union ...	14 Kommission der Europäischen Gemeinschaften ...
2 europäische Gemeinschaft ...	15 vorliegender Fall ...
3 europäisches Parlament ...	16 betreffender Mitgliedstaat ...
4 zuständige Behörde ...	17 Gründung der Europäischen Gemeinschaft ...
5 Amtsblatt der Europäischen Union ...	18 folgende Fassung ...
6 anderer Mitgliedstaat ...	19 Entscheidung der Kommission ...
7 vereinigtés Königreich ...	20 Eg des Europäischen Parlaments ...
8 Ewg des Rates ...	21 Vorschlag der Kommission ...
9 nachstehender Grund ...	22 angefochtene Entscheidung ...
10 Erwägung Nachstehender Gründe ...	23 folgende Verordnung ...
11 europäische Kommission ...	24 vorliegende Verordnung ...
12 staatliche Beihilfe ...	25 vereinigter Staat ...
13 Name der Kommission ...	26 hergestellte Ware ...

Szövegtípus-elemzés (Text type analysis)

A szövegtípus-elemzés a korpusz metaadatairól nyújt felvilágosítást. Segítségével többek között azt is megtudhatjuk, hány dokumentumból áll a korpuszunk, mekkora a terjedelme, milyen forrásból származnak a szövegek, melyik évben jelentek meg, de számos egyéb szűrési és keresési opció is rendelkezésre áll. A megjelenő eredmények száma és részletessége a korpusz annotáltságától függ.

Összegzés

Az ismertető a Sketch Engine nevű korpuszelemző alkalmazás legfontosabb funkcióit mutatta be. Fontos azonban megjegyezni, hogy a leírások nem mindenre terjednek ki, csupán az egyes funkciókban rejlő alapvető lehetőségekre fókuszálnak. A bemutatott műveletek mindegyike tovább pontosítható – ennek részleteiről a rendszer súgó-jában találhatunk információkat –, így némi gyakorlás után a különböző szűrési és keresési beállítások finomhangolásával precízen összeállíthatjuk az aktuális kutatáshoz vagy bármilyen kereséshez szükséges pontos lekérdezést.

IRODALOM

- Klaudy, K. (2005). Párhuzamos korpuszok felhasználása a fordításkutatásban. In Lanstyák, I. & Vančoné Kremmer, I. (szerk.), *Nyelvészetről – változatosan. Segédkönyvek egyetemisták és a nyelvészet iránt érdeklődők számára* (pp. 153–185.) Gramma Nyelvi Iroda.
- Robin, E., Dankó, Sz., Götz, A., Nagy A., L., Pataky, É., Szegh, H. & Zolczer, P. (2016). Fordítástudomány és korpuszkutatás: bemutatkozik a Pannonia Korpusz. *Fordítástudomány*, 18(2), 5–26.
- Robin, E. & Seidl-Péché, O. (2020). Korpuszalapú fordításkutatás Magyarországon. In Robin, E. & Seidl-Péché, O. (szerk.), *Fókuszban a fordított és a tolmácsolt szöveg: Korpuszalapú fordításkutatás Magyarországon* (pp. 9–24.) ELTE BTK Fordítástudományi Doktori Program, MANYE Fordítástudományi Szakosztály. <https://doi.org/10.36252/Nyelvikozvsegedkonyv1.1>
- Seidl-Péché, O. & Robin, E. (2019). Alkalmas-e a korpuszalapú módszer a terminuskezelési stratégiák kutatására? Dróth, J. (szerk.), *Korpusz és kontrasztivitás a szakfordítás oktatásában és gyakorlatában* (pp. 93–104.) Károli Gáspár Református Egyetem, L'Harmattan Kiadó.

Szlávik Szilárd