

Kocsis Ágnes

Informatikai eszközök a nyelvészeti kutatásban

Segédanyag az informatikai eszközök használatához
kognitív nyelvészeknek

Kocsis Ágnes

Informatikai eszközök a nyelvészeti kutatásban

Segédanyag az informatikai eszközök használatához kognitív nyelvészeknek

ELTE BTK Nyelvtudományi Doktori Iskola

2022

A kiadvány az Információs és Technológiai Minisztérium ÚNKP-19-3 kódszámú
Új Nemzeti Kiválóság Programjának szakmai támogatásával készült.



© Kocsis Ágnes, 2022

ISBN 978 963 489 490 2 (pdf)



ELTE | BTK
BÖLCSESZETTUDOMÁNYI KAR

Felelős kiadó: az ELTE Bölcsészettudományi Kar dékánja

Projektvezető: Csanádi-Egresi Nóra

Kiadói szerkesztő: Tihanyi Katalin

Borító, tördelés: Balázs Andrea

Köszönetnyilvánítás

Köszönöm témavezetőmnek, Kugler Nórának, hogy hitt abban, hogy ez a munka megszülethet, és javasolta, hogy gyűjtssem össze ilyen formában a korábbi oktatói munkámból és informatikatanári tevékenységemből származó tapasztalataimat, valamint erőt adott a segédanyag kivitelezéséhez.

Köszönet illeti Simon Gábort is alapos lektori munkájáért, amely lehetővé tette, hogy a kész segédanyag koherensebb és átláthatóbb legyen. Rengeteget tanultam javaslataiból, jegyzeteiből, amelyek későbbi munkám során is hasznosak lesznek, és alapvetően formálták szemléletmódomat.

Köszönöm továbbá az Eötvös Loránd Tudományegyetemnek, és az ELTE Eötvös Kiadónak, hogy lehetővé teszi e munka megjelenését.

Kocsis Ágnes

Tartalom

1. Előszó	6	4. Informatikai eszközök a nyelvészeti kutatások szolgálatában	27
2. A kognitív nyelvészeti kutatások	9	4.1. A keresés és lekérdezés módszertana	27
2.1. A funkcionális kognitív nyelvészet	9	4.1.1. Keresés a weben	27
2.2. A kognitív nyelvészeti kutatások folyamata	11	4.1.2. Reguláris kifejezések	45
2.3. A kognitív nyelvészeti kutatások során alkalmazott adattípusok	11	4.2. Adatbekérő kérdőívek készítése	52
2.4. Introspekció és empirikus adatok viszonya	13	4.2.1. Google Űrlapok	52
3. Empirikus kutatási módszerek a funkcionális kognitív nyelvészetben	16	4.2.2. Online-kerdoiv.com	57
3.1. Gyakori adatgyűjtési módszerek	16	4.3. Táblázatkezelő használata	60
3.1.1. Korpuszhasználat	17	4.4. Adatvizualizáció	75
3.1.2. Kérdőíves felmérések és interjúk	22	4.4.1. Diagramok, gráfok	75
3.2. Egy gyakori kvantitatív adatelemzési módszer: a statisztikai elemzés	24	4.4.2. Infografikák, idővonalak	79
		5. Kutatás és eszközválasztás	80
		6. Összegzés, kitekintés	83
		7. Bibliográfia	84
		8. Melléklet	88

1. Előszó

A kognitív nyelvészet területén kevés összefoglaló kutatómódszertani mű létezik, kiváltképp magyar nyelven. Az összefoglaló munkák inkább elméletiek, mintsem gyakorlatiak, a kutatók által alkalmazott módszerekről, felhasználásuk sokszínűségéről pedig többnyire a kutatási eredményeket közlő tanulmányok módszertani fejezeteiből értesülhetünk. Az egyes tanulmányok azonban átfogó képet ritkán adnak a különböző módszerek, adatok és eszközök általános használati lehetőségeiről. Ez nem is lehet céljuk, azonban a rengeteg kutatási módszer és azok adaptálásának sokasága között szükség van olyan átfogó művekre, amelyek vezérfonalul szolgálhatnak a kutatók számára, felsorolják és bemutatják a módszertani és eszközhasználati lehetőségeket, azok alkalmazási módjait és korlátait.

Jelen kutatómódszertani segédanyag ezt a hiányt igyekszik pótolni a számítógépes kutatástámogatás nézőpontjából: a funkcionális kognitív nyelvtudomány kérdésköreivel foglalkozó empirikus kutatások megalapozásához kínál segítséget, kiemelve azokat a technikai és informatikai lehetőségeket, amelyek megkönnyíthetik a kutatók munkáját. A kötet célja, hogy támogassa a nyelvészeket abban, hogy minél hatékonyabban ki tudják használni az informatikai eszközök és megoldások adta lehetőségeket. A nyelvészeti kutatások során azért alkalmazunk informatikai eszközöket, hogy azok könnyebbé és hatékonyabbá tegyék a kutatómunkát, lerövidítsék az automatizálható folyamatokat, képessé tegyenek minket nagy mennyiségű adat feldolgozására. Fontos kiemelni, hogy az eszközhasználat nem lehet a kutatások célja és nem is szükséges minden kutatáshoz informatikai eszközöket alkalmazni. Ezért csak olyan alkalmazásokat és eljárásokat mutatok be, amelyek hozzájárulhatnak az adatgyűjtési és adatfeldolgozási folyamatok hatékonyságához és javíthatják a kutatás infrastruktúráját.

A segédanyag első fele (**2–3. fejezetek**) a funkcionális kognitív kutatásokkal és az azokra jellemző adatgyűjtési, valamint adatelemzési módszerekkel foglalkozik. Nem ad az egyes módszerekről részletes technikai ismertetőt, inkább a bennük rejlő lehetőségekre hívja fel a figyelmet. E fejezetek tárgyalják a korpuszhasználatot, a kérdőíves és interjú adatfelvételi formákat, valamint röviden kitérnek a statisztikai elemzésekre.

A következő rész (**4. fejezet**) a felsorolt módszerek kivitelezése során használható informatikai eszközöket és eljárásokat ismerteti. Ebben a fejezetben is kritikai viszonyulás érvényesül: a technikai lehetőségeket és funkciókat bemutató szakaszok a programok hiba-lehetőségeire is felhívják a figyelmet. A rövid ismertetők példákon keresztül mutatják be a szoftverek funkcióit, azoknak a kutatásokban való hasznosíthatóságát. A példák többsége konkrét kutatási helyzetekből származik, mint például a szakirodalom-keresés, a nyelvi adatok szűrése, az adatközlők adatainak azonosítóval való ellátása. Az egyes informatikai eszközök és informatikai megoldások kiválasztásának szempontjai a következők voltak:

- ¶ hatékonyan lehessen alkalmazni a kutatás során;
- ¶ megkönnyítse és felgyorsítsa a kutatás folyamatát;
- ¶ amennyiben szoftverről van szó, legyen könnyen hozzáférhető és lehetőleg ingyenes;
- ¶ ne legyen túl bonyolult a használata;
- ¶ sokoldalúan lehessen alkalmazni.

Ezen szempontok alapján a következő eszközök és eljárások kerültek be a segédanyag **4. fejezetébe**:

- ¶ webes keresők;
- ¶ reguláris kifejezések alkalmazása;
- ¶ online kérdőívkészítő alkalmazások;
- ¶ táblázatkezelők;
- ¶ adatvizualizáló eszközök.

Fontos szem előtt tartani, hogy az egyes szoftverek gyorsan változnak, hamar elévülnek, esetleg megszűnnek vagy megszűnik támogatottságuk, és újabbak lépnek a helyükre. Ezért a felhasználónak érdemes nemcsak a konkrét programokkal, hanem azok típusával is megismerkednie, így később könnyebben tud adaptálódni más hasonló szoftverek használatához is.

A segédanyag **melléklete** egy programgyűjtemény, amely összefoglalja az ismertetett programtípusokat és szoftvereket, továbbá olyan eszközöket és alkalmazásokat is tartalmaz, amelyek hasznosak lehetnek ugyan a kutatás szempontjából, mégsem kerültek az ismertetett eszközök sorába.

Az **5. fejezet** összegzi a bemutatott informatikai lehetőségek kutatásban való alkalmazhatóságát, áttekinti, hogy az egyes eszközök a nyelvészeti kutatások mely fázisaiban használhatók hatékonyan, és milyen kutatási módszereknél vehetők igénybe.

A segédanyag kétféleképpen használható. Egyrészt tekinthetünk rá kézikönyvként, amely átfogó képet ad a funkcionális kognitív nyelvészeti kutatások során használható adatgyűjtési és adatfeldolgozási módszerekről, valamint olyan eszközökről, amelyek támogathatják ezen módszerek megvalósítását. Másrészt tekinthetjük segédanyagnak, melynek segítségével képet kaphatunk az egyes eszközök használatáról a technikai jellemzések alapján. Bármelyiket is válassza az olvasó, bízom benne, hogy hasznos információkkal fog gazdagodni.

2. A kognitív nyelvészeti kutatások

2.1. A funkcionális kognitív nyelvészet

A kognitív nyelvészet az 1970-es évek végétől kibontakozó nyelvészeti irányzat (TOLCSVAI NAGY 2013: 28). Több, azonos alapelveken nyugvó elméletet magában foglaló gyűjtőfogalom, elméleti irány, nyelvszeméleti mód, amely a nyelvet az emberi tudás részének, a nyelvhasználatot pedig mentális folyamatok eredményének tekinti (TOLCSVAI NAGY 2013: 42). A kognitív nyelvelméletek a 20. század metatudományos közegében születtek meg. Ebben a tudományfilozófiai közegben az 1950-es '60-as években több formalista irányzat is kialakult, amelyek a nyelvet logikai alapokon írták le, és leválasztották a beszédet és a jelentést a nyelvtanról (például CHOMSKY 1968).

A kognitív nyelvészet kezdetben ezzel a formális elméleti kerettel ellentétben funkcionális alapokon határozta meg elveit. Elutasította a logikai pozitivizmust és amellett érvelt, hogy a nyelvtudományokban nem lehet egyedüliként és kizárólagosan a természettudományos deduktív módszertant érvényesíteni (TOLCSVAI NAGY 2013: 34). A kognitív irányzat tagadja a Chomsky által bevezetett univerzális grammatika létét (CHOMSKY 1965: 6), ehelyett a nyelvi tudást megismerés és interakciók révén szerzett tudásnak tekinti (TOLCSVAI NAGY 2013: 50). A formális teóriákkal ellentétben a funkcionális elméletek nem választják el a beszélőt és a nyelvhasználatot a nyelvtantól és a nyelvleírástól (TOLCSVAI NAGY 2013: 42). Ez a keret úgy tartja, hogy a nyelv alakulásában közrejátszanak nem nyelvi tényezők is: a környezet, az emberi test biológiai, pszichológiai és fejlődési tényezői, valamint a szociokulturális háttér (LANGACKER 1999: 15–16).

A funkcionális kognitív nyelvészet a nyelv és az elme kapcsolatából indul ki, a nyelvet is az emberi társas megismerés részének tekinti (TOLCSVAI NAGY 2013: 50). Ez alapján az embert és az emberi elme világra irányulását helyezi a leírás középpontjába. Az embert kreatív, intencionális lénynek tekinti, aki környezetével és a többi emberrel interakcióra képes (TOLCSVAI NAGY 2013: 37). A nyelv az emberi interakció és kommunikáció során alakul, a nyelvi struktúrák a beszédpartnerek számára hozzáférhetővé tenni kívánt jelentéseket közvetítik, a beszéd során aktuálisan konstruálódnak és a nyelvközösségben konvencionálódnak (GEERAERTS-CUYCKENS 2007: 17).

A funkcionális nyelvelméletek a nyelvet nyelvhasználati adatok megfigyelése alapján írják le, és nem már meglévő elméleteket akarnak velük igazolni (TOLCSVAI NAGY 2013: 33). A nyelvleírás középpontjában a természetesen előforduló nyelvi szerkezetek állnak. Mivel a kognitív elmélet a nyelvet a nyelvi interakciókon keresztül elsajátított tudásnak tekinti, ezért a nyelvi kifejezéseket ezen interakció közegében vizsgálja (LANGACKER 2008: 25). A nyelvi kifejezéseket eszerint kontextusukban kell vizsgálnunk, az egyes kifejezések tipikus kontextusai pedig konvencionálódhatnak (KEMMER-BARLOW 2000: xxi). A nyelvi szerkezeteket sémák alapján konstruálják a beszélők, ezek a sémák pedig bizonyos valószínűség szerint realizálódnak a beszédben (TOLCSVAI NAGY 2013: 54). Mivel a beszélő intenciói, illetve az interakcióban megvalósult közös jelentésképzés kiemelkedő szerepet játszanak a nyelvi szerkezetek létrejöttében, a kognitív nyelvleírás a jelentésből indul ki, a formát és a jelentést együttesen kezeli, amely viszonyban a nyelvi kifejezések megvalósulását a jelentés strukturálja a társas jelentésképzés során (TOLCSVAI NAGY 2013: 55). Az elmélet a nyelvi kifejezéseket tehát nemcsak szerkezetükkel írja le, hanem a leírásba bevonja a konstruálás és a megértés műveletét is (TOLCSVAI NAGY 2013: 57).

Ugyan a kutatási módszerek alapvetően elméletfüggetlenek, az elmélet kérdésfeltevései mégis előnyben részesíthetnek bizonyos módszertani megfontolásokat, hiszen különböző elméleteknek különböző céljaik, fókuszaik lehetnek, ahogyan ezt a formális és a funkcionális megközelítés esetén láthattuk. Így ez a kézikönyv, bár számos más területen is alkalmazható módszert és eszközt sorakoztat fel, ezeket a kognitív nyelvészet szemszögéből mutatja be.

2.2. A kognitív nyelvészeti kutatások folyamata

Langacker a kognitív nyelvészeti kutatási folyamatokat a következők szerint írja le: a kutatás *a)* valamilyen jelenség megfigyeléséből indul ki, ezt követi *b)* a jelenséggel kapcsolatos természetesen előforduló nyelvi adatok gyűjtése, majd *c)* egy olyan elméleti modell kidolgozása, amely igazolja az adatokat és további predikciókat tesz lehetővé. Ezt követi *d)* a validálási szakasz, amely további adatokon ellenőrzi a modellt és az adatok alapján finomítja is azt. Ezután *e)* alternatív modelleket is összehasonlíthatunk, így kiválaszthatjuk, hogy melyik alkalmasabb leginkább a jelenség leírására, végül pedig *f)* a formalizálás következik (LANGACKER 1987: 31). Itt tehát nem egy előzetesen kidolgozott modellt igazol a kutató, hanem az adatok irányítják a modellkidolgozást. A kutatási szakaszok átfedhetik egymást, nem feltétlenül lineáris sorrendben következnek, és akár többszörösen is sorra veheti őket a kutató (LANGACKER 1987: 32). Emellett viszont Langacker számos különböző módszertant megenged a nyelv tanulmányozásához, mivel elveti azt a gondolatot, miszerint egyetlen módszertan mindenféle nyelvészeti kutatásra sikeresen alkalmazható (LANGACKER 1999: 23). Azt is kifejti, hogy bizonyítékainkat célszerű több forrásból gyűjteni (LANGACKER 1999: 26).

Mivel a kognitív nyelvészet magát használatalapúnak vallja, adatait elsősorban a nyelvhasználatból nyeri empirikus módon (GEERAERTS-CUYCKENS 2007: 17; GRONDELAERS et al. 2007: 149; TOLCSVAI NAGY 2013: 32). Az adatokat főként korpuszadatokról szerzi, azonban alkotott példák, kísérletek is lehetnek adatforrások (TOLCSVAI NAGY 2013: 62; GONZALEZ-MARQUEZ et al. 2007: xxii).

2.3. A kognitív nyelvészeti kutatások során alkalmazott adattípusok

A használatalapú kognitív nyelvészetben tehát fontosak a valós nyelvhasználat során létrejövő nyelvi adatok, emellett a kutatások kísérleti adatokra és introspekcióra is támaszkodnak. Ezek az adattípusok, adatforrások eltérő nézőpontokból mutatják be a nyelvet és a mögötte meghúzódó mentális műveleteket. Míg az általában korpuszokból nyert természetesen előforduló írott vagy beszélt

nyelvi adatok elsősorban a nyelvi szerkezetekről és azok kontextusáról adnak információkat, addig az introspekció és a kísérletek során a megismerő és konstruáló műveletekről is többet tudhatunk meg.

Nyelvi adatokat a nyelv bármely aspektusáról, bármilyen kontextusból gyűjthetünk. Az adatok tartalmazhatnak írott szövegeket, hanganyagokat, hanganyagok átíratát; származhatnak különböző térben és időben, különböző regiszterből gyűjtött nyelvi eseményekből. A kognitív nyelvészet vizsgálja az egyéni nyelvváltozatokat, nyelvi reprezentációkat is, nem gondolja, hogy csak egyféle szabályszerűség és megvalósulás léteznék egy-egy jelenség nyelvi kifejezésére vagy egy-egy nyelvi kifejezés adekvátságára vonatkozóan. Ezért szabályok helyett mintázatokot vizsgál, valamint azok előfordulásának körülményeit és gyakoriságát.

A nyelvi mintázatok leírása mellett a kognitív nyelvleírás figyelmet fordít azon kognitív folyamatokra, amelyek a nyelvi sémák és reprezentációik mögött meghúzódnak. A nyelvhasználati adatok mellett ezért kísérleti adatokat is felhasznál az elmeműködés vizsgálatára. Utóbbi adatokat általában pszicholingvisztikai kísérleteken, interjúkon keresztül nyeri. Az elmeműködés tanulmányozására a kísérleti adatok mellett introspekciót is alkalmaznak a kognitív nyelvészek. Ugyan az introspekció kizárólagos alkalmazását elveti a nyelvelmélet, nem zárja ki teljesen, sőt kifejezetten hasznosnak tartja olyan jelenségek vizsgálatánál, amelyek közvetlenül nem hozzáférhetőek a figyelem számára. Ilyen például a szavak jelentésének vizsgálata és a fogalmi metaforák kutatása (TALMY 2007a: xiii).

A nyelvészeti kutatások során használt adatok természetét meghatározzák azok az adatgyűjtési módszerek, eljárások, amelyekkel a kutatók a vizsgált adatokhoz jutnak. Minden eljárás számos tévedési lehetőséget rejt magában, illetve bizonyos módokon torzíthatja az adatokat, és ezáltal az eredményt is. Ahhoz, hogy megfelelő következtetéseket lehessen levonni bizonyos adathalmazokból, tisztában kell lennünk azzal, hogy az egyes eljárások és a kutató egyéni ítéletei miként torzíthatják az adatokat és milyen korlátai vannak az egyes adattípusok alkalmazásának. A következő fejezet arra világít rá, hogyan befolyásolják a kutató intuíciói a kutatás során nyert adatokat és eredményeket.

2.4. Introspekció és empirikus adatok viszonya

A kognitív nyelvészetet gyakran érik vádak azzal kapcsolatban, hogy túl nagy jelentőséget tulajdonít a kutató intuíciójának ahelyett, hogy kizárólag empirikusan szerzett adatokra (korpuszadatok, kísérletekből származó adatok) támaszkodna (GIBBS 2006: 136). A kritikák szerint az introspekciók nem alkalmasak arra, hogy általános érvényű következtetéseket vonhassunk le belőlük, hiszen minden kutató és minden nyelvhasználó introspekciója eltérő lehet (GIBBS 2006: 143). Ezért is alakulnak ki különböző elméletek az egyes nyelvi jelenségekkel kapcsolatban (pl. Lakoff és Johnson [1980] metaforaelmélete vagy Fauconnier és Turner [2002] fogalmi integrációs modellje a fogalmi metaforák magyarázatára). Ezek az elméletek azonban mind hozzátesznek a jelenség leírásához, és nem különülnek el teljesen egymástól. A kognitív nyelvészet azt vallja, hogy a nyelv nem lehet független a beszélők intuíciójától, minden egyes beszélő intuíciója – így a nyelvészé is – a nyelv része (WILLEMS 2012: 673). Az introspekció kiemelkedően fontos a jelentés és a kognitív műveletek vizsgálatánál, hiszen ezek a figyelem számára alig hozzáférhetők: nem tudunk automatikusan számot adni arról, mi játszódik le elménkben, amikor megalkotjuk a beszédet vagy megértjük azt. Ezeknek a folyamatoknak a megfigyelésére lehet hasznos az introspekció. A kognitív nyelvelmélet a kutató intuícióit eltérőnek tartja egy laikusétól, és aktívan fel is használja azokat a kutatásban. Hipotézisek felállításának, kísérletek megtervezésének forrásai lehetnek. Emellett azonban fontos tisztában lenni azzal, hogy az embereknek kevés rálátásuk van a mögöttes kognitív folyamatokra, így a különböző emberek (kísérleti személyek, nyelvhasználók, kutatók) intuíciói eltérhetnek egymástól.

Gibbs (2006) szerint Lewicki és munkatársai (1988) kutatásai kimutatták, hogy az emberek az esetek többségében nem tudnak pontosan beszámolni arról, mi áll döntéseik hátterében, és ez alól a gyakorlott kutatók sem kivételek. Lewicki és munkatársai egy tudatalatti tanulással kapcsolatos kísérletet végeztek el olyan kutatókon, akik PhD-fokozattal rendelkeztek pszichológiából, jártasak voltak a kognitív pszichológiai irányzatokban, és tudták, hogy a kísérletet végzők kutatási területe a tudatalatti tanulás (LEWICKI et al. 1988: 27). A kísérletben részt vevők egy négy részre osztott képernyőt láttak, amelynek mindig másik negyedében jelent meg egy X bizonyos

időközönként, egy bizonyos mintázatot követve. A résztvevők feladata az volt, hogy minél gyorsabban és minél pontosabban nyomják meg a megfelelő negyedhez kijelölt gombot a billentyűzeten (LEWICKI et al. 1988: 27). Ezután interjút készítettek velük tapasztalataikról, amelynek során kiderült, hogy a 9 kutatóból 4 vette csak észre, hogy valamilyen tudatalatti stimulusnak lett kitéve, azonban annak természetéről nem tudtak számot adni. Többen arról számoltak be, hogy egy idő után automatikusan nyomták meg a megfelelő gombokat, és az X megjelenésének mintázatáról is képesek voltak igaz állításokat megfogalmazni, de nem tudták teljesen leírni a mintát (LEWICKI et al. 1988: 33). Ebből arra következtettek, hogy az emberi információfeldolgozás automatikus és tudattalan folyamat, amiről az emberek nem képesek pontosan beszámolni (LEWICKI et al. 1988: 35). Gibbs (2006) szerint ez bizonyítja, hogy még egy bizonyos területen jártas kutatók sem képesek pontosan megragadni azon mentális folyamatokat, amelyek az információfeldolgozás, így a nyelvfeldolgozás közben is lezajlanak elméjükben (GIBBS 2006: 139).

A kognitív pszichológia úgy véli, hogy csak indirekt metódusokkal lehetséges hozzáférni a vizsgált személyek elmeműködéséhez, viselkedésének okaihoz, míg az egyéni beszámolók megtévesztőek lehetnek és torzíthatják a vizsgálat eredményét. Ilyen például a fogalmi metaforák kutatása, amely jórészt egyéni beszámolók és introspekció alapján történik. Ezzel szemben a kognitív pszichológusok kísérletekkel, indirekt módszerekkel vizsgálják ugyanazt a jelenséget. Például Gibbs és O'Brien (1990) úgy vizsgálták az idiómákban megjelenő metaforák mögötti motivációkat, hogy több kísérleti személyen megvizsgálták, milyen mentális képeket társítanak az egyes kifejezésekhez (GIBBS–O'BRIEN 1990: 39).

Megbízhatunk-e tehát az introspekcióban a tudományos kutatás során? Mint mindegyik adatforrást, az introspekciót és az intuíciót is óvatosan érdemes kezelni. Gibbs azt javasolja, hogy a nyelvészek ne hagyatkozzanak pusztán az introspekciókra, hanem próbálják azokat tesztelni (GIBBS 2006: 146). Az alábbi javaslatokat teszi az intuíciók jobb felhasználásának érdekében (GIBBS 2006: 146–147 alapján):

1. A nyelvészek ne csak saját intuícióikra támaszkodjanak, hanem másokat is vonjanak be a vizsgálatokba, és tegyenek fel kérdéseket nekik intuícióikkal kapcsolatban.
2. Ugyanazt a személyt egy későbbi időpontban is kérdezzék meg, mert az intuíciók idővel változhatnak.
3. Nagy mennyiségű nyelvi anyagot vizsgáljanak a jelenséggel kapcsolatban az introspekció alkalmazása mellett.
4. Ne csak alkotott példákat alkalmazzanak, mert ezek általában nélkülözik a kontextust. Valós szövegeket is tanulmányozzanak, a kontextusukkal együtt.
5. A lehető legpontosabban határozzák meg egy nyelvi elem leírásának kritériumait.
6. Az intuíciókat mindig nagy méretű korpuszon tesztelik.
7. Kutatásaikba a lehető legnagyobb mértékben vonják be a kontextust.

Gibbs nézőpontja itt tulajdonképpen az, hogy a nyelvészeti introspekcióknak akkor van létjogosultságuk, ha azok egy kísérlet megtervezését szolgálják, vagy ha autentikus nyelvi adatokat is bevonnak a kutatásba az intuíciók tesztelésére. Talmy is úgy fogalmazza meg álláspontját, hogy minden módszer másra alkalmas, az introspekciónak pedig az a különlegessége, hogy hozzáférést biztosít a jelentéshez és a jelentésképzés kognitív folyamataihoz (TALMY 2007b: 20). Azonban azt is állítja, hogy az introspekció mint módszer elutasítása sosem lehet az introspekciónak mint a vizsgálat tárgyának elutasítása. Az introspekció szerinte egy létező kognitív képesség, amelyet lehet és kell is vizsgálni (TALMY 2007b: 20). Ezen a ponton tulajdonképpen a két nézet megegyezik abban, hogy az introspekciók vizsgálatára további kutatásokat, kísérleteket tartanak célszerűnek végezni nyelvi adatok felhasználásával.

Az introspekció és a kutató egyéni intuíciói fontosak tehát a kutatás során, azokat azonban mindenképpen ellenőrizni kell, és kiegészíteni empirikus módszerekkel; illetve célszerű többféle adattípusra támaszkodnunk egy kutatási kérdés megválaszolásához. A következő fejezet eltérő típusú adatok gyűjtésére és feldolgozására vonatkozó, a kognitív nyelvészetben gyakran alkalmazott empirikus módszereket mutat be.

3. Empirikus kutatási módszerek a funkcionális kognitív nyelvészetben

A kutatás sikerességéhez pontosan meg kell terveznünk, hogy milyen adatokra lesz szükségünk a kutatási kérdésünk megválaszolásához, ezeket az adatokat milyen módszerekkel fogjuk összegyűjteni és milyen módszerekkel fogjuk elemezni. Ez a fejezet olyan adatgyűjtési és adatelemzési módszereket ismertet, amelyeket a funkcionális kognitív nyelvészeti kutatásokban gyakran alkalmaznak.

3.1. Gyakori adatgyűjtési módszerek

A gyűjteni kívánt adatok típusa meghatározza, hogy milyen adatgyűjtési eljárásokat alkalmazhatunk. Az adatgyűjtési eljárásoknak számos típusa létezik, azonban jelen segédanyag olyan módszereket tartalmaz, amelyeket egyrészt a funkcionális kognitív nyelvészeti kutatások során leggyakrabban alkalmaznak, másrészt a kivitelezésük nagymértékben megkönnyíthető bizonyos informatikai eszközök bevonásával. Ez a fejezet csupán néhány módszert ismertet, a bevonható eszközök alkalmazását a **4. fejezet** tárgyalja. Az itt ismertetett adatgyűjtési módszerek a korpuszhasználat, a kérdőíves felmérés és az interjúkészítés.

3.1.1. Korpuszhasználat

A kognitív nyelvészet vizsgálatának középpontjában a természetesen előforduló nyelvi megnyilatkozások állnak. Ezeket általában korpuszokból nyerik a kutatók. A korpusz eredeti használati közegükben, természetesen előforduló szövegek gyűjteménye, amelyeket egy-egy nyelv vagy nyelvváltozat leírására válogatnak össze (SINCLAIR 1991: 171). A mai korpuszok elektronikusan tároltak, adataik számítógépes programok segítségével feldolgozhatók. A modern korpuszok minél nagyobb mértékben próbálják visszaadni a nyelv változatosságát, akár több millió vagy milliárd szóból is állhatnak (SINCLAIR 1991: 171). Céljuk szerint többféle korpuszt különböztethetünk meg. A leggyakoribb az általános korpusz, amely egy bizonyos nyelvet hivatott reprezentálni, nagy mennyiségű szöveget tartalmaz és sokféle kutatási célra alkalmazható (BAKER 2010: 99). Ilyen általános korpusz például a **Magyar Nemzeti Szövegtár** (a továbbiakban: MNSz.) és újabb változata, az **MNSz2**. A kisebb, valamilyen műfajra, időtartamra, helyre, nyelvváltozatra korlátozott korpuszok a specializált korpuszok, amelyek specifikus kutatásokhoz kínálnak segítséget. Gyakran egy általános korpuszsal együtt használják őket, hogy összehasonlíthatóvá váljon a specifikus jelenség a standardhoz képest (BAKER 2010: 99). A korpuszokat megkülönböztethetjük annak alapján is, hogy milyen adatokat tartalmaznak, például beszéltet vagy írottat. Különbséget tehetünk aszerint is, hogy hány nyelvből tartalmaznak adatokat. A csupán egy nyelvből származó adatokat tartalmazó korpuszok egynyelvűek, míg léteznek két- vagy többnyelvű, úgynevezett párhuzamos korpuszok is, mint például a Hunglish Corpus (HALÁCSY et al. 2005).

Azonban a digitalizált szöveghalmazok önmagukban még nem nevezhetők nyelvészeti korpusznak. Azokat alkalmassá kell tenni nyelvészeti vizsgálatokra, vagyis a szövegeket fel kell dolgozni (FÓRIS 2008: 137). Ennek lehetséges módja a szövegek annotálása. Ez azt jelenti, hogy az egyes szövegekhez, azok bekezdéseihez, mondataihoz és szavaihoz nyelvészeti információkat rendelnek, amiket úgynevezett *tag*-ek formájában tárolnak a szöveghez kapcsolva (MITTELBERG et al. 2007: 38). Ezt általában valamilyen jelölőnyelv segítségével teszik, amely alkalmas a korpusz strukturálására és arra, hogy tárolható formátumúvá tegye a korpuszt. A *tag*-ek

jelölhetnek nyelvtani kategóriákat vagy akár stilisztikai tulajdonságokat, metaforákat, prozódiai tulajdonságokat. Ez a fajta annotálás automatikusan és manuálisan is végezhető. Az automatikus gépi annotálás nem hiba nélküli, szükséges annak kézi ellenőrzése (például Szeged Korpusz: CSENDES et al. 2003). Bizonyos tulajdonságokat, például a prozódiát, nehéz géppel annotálni. Az automatikus annotációt általában külön szoftverekkel végzik a korpusz készítői (magyar nyelvre ilyen például az E-magyar, VÁRADI et al. 2018).¹ Fontos, hogy mindig olyan korpuszt válasszunk, amely a kutatási célnak megfelelően van annotálva, mert ez megkönnyítheti a munkánkat.

A korpuszvizsgálatok célja, hogy a nagy mennyiségű elektronikusan tárolt és annotált szöveg mintáit elemezve szabályszerűségeket, trendeket ismerjen fel a kutató a nyelvi produkcióval kapcsolatban (BAKER 2010: 94). A korpuszvizsgálatok tehát lehetővé teszik a kutatók számára, hogy a vizsgált nyelv egy reprezentatív mintáján ellenőrizni tudják feltevéseiket nyelvi jelenségekkel kapcsolatban vagy bizonyos adatminták alapján új elméleteket állítsanak fel, új kutatási kérdéseket nyerjenek (BAKER 2010: 94). A korpuszvizsgálatoknak kétféle megközelítése létezik: a korpuszvezérelt és a korpuszalapú. A korpuszvezérelt kutatások induktívak, céljuk a hipotézisformálás, elméletalkotás (BAKER 2010: 95). Ez azt jelenti, hogy az adatokból levezetett általános szabályszerűségek, következtetések alapján állít fel a kutató egy hipotézist vagy elméleti modellt, és nem feltételez valamely már előzetesen megalkotott modellt vagy magyarázatot. Tehát minden elméletalkotást a korpuszban előforduló adatok elemzésére alapoz, és akár hipotéziseket is ennek alapján állít fel. Az így létrejövő modell később újabb korpuszvizsgálattal vagy más adatgyűjtési módszerrel tesztelhető, finomítható. A korpuszalapú vizsgálatok ezzel szemben egy már adott elmélet adatalapú tesztelésére szolgálnak. A korpuszvezérelt és a korpuszalapú vizsgálatok tehát összefügghetnek, azonban a korpuszalapú vizsgálatok nemcsak korpuszvezérelt módszerrel felállított hipotézisek tesztelésére szolgálhatnak. Az utóbbi, deduktív kutatások hátránya általában, hogy mivel elméletvezéreltek, sokszor figyelmen kívül hagynak az elméletbe nem illeszkedő adatokat (BAKER 2010: 95).

¹ A modern annotáló szoftverek már többnyire mesterséges intelligencián alapulnak, amelyek kézzel annotált tanulókorpusz alapján fejlesztik saját algoritmusukat a fejlesztő által meghatározott kezdeti modellekből kiindulva. Képesek viszonylag nagy pontossággal akár jelentéseket is meghatározni (ilyenek például a szentimentelemző szoftverek).

A korpuszok adatait meghatározott módon válogatják össze a korpusz készítői. Fontos, hogy *a)* a korpusz a vizsgált sokaságra és jelenségre nézve maximális reprezentativitásra törekedjen; *b)* gépek számára olvasható formában legyen tárolva a későbbi feldolgozás érdekében; valamint *c)* lehetőleg kellően nagy méretű legyen ahhoz, hogy lehetővé tegye mind a gyakori, mind a ritka jelenségek vizsgálatát (McENERY–WILSON 2001: 32). A korpusz mérete függ a vizsgálni kívánt jelenségektől is. Míg a jobban reprezentált nyelvi jelenségek vizsgálatához elegendő egy kisebb korpusz is, addig a kevésbé reprezentáltak számára nagyobb korpuszra van szükség (BIBER 1993: 249). Például a grammatikai vizsgálatokhoz többnyire elegendőek a kisebb korpuszok (BIBER 1993: 249), azonban a lexikográfia számára nagyobb méretű korpuszokra van szükség, mert míg a grammatikai jelenségek valószínűbb, hogy előfordulnak kisebb mintában is, addig a szavak közül sok csak egyszer fordulna elő egy ugyanakkora korpuszban, így ezekről nem lehetne általános következtetéseket levonni (KENNEDY 1998: 68). Természetesen a korpuszok mérete nemcsak a nyelvészeti vizsgálatok tárgyától függ, hanem a fellelhető forrásoktól és a kapacitástól is. Sok esetben az egyébként hasznos szövegek nem elérhetőek szerzői jogok miatt (ORAVECZ et al. 2014: 1720), illetve a korpuszépítésre szánt idő véges, így ezek a külső tényezők is korlátozzák, hogy milyen megállapításra juthatunk a korpuszadatok alapján.

A korpuszadatok szűrését segítik a korpuszlekérdező vagy korpuszelemző szoftverek. Fontos leszögezni, hogy a korpusz és a korpuszt lekérdező szoftver nem azonos. Ennek a különbségnek a tudatosítása azért lényeges, mert ez azt jelenti, hogy ugyanazon szöveghalmazt többféle eszköz segítségével tudjuk vizsgálni. Mindegyik más karakterisztikák szerint szűri az adatokat, ami ahhoz is vezethet, hogy a különböző szoftverek máshogyan torzíthatják az adatmintánkat. Bizonyos korpuszok saját lekérdezőfelülettel rendelkeznek, ilyen például az MNSz. Az MNSz.-hez két lekérdezőfelület is kapcsolódik, az egyik egy általános korpuszlekérdező,² míg a másik az igék bővítményszerkezetének vizsgálatára specializálódott (Mazsola, vö. SASS 2009).³ Utóbbi nem használja a korpusz teljes anyagát. Léteznek független elemzőszoftverek is, amelyek bármilyen, akár saját kezűleg összeállított korpusz szűrésére és

2 A <http://clara.nytud.hu/mnsz2-dev/> oldalon található az MNSz2, amely az MNSz megújult változata.

3 A Mazsola lekérdezőfelülete a <http://corpus.nytud.hu/mazsola/> oldalon érhető el.

elemzésére alkalmasak. A legtöbb ilyen lekérdezőfelületen kereshetünk kulcsszavak alapján, illetve alkalmazhatunk reguláris kifejezéseket (lásd **4.1.2. fejezet**), valamint kereshetünk egy speciális korpuszlekérdező nyelv, az úgynevezett *Corpus Query Language* (CQL) segítségével.⁴

Az egyes szoftverek között nagy eltérések lehetnek, nemcsak céljuk, hanem a fejlesztők döntései alapján is. Például a különböző programok eltérő módon értelmezhetik azt, hogy mit tartanak egy szó részének. Ennek szemléltetésére hoz példát Laurence Anthony, aki a 2009-es amerikai elnöki nyitóbeszéd szógyakoriságán keresztül mutatja be, hogy három eltérő korpuszlekérdező program kétféle eredményt ad a *we* szó gyakoriságára (ANTHONY 2013: 150). Ez azért lehetséges, mert míg két programban a szó részének tekintik az aposztrófot, addig a harmadikban nem. Így az első kettő nem számolta bele a *we'll* szóelőfordulásokat a keresésbe, míg a harmadik igen, mert a *we* utáni aposztrófot már külön szóként azonosította. Tehát a lekérdezőszoftverek specifikációi közti különbségek eltérésekhez vezethetnek a végső találati listákban. Emellett a korpuszok építése során sem egyértelmű minden, nagyon sok múlik a fejlesztők egyéni döntésein és interpretációin. A kutatás megtervezésekor célszerű tehát tájékozódni arról, hogy milyen elvek mentén dolgozták fel a korpuszban foglalt szövegeket, és milyen szempontok szerint kereshetők ezek a szövegek.

A legtöbb szoftver lehetőséget biztosít arra, hogy ne csak egyes elszigetelt nyelvi kifejezéseket vagy jelenségeket tudjunk vizsgálni, hanem azokat kontextusukban tudjuk megragadni. Gyakori eszköz az erre a célra szánt konkordancia-készítés és a kollokációk kimutatása. A *konkordancia* a keresett kifejezésekkel együtt kilistázza azok előfordulásainak bizonyos terjedelmű környezetét is, azaz a kifejezés jobb és bal oldalán álló szavakat (SINCLAIR 1991: 32). Ez az úgynevezett 'Key Word in Context' (röviden: KWIC) formátum, amely a mai konkordanciák leggyakoribb formátuma. Ennek segítségével a keresett kifejezés közvetlen környezetéről kaphatunk információt. A *kollokációk* a szavaknak olyan együttállásai, amelyek statisztikailag szignifikánsnak mondhatók (BAKER 2010: 107). Megadják, hogy

4 A CQL használatát jelen segédanyag nem tárgyalja. A CQL reguláris kifejezésekkel való kapcsolatáról és a CQL alapjairól lásd: SASS 2017: 270.

milyen valószínűséggel fordulnak elő bizonyos szavak egymás közelében, általában bizonyos szónyi távolságban. Ezt a távolságot a kutató beállíthatja. A kollokációk tehát megmutatják, hogy melyek azok a leggyakoribb mintázatok, amelyekben az adott szó vagy kifejezés előfordul.

A számítógépes vizsgálatok nagy előnye, hogy gyorsan és pontosan képesek adatok keresésére, szűrésére, kollokációk meghatározására, így olyan adatokat is vizsgálhatóvá tesznek, amelyek az emberi figyelem előtt rejtve maradtak volna. Természetesen ezek az eszközök sem tévedhetetlenek, viszont a segítségükkel egyszerűbben juthatunk olyan információhoz, amelyek bizonyos nyelvi tendenciákat mutatnak meg. A korpuszlekérdezők torzítását csökkentendő érdemes a vizsgálatokat több mintán elvégezni és finomítani a lekérdező kifejezést, hogy minél jobban szűrt találati listát kapjunk.

A korpuszadatok elektronikus volta és egyszerű számszerűsíthetősége lehetőséget ad kvantitatív statisztikai alapú számításokra. Adott számításokat (például gyakorisági számításokat, kollokációk kimutatását) már a korpuszlekérdezők is elvégzik számunkra, így gyorsan juthatunk számszerű információkhoz. Hátrány lehet, hogy komplex jelenségekre nehéz keresni (pl. metafora), azonban a gyakorisági számítások mellett kvalitatív elemzések is végezhetők korpuszokon. Ennek legegyszerűbb módja a kontextusok vizsgálata. Például Partington (1998: 107–119) A FEL JOBB, A LE ROSSZABB metaforákat vizsgálta olyan módon, hogy a *fel* és a *le* kifejezésekre és azok szinonimáira keresve a kapott konkordanciát, a kifejezések kontextusát vizsgálni tudta.

Sokan az internetet is egy hatalmas korpusznak tekintik, és a webes keresést a korpuszkereséssel egyenrangúnak vélik. Az internetes keresés azonban nem feltétlenül reprezentatív adatokból válogat, az eredményt bizonyos mértékig befolyásolhatja a felhasználó kereséstörténete, földrajzi helyzete. Az internet nem strukturált és nincsen annotálva, a webes keresők kimenete pedig nem nyelvi adat, hanem weblapok sokasága. Bizonyos esetekben mégis hasznos lehet az internetes keresés, például a feltételezéseink megerősítésére egy-egy komolyabb kutatás, korpuszvizsgálat előtt. Ritka vagy újabb jelenségeknél sok esetben jobb eredményeket

adhat, mint a korpuszok, azonban ezeket az eredményeket óvatosan kell kezelni (lásd bővebben: MITTELBERG et al. 2007: 43). A hatékony webes keresésekkel a **4.1.1. fejezet** foglalkozik.

A korpuszok jó alapot biztosítanak a használatalapú kognitív nyelvészeti kutatások számára, azonban nem feltétlenül szükséges minden kognitív kutatáshoz korpuszt alkalmazni; illetve az adatok keresése, kiválasztása, elemzése mind interpretatív folyamat, így bizonyos mértékben a kutató intuícióján alapszik (GRONDELAERS et al. 2007: 150). A korpuszvezérelt szemléletmód ezt igyekszik kiküszöbölni.

3.1.2. Kérdőíves felmérések és interjúk

A funkcionális kognitív nyelvészet a nyelvi jelenségek mögött meghúzódó kognitív folyamatokat is vizsgálja, ezért a nyelvi adatok mellett olyan adatokat is felhasznál, amelyekből következtetni lehet ezen folyamatokra. Ezt a célt szolgálhatják a kérdőíves felmérések és az interjúk.

Kérdőívekkel többféle adatot gyűjthetünk. Rákérdezhetünk a válaszadók véleményére bizonyos kifejezésekkel kapcsolatban, kiegészíttethetünk velük mondatokat, nyelvi ítéleteikről kérdezhetjük őket. A kérdőívek kitöltése történhet személyesen papíron vagy digitálisan, valamint online módon digitálisan. Fontos, hogy a kérdéseket és a kérdések sorrendjét, a kitöltés körülményeit pontosan megtervezzük, mert a kutatás során azokon utólag már nem változtathatunk. A kérdések alapvetően zártak vagy nyitottak lehetnek. *Zárt* kérdések esetén a kutató által összeállított válaszlehetőségek közül választhat az adatközlő. Ezeket a válaszokat általában könnyű kódolni és statisztikai számításokkal feldolgozni, azonban a válaszadók ebben az esetben nem tudnak a kész panelektől eltérő választ adni, így nem biztos, hogy az eredmény pontosan tükrözi majd a résztvevők véleményét. Ezzel szemben a *nyitott* kérdésekre az adatközlők önálló választ fogalmazhatnak meg. Ezek a válaszok nehezebben kvantifikálhatók, így a feldolgozásuk több időt vesz

igénybe, azonban olyan információkhoz is hozzájuthatunk, amelyre nem számítottunk korábban (FÓRIS 2008: 126). A kérdőíveink tartalmazhatnak zárt és nyitott kérdéseket vegyesen, vagy csak egyik vagy csak másik típusút. Fontos, hogy a kérdések típusa megfeleljen a kutatás céljainak, és a válaszok feldolgozásának módja is ehhez a célhoz igazodjon. Ha például arra vagyunk kíváncsiak, hogy bizonyos nyelvi kifejezések közül melyeket preferálják az adatközlők, akkor használhatunk zárt kérdéseket, azonban ha az érdekel minket, hogy miért az adott kifejezéseket részesítik előnyben, akkor célszerű nyitott kérdést alkalmazni.

A két típus vegyítésére jó példa Tátrai és Ballagó (2020) kutatása, amelyben kérdőíves vizsgálatokkal igazolják, hogy a beszélőnek a figyelemirányítás kontextusában érvényre jutó szociokulturális pozíciója fontos tényező a stílustulajdonítás folyamatában (TÁTRAI–BALLAGÓ 2020). A kutatás során két egymásra épülő kérdőíves vizsgálatot alkalmaztak. Az első nyitott kérdőív volt, amelyben az adatközlők egyetemi szemináriumok diskurzusrészleteinek stílusát minősítették saját szavaikkal, majd leírták, hogy mely kifejezések alapján jutottak megállapításukhoz (TÁTRAI–BALLAGÓ 2020: 13). A második kérdőíves vizsgálat az elsőre épülő zárt kérdőív volt (TÁTRAI–BALLAGÓ 2020: 2). A második kérdőívben ugyanúgy a szemináriumok diskurzusrészleteit kapták meg az adatközlők, azonban bizonyos részletekben kiemelték azokat a kifejezéseket, amelyeket az első kérdőív adatközlői a stílus szempontjából leginkább feltűnőnek tartottak (TÁTRAI–BALLAGÓ 2020: 15). Az adatközlőknek ezúttal zárt kérdésekre kellett válaszolniuk. Az első kérdőívekre válaszként adott, a szövegek stílusára vonatkozó kifejezéseket használták fel a második kérdőív lehetséges válaszaihoz, így a második kérdőív adatközlői az e kifejezésekből összeállított skálán jelölhették, melyiket tartják a legmegfelelőbbnek a szövegrészlet stílusának jellemzésére (TÁTRAI–BALLAGÓ 2020: 16). Láthatjuk tehát, hogy a nyitott kérdőívek megalapozhatják zárt kérdőívek elkészítését is. Ebben az esetben a zárt kérdőív a nyitott kérdőív válaszainak tesztelésére szolgált.

Az interjúk a kérdőívekhez hasonlóak, amennyiben az adatközlők kérdésekre válaszolnak, azonban itt a kérdezés és válaszadás élőszóban történik. Az interjú tulajdonképpen olyan beszélgetés, amelynek előre meghatározott célja és struktúrája van (KVALE–BRINKMANN 2009: 3). A kutató a beszélgetés felvétele előtt meghatározza, hogy milyen szempontok mentén kívánja irányítani

a beszélgetést annak érdekében, hogy a kutatás céljának megfelelő adatokhoz jusson (FÓRIS 2008: 124). Az interjúk során fényt deríthetünk az adatközlők intuícióira, értékítéleteire, benyomásaira. Emellett nyelvi adatokat is gyűjthetünk interjúk segítségével, ha szituációkkal előidézünk bizonyos nyelvhasználati jelenségeket. Az interjú kvalitatív adatokat szolgáltat, amelyeket nehéz számszerűsíteni, feldolgozásuk a nyitott kérdőívek feldolgozásához hasonló. Az interjúkról általában felvételek készülnek, amelyeket lejegyeznek a kutatók valamilyen átírási rendszerrel. Ezeknek a transzkripcióknak a célja, hogy feldolgozhatóvá tegyék a szöveget, de megőrizték a beszéd szupraszegmentális jegyeit (BORONKAI 2009: 52). Azonban nem létezik egyetlen általános érvényű, objektív átírási rendszer, ezért mindig a kutatás céljának megfelelően és átláthatóan kell átírni a szövegeket (KVALE–BRINKMANN 2009: 186). Így az egyes kategóriák kialakítása és a kategóriába való besorolás is a kutató intuíciójára van bízva. Ez azt jelenti, hogy a rendelkezésre álló adatokból különböző kutatók különböző következtetéseket vonhatnak le. Az interjúknál jellemzően befolyásoló tényező még a kutató jelenléte, mert az adatközlők hajlamosak olyan válaszokat adni, amelyekről azt feltételezik, hogy jobban megfelel a kutató esetleges elvárásainak.

A kérdőívek előnye, tehát, hogy nagy adatmennyiség gyűjthető velük egyszerűen. A válaszok jellemzően számszerűsíthetők és statisztikai számításokkal feldolgozhatók. A számszerűsíthetőség hátránya, hogy a válaszadók sokszor csak korlátozott számú válasz közül választhatnak, és adott esetben a tényleges véleményük, feltételezéseik nem derülnek ki, így bizonyos problémák rejtve maradnak. Az interjúk időigényesebbek, azonban ezek során részletesebb válaszokat kaphatunk a résztvevőktől, így olyan problémákra és jelenségekre is fényt deríthetünk, amelyekre korábban nem gondoltunk, azonban ez az eljárás is torzíthatja az adatokat.

3.2. Egy gyakori kvantitatív adatelemzési módszer: a statisztikai elemzés

Az adatelemzési módszerek közül a statisztikai elemzést emeljük ki, amely kvantitatív adatok elemzésére szolgál. Ezt az elemzési módszert tudjuk leginkább megkönnyíteni informatikai eszközök alkalmazásával. A **4. fejezet**ben felsorolt eszközök többsége jól alkalmazható statisztikai elemzések során.

A statisztikai módszerek célja, hogy segítsék az adatok elrendezését, értelmezését, hipotézisek tesztelését (EDDINGTON 2015: 7). Ezt nagy mennyiségű számszerűsíthető adat feldolgozásával érik el. Az alapvető statisztikai számítások (pl. átlag, medián, szórás, gyakorisági értékek) segítségével könnyen és gyorsan kaphatunk információt kvantitatív adatokról. Olyan mintázatokra bukkanhatunk általuk, amelyek a kézi elemzés előtt rejtve maradtak volna, vagy kézi feldolgozásuk nagyon időigényes lett volna.

A statisztikai módszereknek alapvetően két iránya van, a leíró és az induktív statisztika. A *leíró statisztika* az adatok általános minőségéről ad információt (LEVON 2010: 70), általában az alapvető statisztikai számítások elvégzését és az adatok vizualizálását értjük alatta. Ilyen alapvető számításnak minősül a gyakoriságok, a középértékek és a szóródások kiszámítása. A középérték-mutatók közül a leggyakoribb az átlag, a medián, és a módusz. Az *átlag* az adatok átlagos értékét mutatja meg, a *medián* az adatok középső értékét, a *módusz* pedig a leggyakoribb értéket. A szórás az adatok átlagtól való eltérését mutatja. Ezeket az értékeket szemléletesen ábrázolhatjuk különböző típusú diagramokon. A leíró statisztikai számításokhoz a **4.3.4. fejezet** nyújt segítséget, az Excel táblázatkezelő program beépített statisztikai függvényeinek bemutatásával. Az adatvizualizációban a **4.4. fejezet**ben összegyűjtött alkalmazások segíthetnek.

Míg a leíró statisztika segítségével jellemezni, szemléltetni tudjuk adatainkat, addig az *induktív statisztika* azt a célt szolgálja, hogy meg tudjuk határozni, hogy a populáció egy kisebb mintáján végzett kutatás eredményei milyen mértékben igazak az egész populációra

(EDDINGTON 2015: 18). Ezáltal meghatározhatjuk, hogy azok a mintázatok, amelyeket a vizsgált adathalmazon látunk, valódi mintázatok-e, statisztikailag szignifikánsak-e a teljes populációban (LEVON 2010: 70). A kvantitatív kutatásokban leíró statisztikával kimutatott mintázatok alapján felállíthatunk egy kísérleti hipotézist, valamint ennek a cáfolatát, az ún. *nullhipotézist*, amely azt fogalmazza meg, hogy a kísérleti hipotézis nem igaz (LEVON 2010: 72). Az induktív statisztika azt méri, hogy mekkora a valószínűsége annak, hogy a nullhipotézis igaz. Ezt a valószínűséget százalékban az úgynevezett *p*-értékkel adhatjuk meg (LEVON 2010: 71). Amennyiben ez az érték egy bizonyos küszöb alatt van, elutasíthatjuk a nullhipotézist. Például, ha a küszöbértékünk 5%, és ez alatti az érték, a kísérleti hipotézisünk 95% eséllyel igaznak bizonyul. Ez a fajta kutatási séma a dedukcióból származik. A leggyakoribb induktív statisztikai számítások a *khi*-négyzet teszt (χ^2 próba) és a *t*-teszt. A *khi*-négyzet próba meghatározza, hogy az eloszlás, amit a mintánkon megfigyelünk vajon szignifikánsan különbözik-e attól az eloszlástól, amelyet bármilyen hasonló mértékű populáció esetében tapasztalnánk (LEVON 2010: 74). A *t*-teszt megvizsgálja két mintapopuláció közepét és standard eloszlását, majd megállapítja, hogy ezek szignifikánsan eltérnek-e egymástól (LEVON 2010: 74). Amennyiben a mintapopulációnk adatai nem bizonyultak szignifikánsnak, nem vonhatunk le általános érvényű következtetéseket.

A statisztikai módszerek tehát kvantifikálható adatok elemzésére szolgálnak. Amennyiben leíró jellegű, az adatokat általánosan jellemző eredményeket szeretnénk, akkor leíró statisztikai számításokat kell végeznünk, illetve ábrázolhatjuk adatainkat diagramokon. Ha viszont szeretnénk összehasonlítani az adatainkat más adathalmazokkal, akkor induktív statisztikát kell alkalmaznunk.

4. Informatikai eszközök a nyelvészeti kutatások szolgálatában

4.1. A keresés és lekérdezés módszertana

Keresést nagyon sokféle helyzetben végzünk a kutatások során, ezért kiemelkedően fontos, hogy hatékonyan tudjunk alkalmazni különböző típusú kereséseket. A kutatás megkezdésekor olvasmányokat, szakirodalmat, információkat keresünk. Ezután általában adatokat, majd az adatokban mintázatokat, összefüggéseket próbálunk fellelni. Ez a fejezet bemutatja, hogy hogyan tudjuk egyszerűen és hatékonyan kivitelezni ezeket a kereséseket informatikai eszközök segítségével. A fejezet először a weben való információ- és adatgyűjtési lehetőségeket taglalja, majd a gyűjtött adatokon végezhető hatékony mintázatkeresés egy speciális eszközét, a reguláris kifejezéseket ismerteti.

4.1.1. Keresés a weben

A mindennapi életben és kutatásaink során is gyakran keresünk a weben. Többnyire egyszerű kulcsszavas keresést alkalmazunk, azonban előfordul, hogy a keresett információ egyszerű kereséssel nem áll rendelkezésünkre, vagy túl sok találatot kapunk, amelyek között nehéz eligazodni és releváns információhoz jutni. A legtöbb keresőmotor, online adatbázis lehetőséget biztosít összetett keresésre is. Jelen segédanyag a Google keresőmotor speciális keresési lehetőségeit mutatja be példákkal.

4.1.1.1. Speciális Google-keresés

A *google.com* oldalon alkalmazhatunk egyszerű kulcsszavas keresést. Azonban, ha csak egyszerűen begépeljük a keresett kifejezést, rengeteg olyan találatot is kapunk, amelyek esetleg számunkra nem relevánsak. Például csak egyik szó szerepel a keresett szavak közül a találatban, vagy egyik sem, de rokonértelmű szavak igen. Hogy hatékonyabban juthassunk megfelelő eredményekhez, célszerű ezeket kiszűrni. Ezt megtehetjük a Google speciális keresés funkciójával, amelyet a https://www.google.com/advanced_search⁵ oldalon érünk el, vagy a *google.com* oldalról **Beállítások > Speciális keresés** menüpontokkal navigálva.

A speciális keresési felületen számos lehetőség áll rendelkezésünkre, hogy leszűkítsük találati listánkat. Megadhatjuk, mely szavak szerepeljenek pontosan a keresésben, mely szavakat ne tartalmazzák a találatok, milyen nyelvű oldalak között keressen a keresőmotor, mely országban publikált oldalakat listázzon. Ezen kívül megadhatjuk az utolsó frissítés dátumát és kereshetünk fájltypus szerint is.

A felület egyes mezői közül bármennyit kitölthetünk egyszerre egy keresés alkalmával. Ha például a birtokos szerkezetről keresünk cikkeket, akkor ezt különböző módokon tehetjük meg. Ha az **ezen szavak mindegyikét** címkéjű mezőbe írjuk a **birtokos szerkezet** kifejezést, akkor olyan találatokat fogunk kapni, amelyek tartalmazzák mind a **birtokos**, mind a **szerkezet** szót, azonban nem feltétlenül egyben tartalmazzák a teljes kifejezést. Így nagyjából 68 300 találatot kapunk.⁶ Ha ugyanezt a **pontosan ezt a szót vagy kifejezést** mezőbe írjuk, akkor csak azokat a találatokat kapjuk meg, ahol pontosan egymás mellett szerepel a két szó. Így már csak körülbelül 7 420 találat lesz. Ha emellett abban is biztosak szeretnénk lenni, hogy csak az elmúlt egy évből kapunk találatokat, akkor kitölthetjük az **utolsó frissítés**-t is. Itt az **elmúlt év** lehetőséget kell választani. Ebben az esetben a találatok száma 53. Láthatjuk, hogy ezekből számos találat általános iskolai nyelvtani összefoglaló.

5 Utolsó elérés dátuma: 2022. 05. 29.

6 A keresés 2020. augusztus 17-én történt.

Speciális keresés

Ilyen oldalak keresése:

ezen szavak mindegyikét:

pontosan ezt a szót vagy kifejezést:

ezen szavak bármelyikét:

ne tartalmazza ezen szavak egyikét sem:

számok ettől:

eddig:

Ennek elvégzéséhez a keresőmezőben.

Írja be a fontos szavakat: emeletes esküvői torta

A pontos kifejezéseket tegye idézőjelbe: "esküvői torta"

Az összes kívánt szó közé írja be a(z) OR szót: miniatűr OR szabványos

Az elkerülendő szavak elé tegyen mínuszjelet: -rágcsáló, -"Jack Russell"

Tegyen két pontot a számok közé és tüntessen fel egy mértékegységet: 10..35 kg, \$300..\$500, 2010..2011

Ezután szűkítse a találatokat eszerint:

nyelv:

A kiválasztott nyelven írott oldalak keresése.

régió:

Egy adott régióban közzétett oldalak keresése.

utolsó frissítés:

A megadott időn belül frissített oldalak keresése.

webhely vagy domain:

Keresés egy oldalon (pl. wikipedia.org), illetve a találatok korlátozása egy domainre, például: .edu, .org vagy .gov

megjelenő kifejezések:

Kifejezések keresése az egész oldalon, az oldal címében vagy webcímében, illetve a keresett oldalra mutató linkekben.

[Biztonságos Keresés:](#)

Adja meg a [Biztonságos Keresés](#) korhatáros találatok szűrésére vonatkozó beállításait.

fájltípus:

A kívánt formátumú oldalak keresése.

[felhasználási jogok:](#)

Szabaddon felhasználható oldalak keresése.

Speciális keresés

1. ábra.
A Google speciális
keresési felülete

Amennyiben ezeket ki szeretnénk szűrni, a **ne tartalmazza ezen szavak egyikét sem** mezőbe beírhatjuk az **általános iskola, iskola** kifejezéseket (lásd a 2. ábrát). Így 28 találatunk lesz. Ez a mennyiség már sokkal könnyebben áttekinthető, azonban még tovább szűkíthető, ha például csak a pdf-formátumú találatokra keresünk, ugyanis nagyobb az esélye annak, hogy ezek tanulmányok lesznek és nem blogbejegyzések. Ekkor már csak 3 találatot kapunk.

2. ábra. Speciális keresés több feltétel megadásával

Speciális keresés

Ilyen oldalak keresése:

ezen szavak mindegyikét:

pontosan ezt a szót vagy kifejezést:

ezen szavak bármelyikét:

ne tartalmazza ezen szavak egyikét sem:

számok ettől:

eddig:

Ezután szűkítse a találatokat eszerint:

nyelv:

régió:

utolsó frissítés:

webhely vagy domain:

megjelenő kifejezések:

Biztonságos Keresés:

fájltípus:

felhasználási jogok:

bármely nyelv

bármelyik régió

elmúlt év

bárhol az oldalon

A leginkább releváns találatok megjelenítése

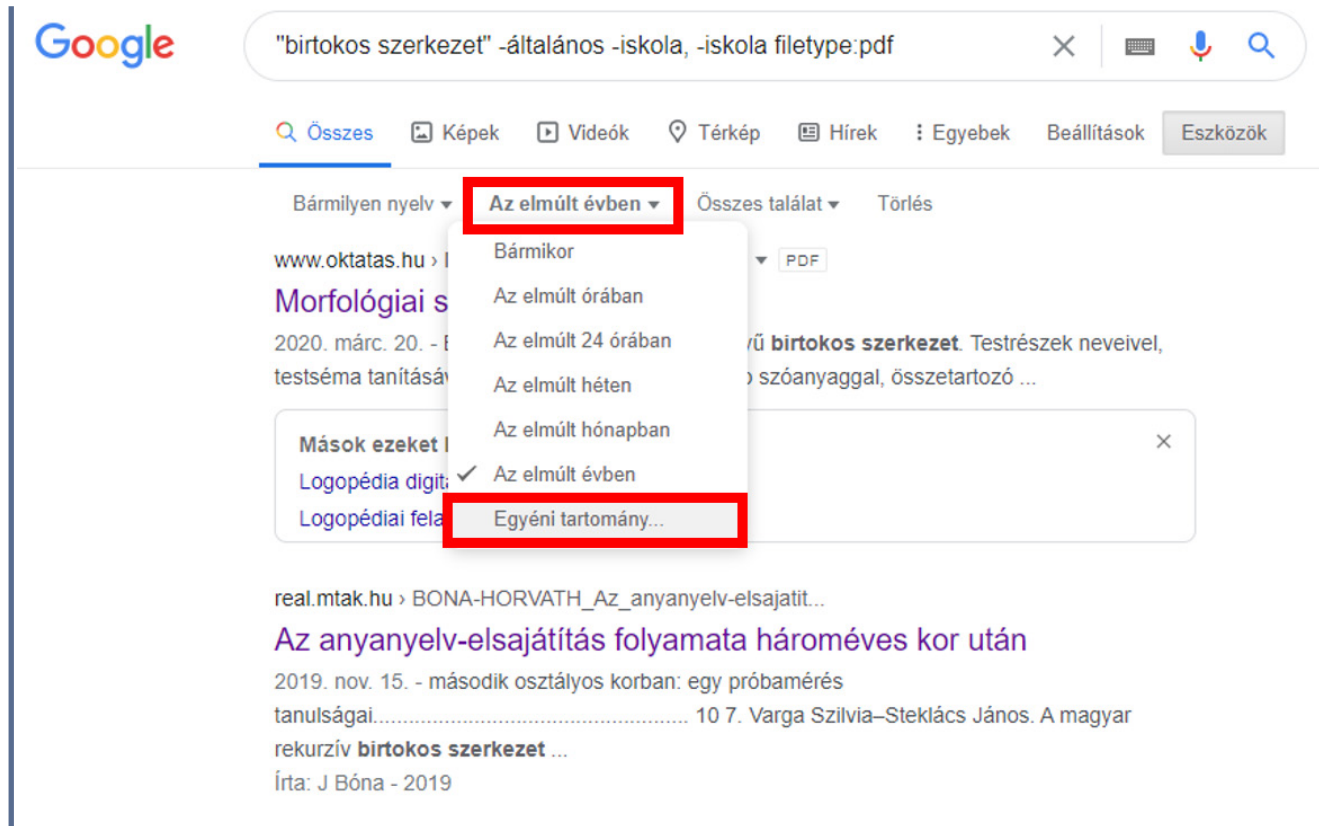
Adobe Acrobat PDF (.pdf)

nincs licenc szerinti szűrés

Speciális keresés

Természetesen ez csak egy példa a több-tényezős keresésre, előfordulhat, hogy a három találat közül egyik sem tartalmaz olyan információt, amire szükségünk van. Ekkor finomíthatjuk a keresést újabb kulcsszavak megadásával vagy az időintervallum kiszűrésével. Utóbbit megtehetjük úgy, hogy a keresés elküldése után a találatoknál az **elmúlt egy évben** lehetőséget **egyéni tartomány**-ra módosítjuk, majd a felugró ablakban megadjuk a pontos dátumokat (lásd a 3. ábrát).

3. ábra. Egyéni dátumtartomány beállítása



A mezők jobb oldalán lévő szöveg (vö. 1. és 2. ábra) segíti a pontos kitöltést, azonban ez sokszor megtévesztő. Több esetben nem azt írja le, hogyan kell a mezőt kitölteni, hanem hogy hogyan alkalmazhatjuk ugyanezt a keresést az egyszerű keresőfelületen operátorok segítségével. Például a **számok ettől: 10 eddig: 15** mező kitöltése helyett az egyszerű keresésénél a *google.com* oldalon azt is írhatjuk, hogy **10..15**. Ekkor ugyanazt kapjuk, mintha a speciális keresőmezőt töltöttük volna ki. A továbbiakban ezeket az operátorokat és rövidítéseket részletezem.

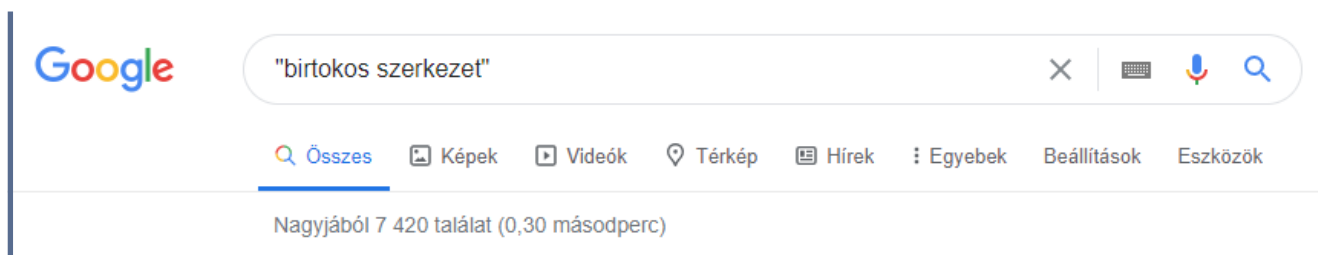
4.1.1.2. Google operátorok

A 3. ábrán láthatjuk, hogy a keresőmezőbe bekerültek a **"Speciális keresés"** felületén kitöltött keresési feltételeink, azonban egyéb jelek is társultak hozzájuk, mint az idézőjel, kötőjel és a **filetype:** parancs. Ezeket együttesen operátoroknak hívjuk, és azt a célt szolgálják, hogy minél pontosabban tudjunk feltételeket megadni a kereséseinkhez. Az operátorok a speciális keresőfelülethez képest több lehetőséget biztosítanak, illetve jobban lehet kombinálni őket, így a találatok listáját is hatékonyabban tudjuk szűkíteni velük, ha megfelelően használjuk őket. Az alábbiakban bemutatok néhány hasznos operátort. Mivel előfordul, hogy a Google változtat az operátorain, így itt csak a viszonylag állandóak szerepelnek. Az operátorokhoz a <https://support.google.com/websearch/answer/2466433?hl=hu> oldalon találunk segítséget. Fontos még megemlíteni, hogy bár a keresések során alapvetően nem számítanak a kis- és nagybetűk, azonban az operátoroknál igen.

"" keresés összetartozó kifejezésekre

A 3. ábrán láthattuk, hogy a **pontosan ezt a szót vagy kifejezést** mezőbe írt keresőszavakat a Google felülete idézőjelek közé helyezte. Ez azt jelenti, hogy az egyszerű *google.com* keresőbe írt **„birtokos szerkezet”** kifejezés ugyanazt az eredményt adja vissza, mintha a speciális felületen a **pontosan ezt a szót vagy kifejezést** mezőbe írtuk volna.

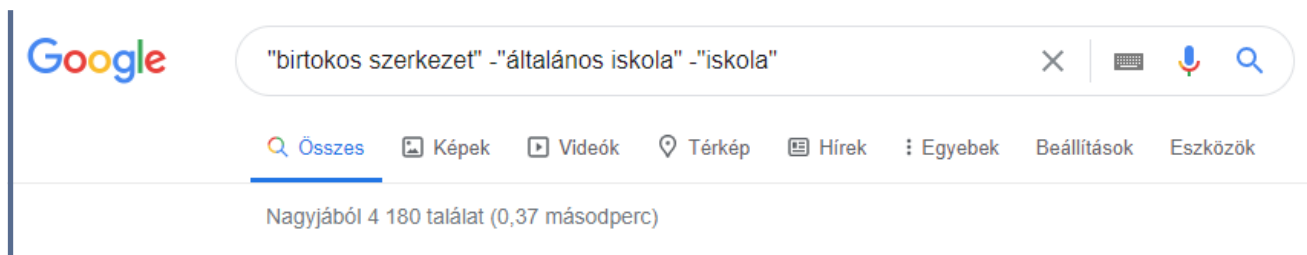
4. ábra. Összetartozó
kifejezések keresése



– kifejezés kizárása a keresésből

A 3. ábrán azt is láthatjuk, hogy a **ne tartalmazza ezen szavak egyikét sem** mezőbe írt kulcsszavakat mínuszjellel (vagy kötőjellel) jelöli az egyszerű keresőmező. A mínuszjel tehát azt a célt szolgálja, hogy kizárjunk a keresésből bizonyos kifejezéseket. Ha például nem szeretnénk, hogy generatív elmélettel kapcsolatos találatokat is tartalmazzon a keresésünk, akkor beírhatjuk a „- **generatív**” kifejezést. Jelen esetben az *iskola* és *általános iskola* kifejezéseket szeretnénk volna kiszűrni, azonban láthatjuk, hogy az *általános* és az *iskola* szavakat külön kifejezésként kezeli a felület. Ha több szóból álló kifejezéseket szeretnénk kiszűrni, akkor ismét alkalmazhatjuk az idézőjeleket: **-“általános iskola”**.

5. ábra. Kifejezések
kizárása a keresésből

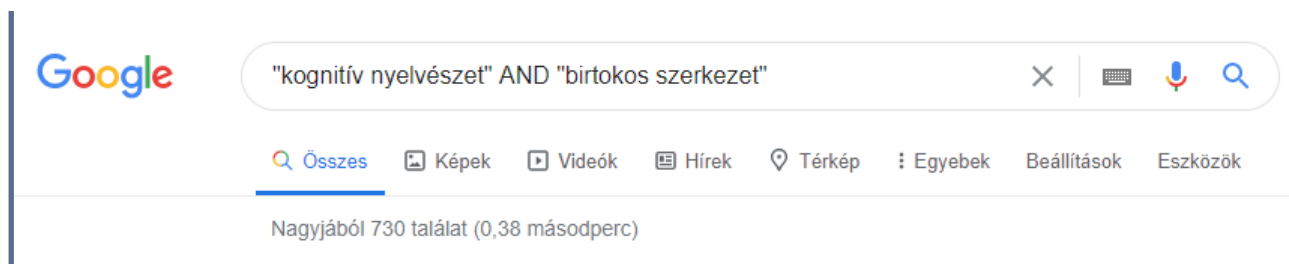


AND mindkét szó szerepeljen a találatban

Az 'AND' operátor lehetővé teszi, hogy a keresőmotor olyan találatokat adjon, amelyekben az 'AND' elé és mögé írt kifejezés együttesen szerepel. Például, ha csak kognitív nyelvészeti munkákra szeretnénk keresni a birtokos szerkezettel kapcsolatban, akkor azt többféle módon tehetjük meg. Egyrészt beírhatjuk a keresőbe a **kognitív nyelvészet birtokos szerkezet** kifejezéseket. Ekkor azonban olyan találatok is listázódnak, amelyekben vagy egyik vagy másik szó szerepel csak. Beírhatjuk idézőjelek közé a kifejezéseket: **"kognitív nyelvészet birtokos szerkezet"**, azonban így csak olyan találatokat kapunk, ahol pontosan egymás mellett szerepelnek ezek a szavak.

Ha azt szeretnénk, hogy csak olyan találatokat kapjunk, amelyekben mindkét kifejezés egyszerre szerepel, de nem egymás mellett, akkor alkalmazhatjuk az *AND* operátort: „*kognitív nyelvészet*” *AND* „*birtokos szerkezet*”.

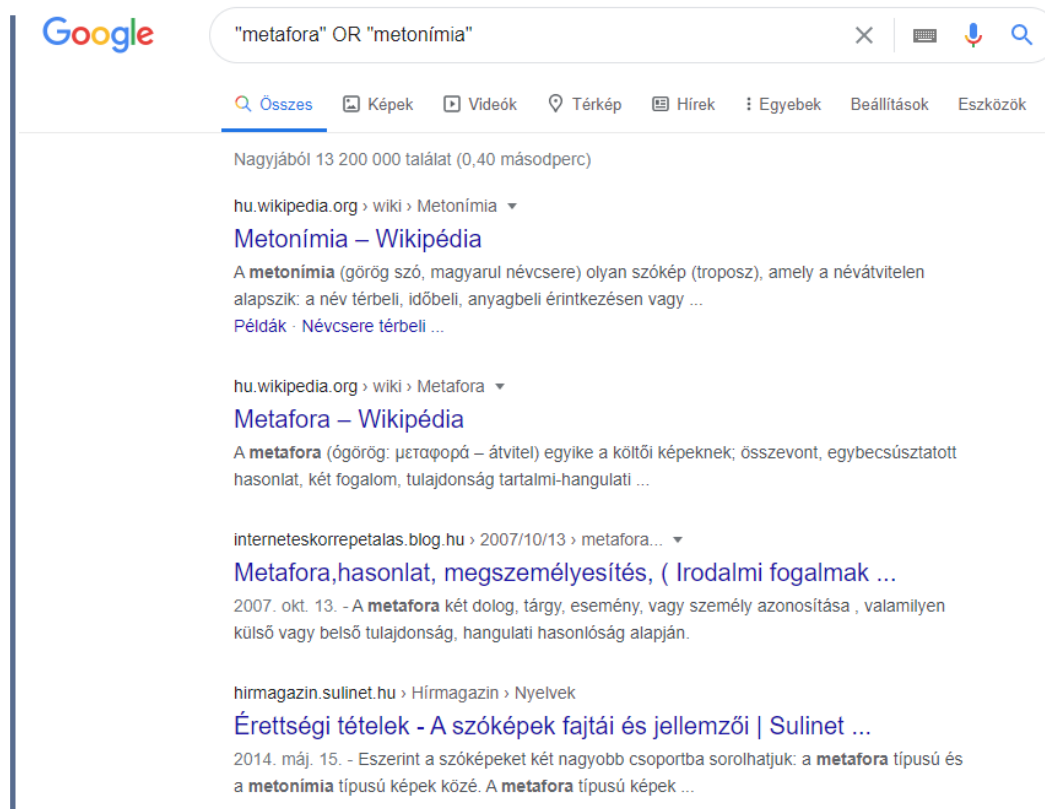
6. ábra. Keresés
'AND' operátorral



OR egyik vagy másik szó szerepeljen a találatban

Az 'OR' operátor lehetővé teszi, hogy olyan találatok is szerepeljenek a találati listában, amelyekben csak az OR elé vagy mögé írt kifejezés fordul elő. Ez nem zárja ki azokat a találatokat sem, amelyekben mindkét kifejezés előfordul. Ha például metaforával és metonímiával kapcsolatos találatok is érdekelnek minket, akkor használhatjuk az alábbi kifejezést: „*metafora*” OR „*metonímia*”.

7. ábra. Az 'OR'
operátor használata



* hiányzó karakterek, szavak

A * operátorral helyettesíthetünk karaktereket, szavakat a keresőkifejezésünkben. Ha például azt írjuk be, hogy **cognitive grammar *** **langacker**, akkor találatul kapjuk Langacker több könyvcímét is, amelyben a kognitív grammatikáról ír. Ez azért jobb, mint a * nélküli kifejezés, mert ez kizárja azokat a találatokat, amelyekben mindkét kifejezés megtalálható egymástól függetlenül. A 8. ábra a két keresés találatai közti különbséget mutatja be. A különbség nem olyan számottevő, de a *-os kifejezéssel ki tudtuk szűrni például a Wikipédia-találatot. Ennek az operátornak nagy hasznát vehetjük még azokban az esetekben, amikor nem jut eszünkbe egy szó például egy könyv címében, és ezt helyettesítjük vele.

8. ábra. Keresési eredmények
‘*’ operátorral és anélkül

The image shows two side-by-side Google search results. The left search is for "cognitive grammar langacker" and the right search is for "cognitive grammar * langacker". The right search results are more specific, showing books by Langacker.

Left Search: cognitive grammar langacker

Nagyjából 238 000 találat (0,43 másodperc)

Tudományos cikkek - "cognitive grammar langacker"

Cognitive grammar: A basic introduction - Langacker - Idézetek száma: 4799

Essentials of cognitive grammar - Langacker - Idézetek száma: 222

Foundations of cognitive grammar: Theoretical ... - Langacker - Idézetek száma: 18107

en.wikipedia.org > wiki > Cognitive_gr... > Oldal lefordítása

Cognitive grammar - Wikipedia

Cognitive grammar is a cognitive approach to language developed by Ronald Langacker, which hypothesizes that grammar, semantics, and lexicon exist on a ...

oxford.universitypressscholarship.com > ... > Oldal lefordítása

Cognitive Grammar: A Basic Introduction - Oxford Scholarship

Ronald Langacker. Abstract. Cognitive Grammar is a radical alternative to the formalist theories that have dominated linguistic theory during the last half century.

www.amazon.com > Cognitive-Gramm... > Oldal lefordítása

Cognitive Grammar: A Basic Introduction ... - Amazon.com

Amazon.com: Cognitive Grammar: A Basic Introduction (9780195331967): Langacker, Ronald W.: Books.

edisciplinas.usp.br > pluginfile.php > mod_folder > content | PDF

Cognitive Grammar: An Introduction - e-Disciplinas

Langacker, Ronald W. Cognitive grammar : a basic introduction / Ronald Langacker. p. cm. Includes bibliographical references. ISBN 978-0-19-533195-0; ...

Right Search: cognitive grammar * langacker

Nagyjából 304 000 találat (0,51 másodperc)

Tudományos cikkek - "cognitive grammar * langacker"

Essentials of cognitive grammar - Langacker - Idézetek száma: 222

Foundations of cognitive grammar: Theoretical ... - Langacker - Idézetek száma: 18107

... of Cognitive Grammar: descriptive application. Volume ... - Langacker - Idézetek száma: 5006

www.amazon.com > Cognitive-Gramm... > Oldal lefordítása

Cognitive Grammar: A Basic Introduction ... - Amazon.com

Amazon.com: Cognitive Grammar: A Basic Introduction (9780195331967): Langacker, Ronald W.: Books.

www.amazon.co.uk > Essentials-Cognit... > Oldal lefordítása

Essentials of Cognitive Grammar: Amazon.co.uk: Langacker ...

Buy Essentials of Cognitive Grammar by Langacker, Ronald W. (ISBN: 9780199937356) from Amazon's Book Store. Everyday low prices and free delivery on ...

www.amazon.de > Essentials-Cognitive... > Oldal lefordítása

Essentials of Cognitive Grammar: Amazon.de: Langacker ...

Essentials of Cognitive Grammar | Langacker, Ronald W. | ISBN: 9780199937356 | Kostenloser Versand für alle Bücher mit Versand und Verkauf durch Amazon.

www.amazon.fr > Cognitive-Grammar-... > Oldal lefordítása

Cognitive Grammar: A Basic Introduction - Langacker, Ronald ...

Noté /5. Retrouvez Cognitive Grammar: A Basic Introduction et des millions de livres en stock sur Amazon.fr. Achetez neuf ou d'occasion.

intitle: és allintitle: keresés az oldal címében

Ha olyan találatokat szeretnénk csak eredményül kapni, amelyeknél a *birtokos szerkezet* kifejezés a cikkek címében szerepel, akkor az **'allintitle: birtokos szerkezet'** keresést használhatjuk (lásd 9. ábra). Az **'intitle:'** és az **'allintitle:'** között csupán annyi a különbség, hogy az **'intitle:'** csak a mögötte álló egyetlen szóra keres, míg az **'allintitle:'** több szóra is képes keresni.

9. ábra. Az 'allintitle:' operátor használata

intext: és allintext: keresés az oldal szövegében

Ez a két operátor annyiban különbözik az előző kettőtől, hogy nem a cikkek címében, hanem azok szövegében keres a megadott kifejezésekre. Ez azt jelenti, hogy a címekben lévő találatokat nem fogja kiadni.



allintitle: birtokos szerkezet



Összes

Képek

Videók

Hírek

Térkép

Egyebek

Eszközök

Nagyjából 178 találat (0,44 másodperc)

<https://www.magyarora.com> > bariton_birtokos_uj PDF

A birtokos szerkezet - Magyaróra

FIGYELEM, FIGYELEM! A számnevek után egyes számot használunk: Van öt macskám / három lányom / néhány jó barátom / (és egy-két dal gitárom.
5 oldal

<https://angol.hu> > klm

13 birtokos szerkezet különleges esetei - angol.hu

13 birtokos szerkezet különleges esetei. Ld még 11, 12, 14. (1) "Elváló" birtokos. Az 's-t nem maga a birtokost kifejező főnév kapja, hanem a teljes főnévi ...

A(n), one, every, each, (n)either: friend Some/any, no, two, three, many, few: friends
Határozatlan névelő(féle): főnév

<https://angol.hu> > klm

12 birtokos szerkezet - angol.hu

12 birtokos szerkezet ... (1) Az angolban az egyetlen eset, amely felismerhető végződéssel rendelkezik, a birtokos eset, pl.: Rosemary's dog. A birtokviszony ...

<https://nevmasblog.wordpress.com> > nevmasok > birtok...

Birtokos szerkezet | Névmásblog

2020. jan. 21. — A számbeli egyeztetés hiánya a magyarban. A magyarban az alany és az ige egyezni szoktak számban és személyben, ahogy azt az alábbi példasor ...

<https://osvenyek.elte.hu> > media > Birtokos szerke... PDF

Birtokos szerkezet - Ösvények

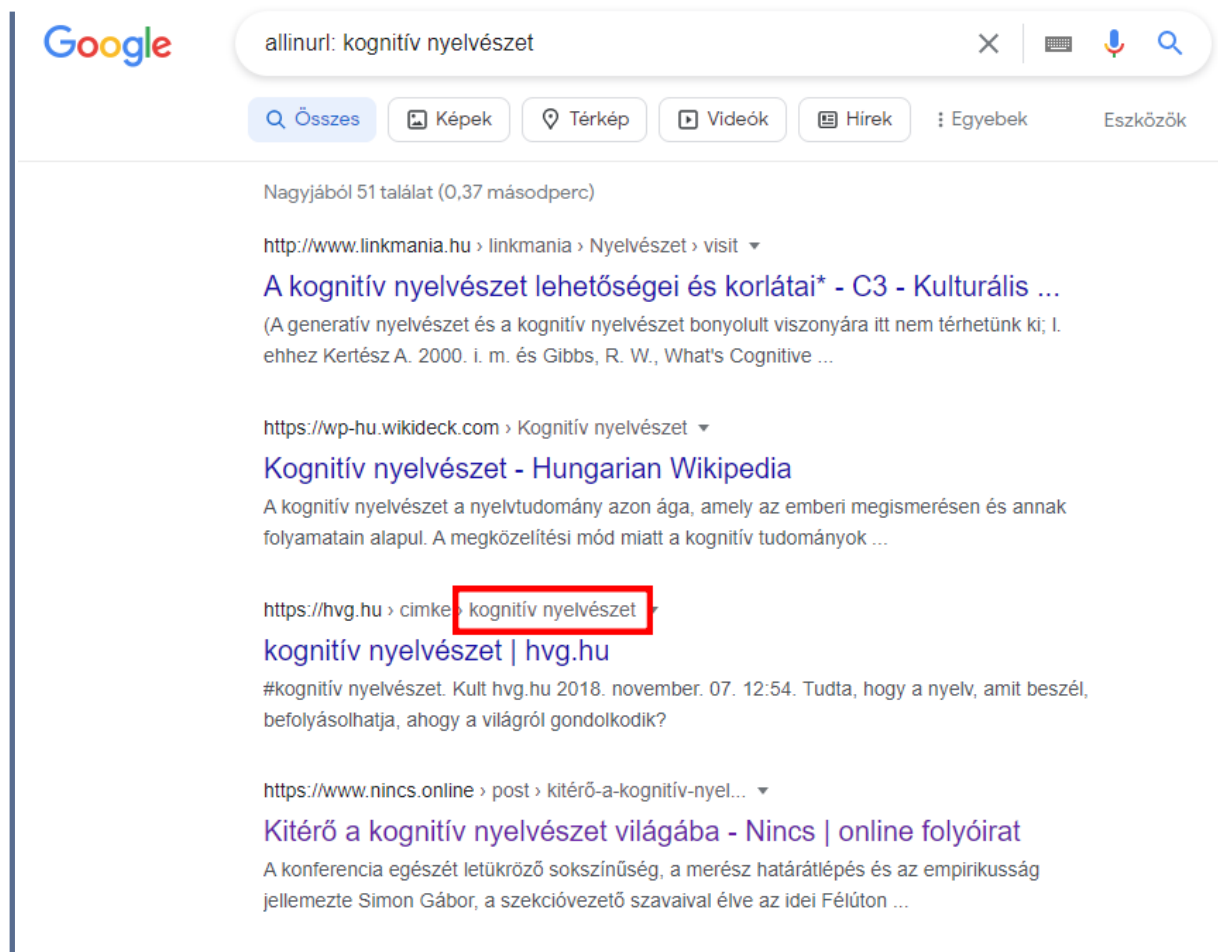
A tanulmány fő témája a magyar birtokos személyjeles konstrukciók vizsgálata. Az elemzés kiindulópontjában a Langacker-féle referenciapont-modell áll, ...
9 oldal

inurl: és allinurl: keresés az oldal URL-jében

Szintén hasonló az **'intitle:'** és az **'intext:'** operátorokhoz, azonban ezek a keresőkifejezések az oldalak URL-címében⁷ keresnek megadott szavakra. A 10. ábrán lévő példa azt mutatja, hogyan kereshetünk olyan oldalakra, amelyeknek az URL-jében szerepel a **kognitív nyelvészet** kifejezés. A pirossal jelzett területen láthatjuk az URL azon részét, amely tartalmazza a kifejezést. Ennek segítségével olyan oldalakat is elérünk, amelyek például előadásokat, ismertetőket, blogbejegyzéseket tartalmaznak az adott témával kapcsolatban.

10. ábra. Keresés az URL-ben

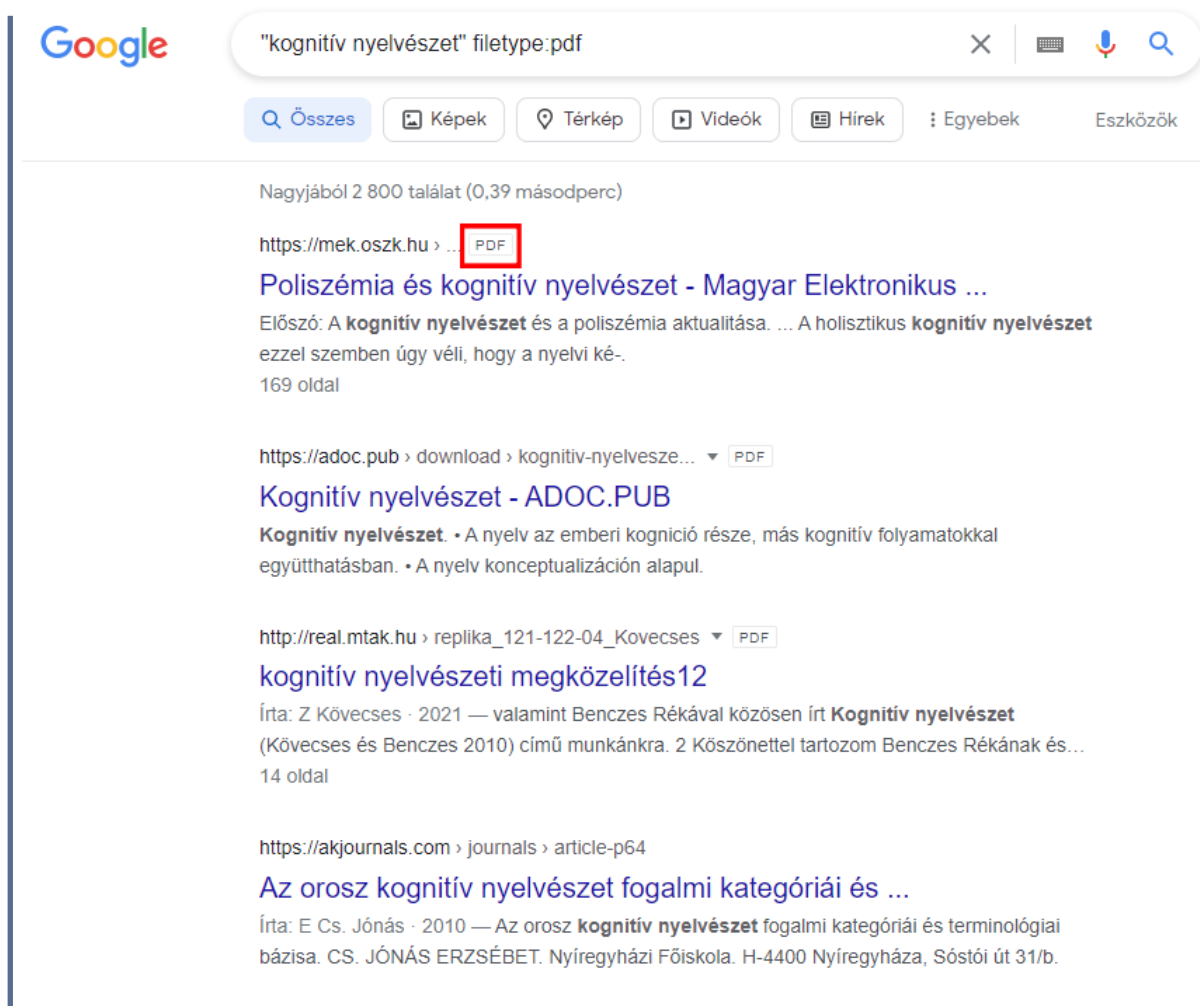
⁷ URL: A Uniform Resource Locator rövidítése, tulajdonképpen a weboldal címe, amelyet a címsorba írunk.



filetype: keresés fájltypus szerint

A speciális kereséses példánkban szűkítettük a keresésünket pdf-fájlokra. Ezt az egyszerű keresésénél a **'filetype:'** paranccsal tehetjük meg. A 11. ábrán pirossal jelölt terület mutatja a talált dokumentum formátumát. Természetesen nemcsak pdf-re, hanem egyéb szöveges fájltypusra is kereshetünk, pl. docx-re, odf-re.

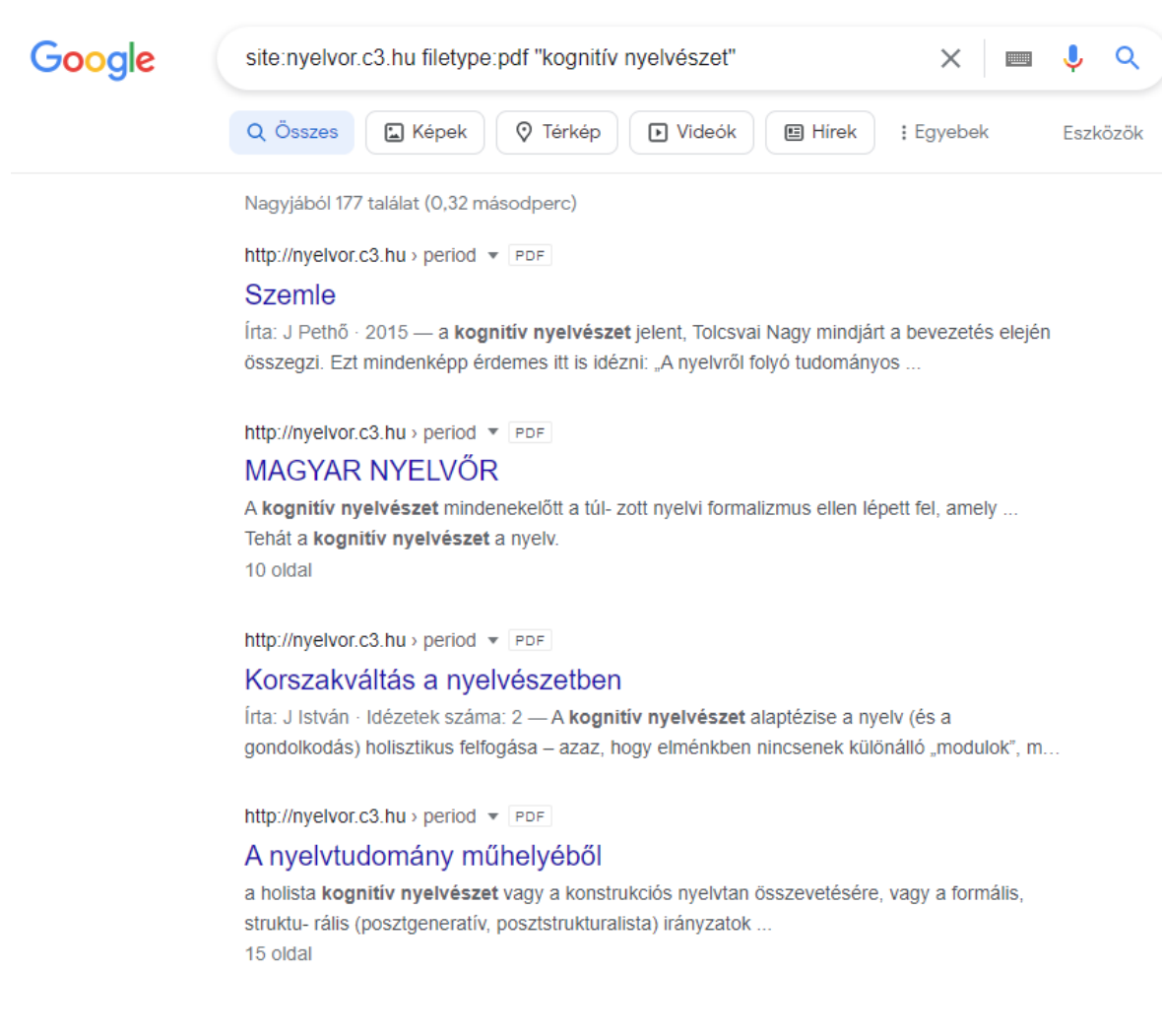
11. ábra. Keresés fájltypusra



site: keresés egy oldal teljes tartalmában

Az egyik leghasznosabb keresőoperátor a 'site:', amely lehetővé teszi, hogy egy adott weboldal tartalmában keressünk. Használata főleg akkor jelenthet segítséget, amikor az adott oldal nem tartalmaz keresőfunkciót. Vigyázni kell azonban vele, mert előfordulhatnak úgynevezett indexelési hibák, amikor a Google nem jól azonosítja az oldal lapjait. Ekkor általában túl sok vagy túl kevés találatot kapunk. Természetesen nem minden oldal egyformán kereshető így, hiszen bizonyos oldalakat nem ér el a Google sem. A **'filetype:'** operátorral és kulcsszavakkal kombinálva fájlokat is kereshetünk az adott oldalon. A 12. ábrán látható keresés a Magyar Nyelvőr weboldalán található „**kognitív nyelvészet**” kulcsszavakat tartalmazó pdf-dokumentumokra ad találatot.

12. ábra. Keresés konkrét weboldalak tartalmában



related: hasonló tartalomra keresés

A 'related': operátor segítségével egy bizonyos weboldalhoz hasonló tartalommal rendelkező oldalakat kereshetünk. Például a *Magyar Nyelvőr*-höz hasonló oldalakat kereshetünk a 13. ábrán látható módon.

13. ábra. Hasonló tartalmú weboldalak összegyűjtése

Google search results for the query `related:http://nyelvor.c3.hu/`.

6 találat (0,20 másodperc)

- [hu.wikipedia.org > wiki > Magyar_Nyelvőr](http://hu.wikipedia.org/wiki/Magyar_Nyelv%C3%B6r)
Magyar Nyelvőr – Wikipédia
Története[szerkesztés]. A folyóiratot Szarvas Gábor Hunfalvy Pálnak, Budenz Józsefnek és Gyulai Pálnak a kezdeményezésére, az MTA támogatásával indította ...
- [adtplus.arcanum.hu > collection > MagyarNyelvőr](http://adtplus.arcanum.hu/collection/MagyarNyelvor)
Magyar Nyelvőr 1872-2018 | Arcanum Digitális Tudománytár
Hunfalvy Pál, Budenz József, Gyulai Pál kezdeményezésére, a MTA támogatásával indította útjára a folyóiratot Szarvas Gábor. A Magyar Tudományos ...
- www.nytud.hu
MTA Nyelvtudományi Intézet
MTA Nyelvtudományi Intézet : nyelvtudomány, nyelvészet.
- www.nyest.hu
Nyelv és Tudomány- Főoldal
Árpád-házi kéktúra Afganisztánból Baskortosztánba. III. Béla királyunk ősei 4500 éve Baktriából indultak Baskíriába. Onnan pedig 2000 éve továbbálltak.
- [adtplus.arcanum.hu > collection > MagyarNyelv](http://adtplus.arcanum.hu/collection/MagyarNyelv)
Magyar Nyelv 1905-2017 | Arcanum Digitális Tudománytár
Magyarország legnagyobb és folyamatosan bővülő digitális periodika adatbázisa, amely a teljesség igényével teszi hozzáférhetővé több száz hazai ...
- [www.kislexikon.hu > magyar_nyelvor](http://www.kislexikon.hu/magyar_nyelvor) [Oldal lefordítása](#)
Magyar Nyelvőr - Lexikon :: - Kislexikon
magyar nyelvészeti folyóirat, melyet a magyar tudományos akadémia nyelvtudományi bizottságának megbízásából és az akadémia anyagi támogatásával ...

define: keresés definíciókra

Ez a funkció csak angolul működik igazán jól. A 'define:' operátor segítségével egy fogalom meghatározására kereshetünk rá. A Google első találatnak egy szótári szócikket ad, amely definiálja az adott kifejezést, majd a további találatok is kapcsolódnak a keresett fogalom definíciójához (14. ábra). Segítségével több tudományos oldalra is eljuthatunk a megfelelő szűrésekkel.

14. ábra. A 'define:' operátor használata

The screenshot shows a Google search for 'define:cognition'. The search bar at the top contains 'define:cognition'. Below the search bar, there are tabs for 'Összes', 'Képek', 'Hírek', 'Videók', 'Térkép', 'Egyebek', 'Beállítások', and 'Eszközök'. The search results show 'Nagyjából 57 600 000 találat (0,38 másodperc)'. The first result is from Oxford Languages, titled 'cognition' with the phonetic transcription /kɒɡˈnɪʃ(ə)n/. The definition is: 'the mental action or process of acquiring knowledge and understanding through thought, experience, and the senses.' Below the definition are buttons for 'Rokonértelmű:' (perception, discernment, awareness, apprehension, learning) and 'Rokonértelmű:' (perception, discernment, awareness, apprehension). There is a dropdown menu for 'A(z) cognition kifejezés fordítása a következőre:' with 'Magyar' selected. Below this, it says 'noun' and '1. megismerés'. There is a link for 'További meghatározások és etimológia'. At the bottom, there is a section 'Mások ezeket a kérdéseket is felteszik' with four questions: 'What is cognition in simple terms?', 'What is an example of cognitive?', 'What does cognitive status mean?', and 'What are the types of cognition?'. Below this, there are links to 'en.wikipedia.org' and 'www.verywellmind.com' for 'Cognition - Wikipedia' and 'What Is Cognition? - Verywell Mind' respectively.

szám..szám keresés intervallumon belül

Már említettük, hogy számtartományra úgy is kereshetünk, hogy a kezdő- és végpontot jelölő számok közé két pontot teszünk. A módszer segítségével évszámokra is kereshetünk a weboldalak szövegében. Például: *"cognitive linguistics" 2000..2020*.

15. ábra. Keresés számsorozat elemeire

Google "cognitive linguistics" 2000..2020

Nagyjából 376 000 találat (0,59 másodperc)

en.wikipedia.org › wiki › Cognitive_li... ▼ Oldal lefordítása
Cognitive linguistics - Wikipedia
Cognitive linguistics is an interdisciplinary branch of linguistics, combining knowledge and ...
34 (6): 656–664. doi:10.1016/j.langsci.2012.04.007. ^ Greenwood ...
Background · Approaches · Cognitive Linguistics ... · Integrative cognitive ...

www.degruyter.com › cogl - Oldal lefordítása
Cognitive Linguistics | De Gruyter
Please note that our journal has adopted a new Open Data Policy effective as of 1 July 2020.
Your benefits: State-of-the-art linguistic research on the interaction ...

www.cognitivelinguistics.org › about-c... ▼ Oldal lefordítása
About Cognitive linguistics - Cognitive Linguistics
Cognitive Linguistics grew out of the work of a number of researchers active in the ... In 2005
the Governing Board voted to take the conference to China, and a ...
Írta: M Logo - Kapcsolódó cikkek

www.psychologytoday.com › intl › blog ▼ Oldal lefordítása
What Is Cognitive Linguistics? | Psychology Today
Cognitive linguistics is a radical and exciting approach to language and mind. Find out what
makes it new and unique. ... Posted Jul 12, 2019. Facebook icon ...

www.oxfordhandbooks.com › oxfordhb ▼ Oldal lefordítása
Introducing Cognitive Linguistics - Oxford Handbooks
This article also discusses how cognitive linguistics and generative grammar can both ... In
2003, a second journal specifically devoted to research in Cognitive ...
Írta: D Geeraerts - Idézetek száma: 268 - Kapcsolódó cikkek

www.thoughtco.com › what-is-cognitiv... ▼ Oldal lefordítása
Definition and Discussion of Cognitive Linguistics - ThoughtCo
2020. febr. 12. - Cognitive linguistics is a cluster of overlapping approaches to the study of ...
In the introduction to Cognitive Linguistics: Basic Readings (2006), ...

www.researchgate.net › publication › 2... - Oldal lefordítása
(PDF) Introducing Cognitive Linguistics - ResearchGate
2020. ápr. 14. - cognitive linguist. The journal Cognitive Linguistics, which was founded by Dirk
Geeraerts in 1990, is the official journal of the ICLA. In 2003, a ...

4.1.1.3. Keresés tudományos adatbázisban

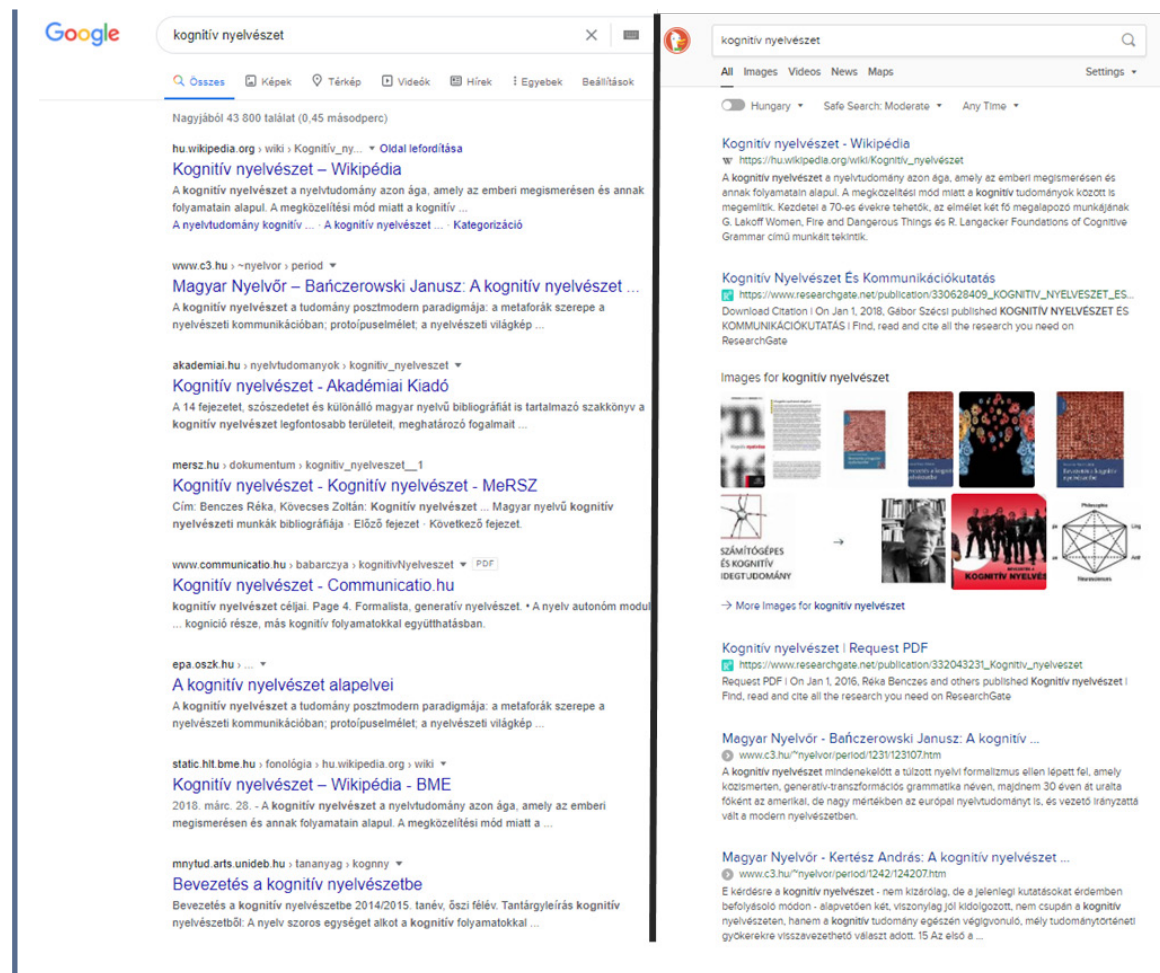
Szintén a Google egyik keresőfelülete a **Google Scholar**, amely kifejezetten tudományos munkák keresését teszi lehetővé. Itt is alkalmazhatjuk az előbbi keresőoperátorokat, azonban a speciális keresési felülethez hasonlóan, ezen a felületen is van lehetőségünk szűrésre. A bal oldali részen (az ábrán sárgával jelölve) dátumra szűrhetünk, illetve bekapcsolhatjuk a szabadalmak és idézetek számának megjelenítését, így láthatjuk, hogy hányszor hivatkoztak eddig az adott közleményre (pirossal jelölve). A találatok címében fel van tüntetve a műfaj (zölddel jelölve), tőlük jobbra pedig a formátumuk és a tanulmányt tartalmazó oldal URL-je (késsel jelölve).

16. ábra. A Google Scholar keresőfelülete

The screenshot displays the Google Scholar search results for the query "Ronald W. Langacker". The search bar at the top shows the query and a magnifying glass icon. Below the search bar, the results are listed. On the left side, there is a sidebar with search filters, including "Bármikor" (Anytime), "2020 óta" (Since 2020), "2019 óta" (Since 2019), "2016 óta" (Since 2016), and "Egyéni tartomány..." (Individual range...). There are also checkboxes for "szabadalmak is" (Include patents), "idézetek megjelenítése" (Show citations), and "Értesítés létrehozása" (Create alert). The search results are sorted by "Rendezés relevancia szerint" (Sort by relevance). The first result is "Cognitive grammar: A basic introduction" by Ronald W. Langacker, published in 2008. It is a book from Google Books. The result shows a citation count of 4779 and a link to "Kapcsolódó cikkek" (Related articles). The second result is "Foundations of cognitive grammar: Theoretical prerequisites" by Ronald W. Langacker, published in 1987. It is a book from Google Books. The result shows a citation count of 18064 and a link to "Kapcsolódó cikkek" (Related articles). The third result is "Foundations of Cognitive Grammar: descriptive application. Volume 2" by Ronald W. Langacker, published in 1987. It is a book from Google Books. The result shows a citation count of 4996 and a link to "Kapcsolódó cikkek" (Related articles). The fourth result is "Concept, image, and symbol: The cognitive basis of grammar" by Ronald W. Langacker, published in 2002. It is a book from Google Books. The result shows a citation count of 4503 and a link to "Kapcsolódó cikkek" (Related articles). The fifth result is "Grammar and conceptualization" by Ronald W. Langacker, published in 2000. It is a book from Google Books. The result shows a citation count of 3223 and a link to "Kapcsolódó cikkek" (Related articles). The sixth result is "Reference-point constructions" by Ronald W. Langacker, published in 1993. It is a book from Degruyter.com. The result shows a citation count of 1311 and a link to "Kapcsolódó cikkek" (Related articles). On the right side, there is a link to "[PDF] academia.edu".

4.1.1.4. A személyre szabott találatok elkerülése: Duck Duck Go

A Google találatait gyakran befolyásolják a rólunk gyűjtött adatok: mire kattintottunk korábban, milyen tartalmakat keresünk a legtöbbször, hol és mit vásároltunk, hol élünk stb. Számos hasonló adat alapján a Google rangsorolja a találatokat, és azokat helyezi előrébb a rangsorban, amelyek ezen adatok alapján relevánsnak tűnnek a felhasználó számára. Ez gyakran sokat segít a megfelelő információk megtalálásában. Amennyiben viszont a találatok ilyenfajta súlyozását szeretnénk elkerülni, és olyan találatokat szeretnénk kapni, amelyek bárkire bárhol a világban érvényesek lehetnek, akkor használhatjuk a **Duck Duck Go** keresőmotort. A 17. ábra a Google és a Duck Duck Go találatai közti különbségeket mutatja. Itt láthatjuk, hogy bár hasonló találatok is előfordulnak, ezek eltérő sorrendben jelennek meg. A Google-találatok esetében befolyásolja a sorrendet a relevancia, vagyis hogy az adott személynek mely adatok lehetnek fontosabbak.



17. ábra. A Google és a Duck Duck Go keresőmotorok találati listái

4.1.2. Reguláris kifejezések

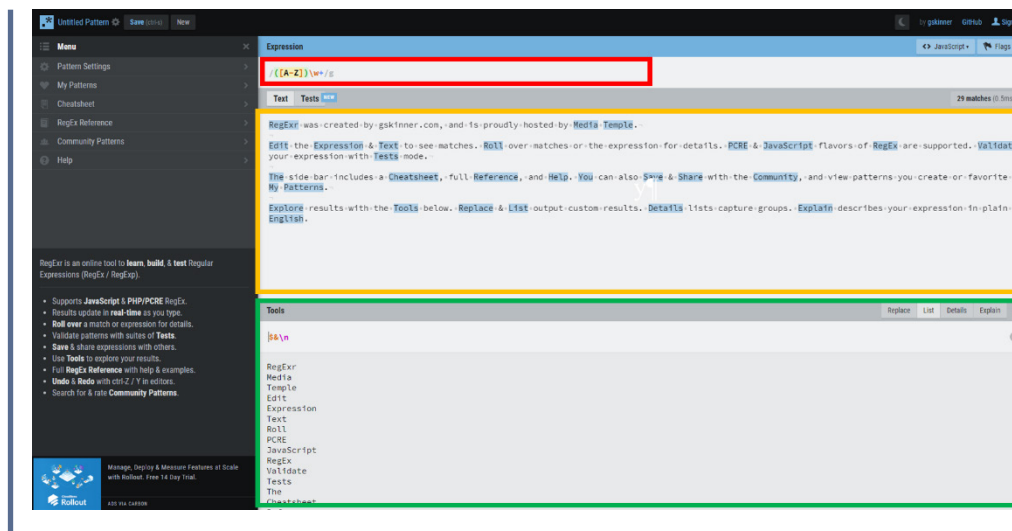
Míg a webes keresések főleg a szakirodalom felkutatásához és esetlegesen az adatgyűjtéshez nyújthatnak segítséget, addig a reguláris kifejezések más természetű keresést tesznek lehetővé számunkra. Segítségükkel az összegyűjtött nyelvi adatok halmazában mintázatok, összefüggéseket kereshetünk. Alkalmazhatók a legtöbb korpuszlekérdező felületen is. A *reguláris kifejezések* (vö. ang. *regular expressions*, röviden: *regex*) formális kifejezések, amelyek segítségével bizonyos mintázatokkal való egyezésekre kereshetünk szöveges adathalmazban (KARTTUNEN et al. 1997: 2). Szintaxisuk egyszerű, azonban a kifejezések összeépíthetők egymással, így bonyolultabb mintázatok keresésére is lehetőséget adnak (KARTTUNEN et al. 1997: 2). Mintázatok alatt például dátumok, e-mail címek, telefonszámok sémája, bizonyos karakter-együttállások értendők. A *regex*et használják például weboldalak regisztrációs felületén az e-mail címek, telefonszámok validálására. Ez azt jelenti, hogy ellenőrzik, hogy a begépelt adat egyezik-e az e-mail címek vagy telefonszámok sémájával. Az e-mail címek például jellegzetesen tartalmazznak @ jelet és egy tartománynevet (pl. @elte.hu), a telefonszámok jellegzetesen tartalmazznak országhívót és bizonyos számú egész számot. A *regex* operátorok hasonlóan működnek, mint a Google operátorai, azonban sokkal több lehetőséget biztosítanak a keresésre. Segítségükkel nem egyszerű adatkeresést hajtunk végre, mint a kulcsszavas keresés esetében, hanem adatbázisokban komplex összefüggések alapján kereshetünk, vagy akár adatbázisokat építhetünk az így rendszerezett adatokból.

A reguláris kifejezéseket ugyan a programozásban is sokoldalúan használják, egyszerű felületek segítségével programozás nélkül, az alapvető operátorok ismeretével is nagy hasznát vehetjük ennek a formalizálási lehetőségnek. Az alábbiakban az alapvető operátorokat és műveleteket mutatom be. A különböző felületeken eltérően működhetnek bizonyos kifejezések, de az alább ismertetett operátorok többnyire minden hasonló webes felületen ugyanúgy működnek.

4.1.2.1. Hol használható reguláris kifejezés?

A reguláris kifejezések alkalmazásához olyan felületre van szükségünk, amely támogatja használatukat. A mindennapokban gyakran igénybe vett programok közül a Libre Office Writer, az MS Office Word bizonyos verziói és a Notepad++ teszik ezt lehetővé. Ezen kívül a legtöbb korpuszlekérdező-felületen is tudjuk őket használni. Jelen segédanyag egy kezdőbarát online felületet mutat be bővebben, amely lehetővé teszi nagyobb mennyiségű szöveges adat importálását és exportálását, valamint tartalmaz segédletet is a *regex* használatához.

Ez az oldal a <https://regexr.com/>. Mivel online felület, ezért bármilyen operációs rendszeren egyszerűen használhatjuk. Az oldal akár regisztráció nélkül és ingyenesen működik. A felső sávba (a 18. ábrán pirossal jelölve) írhatjuk a kereső-kifejezést. Alatta helyezhetjük el egyszerűen bemásolva (Ctrl+C, Ctrl+V) a szöveget, amelyben keresni kívánunk (az ábrán sárgával jelölve). A legalsó sávban (zölddel jelölve) a **List** menüre kattintva kilistázhatjuk a találatokat, ezt innen egyszerűen ki is másolhatjuk, elmenthetjük. A **Replace** menü segítségével kicserélhetünk bizonyos karaktersorozatokat egy másik karaktersorozatra. Megadhatjuk a kicserélendő karaktersorozat mintázatát, illetve, hogy ezeket mire cserélje ki a program. A *Details* és az *Explain* megmutatja, hogy az adott találatra miért érvényes a fent megadott reguláris kifejezés a szabályok szerint. A 18. ábrán az **Expression** sávba írt kifejezés a szövegben található nagy kezdőbetűs szavakat listázza ki.



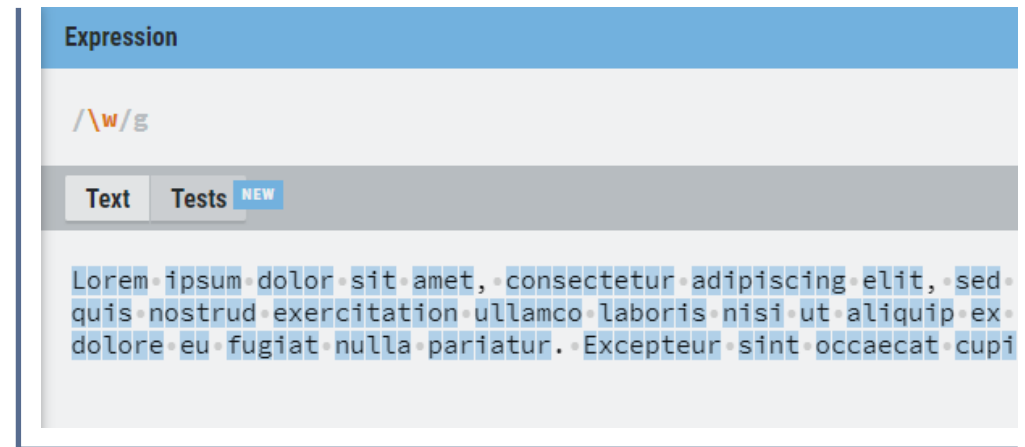
18. ábra. A regexr.com felülete

4.1.2.2. Típusegyezések

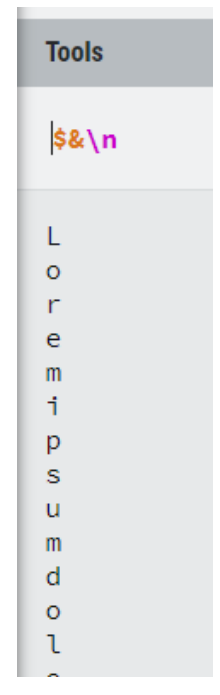
A legegyszerűbb keresés, ha egy bizonyos típusú adatra keresünk a szövegben, például szóra, egész számra, szóközre stb.

- .** Bármilyen karaktert helyettesíthetünk vele.
- \w** Szavakat alkotó karaktereket helyettesít. Ékezetes betűkkel nem mindig működik.
- \W** Visszaad minden olyan szövegrészt, ami nem számít szóznak.
- \s** Visszaad minden üres helyet (*whitespace*), például szóközt, tabulátort, sorvégi entert.
- \S** Minden olyan karaktert visszaad, ami nem a \s-nél meghatározott üres hely.
- \g** Nemcsak az első találatot listázza, hanem az egész szövegben keres.

A 19. ábrán láthatjuk, hogy a `\w` operátort beírva a program minden szóalkotó karaktert kékkel jelölt. Listaként ez a 20. ábrán látható módon jelenik meg.



19. ábra. Keresés a `\w` operátorral



20. ábra. A `\w` keresési eredménye kilistázva a Regexp felületén

4.1.2.3. Karakteregyezések, karakterosztályok

. (pont)

A pont karaktert helyettesítő operátor. Bármilyen karaktert helyettesíthet. Így például a *.al* kifejezés visszaadja a *hal*, *bal* és *dal* szavakat is.

A karakterosztályokat elhatároló operátorokkal bizonyos típusú karaktercsoportokra kereshetünk.⁸

[karakterek]

A szögletes zárójelben felsorolt karakterek valamelyikével megegyező karaktert adja vissza. Ha például arra keresünk, hogy *b[éa]ka*, akkor a *béka* és *baka* szavakat is visszaadja, azonban a *bika* szót már nem.

[^karakterek]

A *^* után a szögletes zárójelek között felsorolt karakterek egyikével sem egyező karaktereket ad vissza. Például a *b[^éa]ka* kifejezés visszaadja a *bika* szót, de a *baka* és a *béka* nem fog a találatok közt szerepelni.

[karakter - karakter]

A szögletes zárójelben kötőjellel összekapcsolt karaktertartomány az első paraméterben meghatározott karaktertől a második paraméterben meghatározottig bármelyik karakterre adhat találatot az angol ábécé szerint. Fontos, hogy az ékezetes betűk ebben az ábécében a *z* után jönnek. Tehát ha azt íránk be, hogy *b[a-i]ka*, akkor erre mind a *baka*, mind a *bika* találatokat megkapjuk, azonban a *békát* nem találjuk meg a listában.

⁸ A példákat a http://vbence.web.elte.hu/regex_leiras.html oldalon található leírás példái inspirálták (Utolsó elérés: 2022. 08. 16.).

[szám-szám]

Az első és a második megadott szám közti számokra keres rá, a végpontokat is beleértve. Tehát a *[0-5]* kifejezés visszaadja az összes számot 0-tól 5-ig. Mivel karakterekre szűr a kifejezés, a többjegyű számokat is visszaadja karakterenként, tehát például az 55-öt két darab ötösként.

4.1.2.4. Mintázat többszörözése

Az alábbi operátorok azt jelzik, hogy az előttük lévő *regex*-mintának valamilyen módon többször kell szerepelnie a találatban.

- ?** Az előtte lévő minta 0 vagy 1 alkalommal forduljon elő a találatban. Például a *.?al* kifejezés nemcsak a *hal*, *bal*, *dal* karaktercsoportokat adja vissza, hanem az *al*-t is.
- +** Az előtte álló minta egy vagy több alkalommal forduljon elő a találatban. Például az *ál+* kifejezés visszaadja az *áll* és az *ál* karakterláncot is.
- *** Az előtte lévő minta 0 vagy több alkalommal forduljon elő a találatban. Például az *ál** már nemcsak az *áll* és *ál*-t adja vissza, hanem önmagában az *á*-t is, mivel megengedi azt is, hogy egyszer sem szerepeljen az *l* a karaktersorozatban.
- {n}** Egy karakter *n*-szer forduljon elő. Például az *ar{2}a* kifejezés visszaadja az *arra* szót, de az *ara*-t nem.

{m,n} A karakter előfordulásának számát minimum (m) és maximum (n) értékekkel is megadhatjuk. Például a `\w{1,3}ék` kifejezéssel olyan szóra keresünk, amelyben az -ék végződés előtt legalább egy, legfeljebb három karakter áll. Így ez a kifejezés visszaadhatja a *lék*, *fék*, *szék*, *Eszék* szavakat is.

4.1.2.5. Logikai operátorok

| és () A `|` operátorral vagy-kapcsolatot tudunk kifejezni. Ez azt jelenti, hogy a `|` két oldalára írt karaktereknek vagy egyike vagy másika fog bekerülni a találatokba. A kifejezést a gömbölyű zárójelek segítségével határolhatjuk el, ha szóbelseji karakterekre keresünk. Például a `ka(lap|bát)` kifejezésre megkapjuk a *kalap* és a *kabát* szavakat is vagy a `fotel(ek|ok)`-ra a *fotelek* és a *fotelok* szavakat is. A reguláris kifejezések kreatív használhatóságát mutatja, hogy az utóbbi kifejezést máshogyan is felírhatjuk: `fotel(e|o)k`.

4.1.2.6. Escape

Ha olyan karakterekre szeretnénk keresni a szövegben, amelyek *regex* operátorok is, akkor jelölnünk kell a reguláris kifejezésünkben, hogy ezt a karaktert nem operátorként szeretnénk használni. Ezt a műveletet hívjuk *escape*-elésnek, amelyet a `\` operátorral, vagyis *escape* karakterrel tudunk elvégezni. Tehát ha például mondatvégi kérdőjelekre akarunk keresni, akkor a `\?` kifejezést kell használnunk a sima kérdőjel helyett, hiszen a `?` *regex* operátor.

4.1.2.7. Reguláris kifejezések használata korpuszlekérdezőkben

Ahogy a **3.1.1. fejezet**ben említettük, reguláris kifejezéseket alkalmazhatunk olyan korpuszlekérdező-felületeken is, mint amilyen például az MNSz2. Ilyen esetben a lekérdezés típusát át kell állítanunk szótőre, szóalakra vagy részletes keresésre, és máris kiegészíthetjük keresésünket reguláris kifejezésekkel. Ha például az összes *poharat* és *pohárt* adatot ki szeretnénk listázni, akkor ezt megtehetjük a „*poh[aá]ra?t*” kifejezés beírásával (lásd 21. ábra).

The screenshot shows the MNSz2 corpus search interface. At the top, there are three dropdown menus: 'korpusz:' set to 'MNSZ2', 'lekérdezés típusa:' set to 'szóalak', and 'szóalak:' containing the regular expression 'poh[aá]ra?t'. To the right of the 'szóalak:' field is a checkbox labeled 'kisbetű/nagybetű számít:'. Below these fields are two buttons: 'Konkordancia készítése' and 'Törlés'. The main area of the interface displays a list of search results, which are excerpts from a text. The word 'poharat' is highlighted in red in several places, and 'pohárt' is highlighted in green in one place. The text is a Hungarian narrative about a person drinking from a cup.

21. ábra. Reguláris kifejezések használata az MNSz2-ben

4.2. Adatbekérő kérdőívek készítése

A kérdőíves adatfelvétel módszereivel a **3.1.2. fejezet** foglalkozik. Ebben a fejezetben a kérdőívkészítés technikai-informatikai hátterét tekintjük át, két gyakran alkalmazott szoftverre fókuszálva. Ezek az online alkalmazások megkönnyítik a kérdőívek készítését, a kutatási adatok feldolgozását, kódolását és a kérdőív terjesztését a potenciális adatközlők körében.

Az online eszközök előnye, hogy bármilyen platformról, operációs rendszerről elérhetőek mind a kérdések, mind a megválaszolt kérdőívek adatai. Ezeket az adatokat a szoftverek vizuálisan is megjelenítik számunkra diagramok formájában, amelyeket elmenthetünk, felhasználhatunk, az adatok letöltésével pedig saját kimutatásokat is készíthetünk. A legtöbb program lehetővé teszi, hogy csak egy bizonyos link birtokában férhessenek hozzá az adatközlők a kérdőívhez, így a kérdőíves vizsgálatban résztvevők köre lehet nyitott és zárt is. A kitöltés helyét bízhatjuk akár az adatközlőkre, de zárt térben, ellenőrzött körülmények között is történhet a válaszadás.

Az online kérdőívkészítőknek hátránya lehet azonban, hogy nem feltétlenül tartalmazznak minden számunkra szükséges választípust, illetve a megosztható linket a kutatásban résztvevők továbbíthatják másoknak. Ez utóbbi eset megelőzésére egyes kérdőívkészítőkben jelszavas védelmet alkalmazhatunk, a jelszót pedig megváltoztathatjuk bizonyos időközönként.

4.2.1. Google Űrlapok

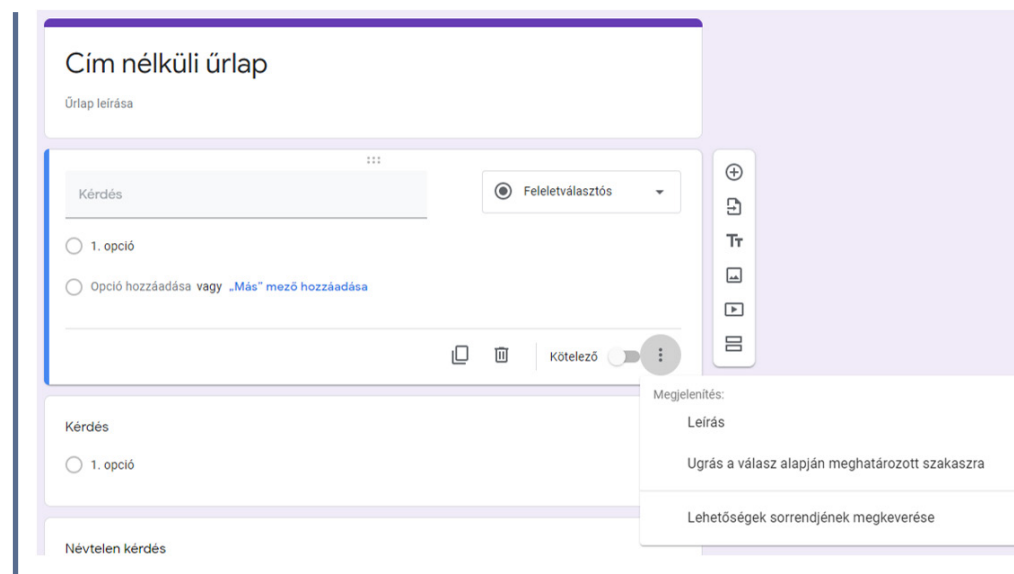
Talán a legismertebb és leginkább alkalmazott kérdőívkészítő alkalmazás a *Google Űrlapok*. Előnye, hogy egyszerűen kezelhető, használata ingyenes, magyar nyelven is elérhető, és az általa nyert adatokat egyszerűen letölthetjük Excel-táblázatként. Használatához

csupán Google-fiók regisztrálása szükséges. Hátránya, hogy a válaszokhoz nem lehet időkorlátot hozzárendelni, tehát azt például nem szabályozhatjuk, hogy maximum mennyi időt tölthetnek a válaszadók egy kérdés megválaszolásával.

A Google Űrlapok elkészítése

A Google Űrlapokat a <https://docs.google.com/forms> oldalon érhetjük el, Google-bejelentkezés után. A kezdőlapom láthatók az esetlegesen korábban készített kérdőívek. A jobb alsó sarokban található „+” jel segítségével készíthetünk újabb kérdőívet. A cím és az űrlap leírása után adhatjuk hozzá a kérdések (22. ábra), amelyeket panelekkel csoportosíthatunk. Minden egyes kérdésnél beállíthatjuk, hogy kötelező-e az adott kérdés kitöltése (Kötelező kapcsoló); előre meghatározott válaszok esetén a válaszlehetőségek véletlenszerű sorrendben jelenjenek-e meg (a három pöttyre kattintva a 'Lehetőségek sorrendjének megkeverése' menüt kell választani); illetve a válasz alapján melyik kérdéscsoportra vagy kérdésre ugorjon a kérdőív (ugyanitt az 'Ugrás a válasz alapján meghatározott szakaszra' menüpontra kattintva).

22. ábra. Kérdés létrehozása
a Google Űrlapokban



Az oldal jobb középső részén található **Beállítások** menüre kattintva találjuk a teljes kérdőívre vonatkozó beállításokat (23. ábra). Itt beállíthatjuk, hogy a kérdőív gyűjtse-e a kitöltők e-mail-címeit, ugyanazon személy többször is kitöltheti-e az űrlapot, illetve hogy a kitöltő személyek láthatják-e az eredményeket. A **Prezentáció** panelen beállíthatjuk, hogy látszódjon-e a kitöltöttség állapotát mutató folyamatjelző, a kérdések össze legyenek-e keverve, illetve a kitöltést követően milyen üzenetet lássanak a kitöltők. Az űrlapunkat tesztnek minősíthetjük, ami azt jelenti, hogy pontszámokat és érdemjegyeket rendelhetünk az űrlaphoz. Ebben a módban lehetőség van arra, hogy a kérdésekhez helyes válaszokat adjunk meg, amelyek alapján a program kiértékeli a kitöltők válaszait, majd egy érdemjeggyel értékeli azokat.

23. ábra. Az űrlapbeállítások panelja

Kérdések Válaszok 3 Beállítások

Beállítások

Beállítás tesztként
Hozzárendelheti a pontszámokat, megadhatja a válaszokat, és automatikus visszajelzést küldhet

Válaszok
A válaszok összegyűjtési és védelmi módjának kezelése

E-mail-címek begyűjtése

A válaszok másolatának elküldése a válaszadóknak
Megköveteli az E-mail-címek begyűjtése beállítását

Válasz szerkesztésének engedélyezése
A válaszok módosíthatók a beküldésük után

BEJELENTKEZÉS SZÜKSÉGES

Korlátozás egy válaszra

Prezentáció
Az űrlap és a válaszok megjelenítési módjának kezelése

ŰRLAP PREZENTÁCIÓJA

Folyamatjelző megjelenítése

Kérdések sorrendjének összekeverése

BEKÜLDÉS UTÁN

Megerősítő üzenet
Választ rögzítettük.

Link megjelenítése, melyen újabb válasz küldhető be

Eredmény-összefoglaló megtekintése
Az eredmények összefoglalójának megosztása a válaszadókkal

KORLÁTOZÁSOK:

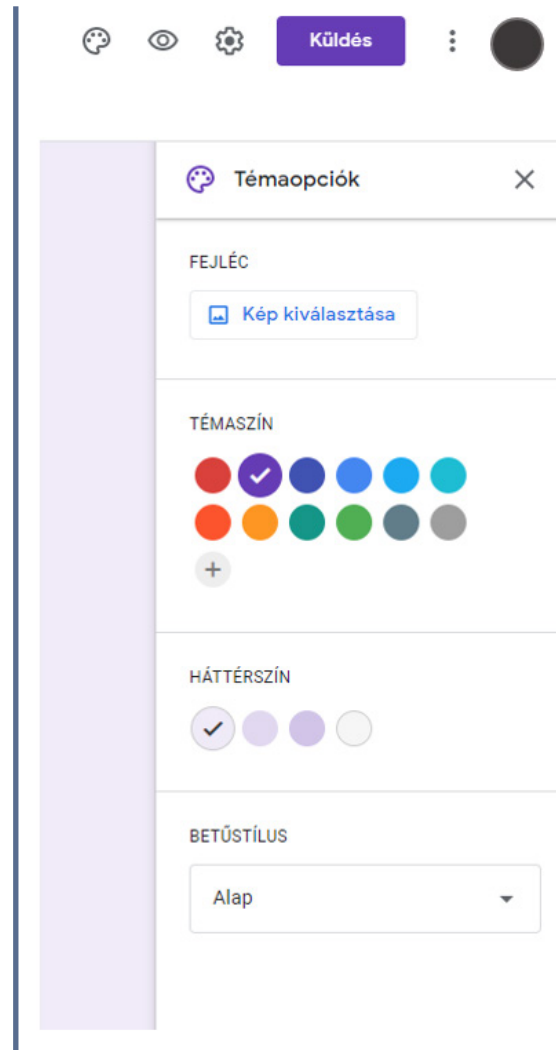
Automatikus mentés letiltása az összes válaszadónál

Az oldal jobb felső részén található festékpalletta 🎨 és szem 👁 ikon segítségével az űrlap kinézetét állíthatjuk be. A festékpalettára 🎨 kattintva kiválaszthatjuk az űrlap háttérszínét, betűtípusát, valamint képet szűrhatunk be fejlécként. A szem 👁 ikonra kattintva az űrlap előnézetét kapjuk, így ellenőrizhetjük, hogy valóban úgy néz-e ki és úgy működik-e a kérdőívünk, mint ahogyan elterveztük (24. ábra).

Választípusok és adattípusok

A kérdéseknél beállíthatjuk, hogy milyen típusú választ várunk a kitöltőktől. Alapvetően a *Feleletválasztós* lehetőség van kijelölve, azonban emellett még több más típus közül választhatunk. Az alábbiakban ezeket a lehetőségeket részletezem, illetve azt, hogy az egyes választípusok milyen adatokat adnak vissza, és ezek hogyan jelennek meg táblázatkezelőben.


- ¶ A *Rövid válasz* típusnál a kitöltő egy rövid egysoros választ adhat meg. A begyűjtött adat tehát ebben az esetben szöveges.
- ¶ A *Bekezdés* szintén hasonló, azonban ebben az esetben hosszabb szöveges válasz is adható.
- ¶ A *Feleletválasztós* lehetőségnél több lehetséges választ is megadhatunk, illetve az *Egyéb* lehetőség engedélyezésével a válaszadó egyéni választ is megadhat. Ezek közül csak egyet választhat a kitöltő. A kapott adat tehát egy véges elemhalmazból származik, amely lehet szöveg vagy szám.



24. ábra. Az űrlap kinézetének beállításai

- ¶ A *Jelölőnégyzetek*nél szintén több válaszlehetőséget adhatunk meg, azonban a kitöltő itt több lehetőséget is kiválaszthat egyszerre. Ebben az esetben a kapott adat egy karaktersorozatokból vagy számokból álló lista lesz.
- ¶ A *Legördülő listát* választva az előre megadott válaszlehetőségek közül egyet választhat a kitöltő. Tulajdonképpen hasonló a *Feleletválasztós* kérdéshez, azonban ennél a lehetőségnél a válaszok rendezettebben jelennek meg nagyobb számú opció esetén is. A másik különbség, hogy itt nem adhatunk meg *Egyéb* lehetőséget. A válaszokból származó adatok tehát itt is egy véges listából kiválasztott karaktersorozatok vagy számok lesznek.
- ¶ A *Fájl feltöltése* lehetőség képek, dokumentumok feltöltését teszi lehetővé a kitöltők számára. Ez akkor lehet hasznos, ha például valamilyen rajz, ábra elkészítését kérjük egy kísérlet során. A kapott adat ebben az esetben egy fájl, amely lehet kép, szöveges dokumentum vagy egyéb fájltypus.
- ¶ A *Lineáris skálánál* a válaszadó egy Likert-skálán jelölheti be a választát. A skála két végéhez leírást is rendelhetünk. Itt a kapott adat egy darab egész szám lesz.
- ¶ A *Feleletválasztós rácsot* akkor célszerű alkalmazni, ha hasonló szempontok alapján szeretnénk értékelni bizonyos állításokat, fogalmakat. A rács sorai az egyes állítások, fogalmak stb. lesznek, az oszlopok pedig a kiválasztható tulajdonságok, válaszok, amelyekből csak pontosan egyet választhat az adatközlő.
- ¶ A *Jelölőnégyzetrács* hasonló a *Feleletválasztós rács*hoz, itt azonban soronként több választ is megadhat a kitöltő.
- ¶ A *Dátum* és az *Idő* válaszlehetőséggel a kitöltő egy dátumot vagy időpontot adhat meg. A kapott adat dátumformátumú lesz.

Adatok és grafikonok exportálása

A *Google Űrlapok* nagyon egyszerűvé teszi az adatok gyűjtését és elemzését. A válaszok Google-táblázatba illeszthetők vagy letölthetők .csv formátumban, amelyet importálhatunk táblázatkezelőbe. Ha a kérdőívszerkesztőben a **Válaszok** fület választjuk, a felső sávban a  ikonra kattintunk és az 'Új táblázat létrehozása'-t választjuk, majd a **Létrehozás** gombra kattintunk, megnyílik egy Google táblázat az összes adatunkkal. Ezt a táblázatot letölthetjük többféle formátumban a **Fájl » Letöltés** menüpontokat választva. Ha itt a .xlsx formátumot választjuk, akkor a dokumentum közvetlenül megnyitható lesz Excelben. A táblázatot .csv formátumban is letölthetjük, a **Válaszok** fülön a három pöttyre kattintva a **'Válaszok letöltése' (.csv)** lehetőséget választva. Ez a formátum is importálható táblázatkezelőbe. Ennek módját a **4.4.1. fejezet** részletezi.

A **Válaszok** fülön automatikusan létrehozott grafikonokat is találunk a válaszadatokból létrehozva. Ezeket a másolás ikonra kattintva tudjuk képként beilleszteni szöveges dokumentumba vagy prezentációba a Ctrl + V billentyűkombinációval.

4.2.2. Online-kerdoiv.com

A *Google Űrlapok* ugyan egyszerű és ingyenes megoldást biztosít egyszerűbb kérdőívek elkészítéséhez és feldolgozásához, az alkalmazásból számos olyan funkció hiányzik, amelyek egy tudományos kutatásnál hasznosak lehetnek. Ilyen például az időkorlát megadása és a kérdőív védelme jelszóval. Az *online-kerdoiv.com* több ilyen funkciót is lehetővé tesz, ezért érdemes megemlíteni.

Az *online-kerdoiv.com* ingyenes változata mindössze 3 kérdőív létrehozását teszi lehetővé, kérdőívenként 25 kérdéssel, kérdésenként 25 válaszlehetőséggel, és mindössze 25 kitöltéssel. Azonban akár kéthónapos előfizetést is vásárolhatunk, tehát nem szükséges egész éves előfizetés, ha csak rövidebb ideig szeretnénk használni.

Az elv nagyon hasonló, mint a *Google Űrlapok*nál. A bejelentkezés után láthatjuk a már létrehozott kérdőíveinket, az **Új kérdőív** gombra kattintva pedig újat hozhatunk létre. Az új kérdőív lehet teljesen új, egy korábbi másolata vagy sablonra épülő. A kérdőív szerkesztésének megkezdésekor bekapcsolhatunk GDPR-beállításokat, amelyek segítségével hozzájáruló nyilatkozatot fogadtathatunk el a kitöltőkkel. Ugyanebben a menüben beállíthatunk jelszót a kérdőívhez, illetve számos beállítást végezhetünk a megjelenítéssel és a működéssel kapcsolatban. A megjelenítésnél a *Google Űrlapok*hoz hasonlóan beállíthatjuk a kérdések randomizálását és az állapotjelzőt, amely azt jelzi a kitöltőnek, hogy éppen hol tart a kérdések megválaszolásában. A működésnél meghatározhatjuk többek között a kérdőív lezárásának időpontját, valamint időkorlátot rendelhetünk az egész kérdőívhez, de akár az egyes kérdésekhez is. A **Teszt, vizsgáztatás** menü lehetőséget ad arra, hogy az adott kérdőív tesztként működjön. Ez is nagyon hasonló a *Google Űrlapok*hoz.

Az oldalsó menüben a **'Kérdések'** menüpontra lépve térhetünk vissza a kérdésszerkesztő felülethez. Az egyes kérdéseknél szintén beállítható időkorlát, illetve adhatunk azonosító kódot a kérdésnek, amely később az adatfeldolgozás során hasznos lehet. Itt is bekapcsolhatjuk az **'Egyéb'** válaszlehetőséget. A kérdés típusát is beállíthatjuk, itt azonban több típus közül választhatunk. A *Google Űrlapok*hoz hasonlóan itt is találkozunk *egysoros válasz*, *többsoros válasz*, *dátum*, *lista* lehetőségekkel. Ezekon kívül az alábbi válaszokat és adattípusokat állíthatjuk be.

25. ábra. Az **online-kerdoiv.com** kérdőívkészítő felülete

The screenshot shows the 'Kérdőívek' (Questionnaires) management interface. The left sidebar contains navigation options: 'Kérdések' (Questions), 'Beállítások' (Settings), 'Kinézet' (Appearance), 'Szövegek' (Texts), and 'Események' (Events). The main area is titled 'KÉRDŐÍV' and shows a questionnaire named 'Próba'. Below the title, it states 'Egy oldalon lesz az összes kérdés. Módosítás >'. The 'Bevezető oldal' (Introductory page) section is active, with a 'Kikapcsolása' (Deactivate) button. A text box for the introductory page is provided, along with a '+ kérdés' (Add question) button. Below this, the 'Kérdések (5)' (Questions (5)) section is visible, showing a list of questions. The first question is '1. Kérdés szövege' (Question text), and it has four answer options: '1. válasz' (selected), '2. válasz', '3. válasz', and an empty option.

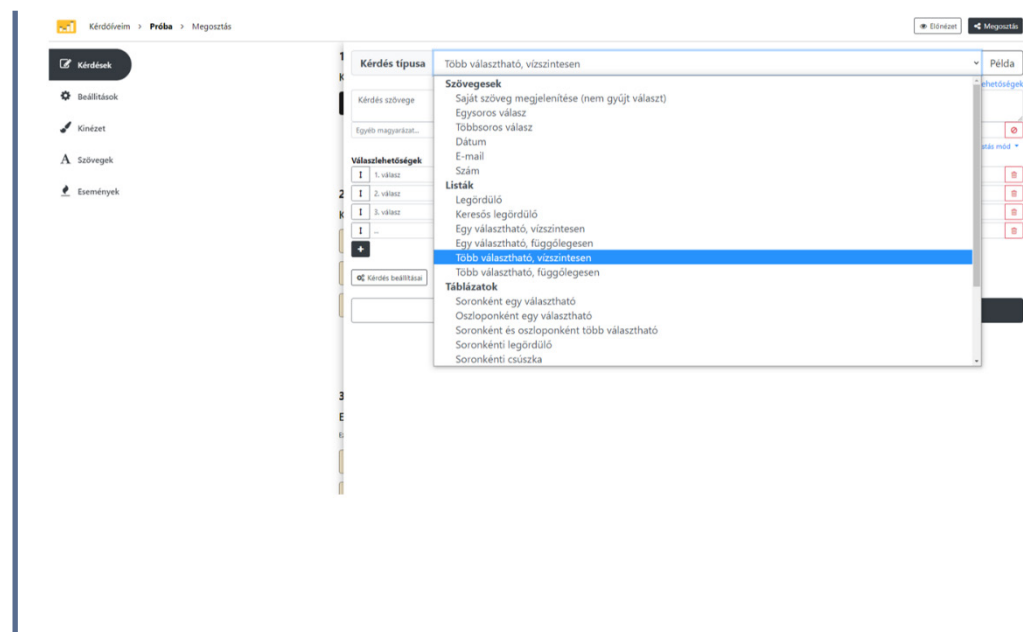
Válasz- és adattípusok

Szövegesek: *E-mail, Szám*

Listák: *Legördülő, Keresős legördülő, Egy választható vízszintesen, Egy választható függőlegesen, Több választható vízszintesen, Több választható függőlegesen*

Táblázatok: *Soranként egy választható, Oszloponként egy választható, Soranként és oszloponként több választható, Sorankénti legördülő, Sorankénti csúszka, Cellánkénti mező*

Speciálisak: *Egyszerre több adat bekérése, Csúszkás értékelőskála, Rangsorolás, Kategorizálás, Értékthalmaz*



26. ábra: Az **online-kerdoiv.com** kérdésszerkesztője a választípusokkal

Adatok és grafikonok exportálása

A kezdőlap a kérdőív melletti **Műveletek** legördülő menüre kattintva, a **'Statisztikák'** menüt választva tudjuk megtekinteni a válaszokat és az azokból generált grafikonokat, statisztikákat. Itt kérdés, válasz, időintervallum és sorszám alapján szűrhetjük a válaszokat. Az eredmények letöltése sajnos csak a fizetős verziókban érhető el, ezekben azonban többféle formátumban tölthetjük le őket, többek között Excelben megnyitható .csv-ben is.

4.3. Táblázatkezelő használata

A kutatások során sokoldalúan alkalmazhatunk táblázatkezelőt. Segítségével rendezhetjük, elemezhetjük az összegyűjtött adatokat, ábrázolhatjuk eredményeinket. Ez a fejezet a táblázatkezelő szoftverek azon funkcióit mutatja be, amelyek segíthetik nyelvészeti kutatások kivitelezését. A konkrét szoftverek közül elsősorban a Microsoft Excel-re támaszkodik, azonban a bemutatott funkciók egyéb táblázatkezelő szoftverek (pl. Libre Office, Google Táblázatok) esetében is hasonlóan működnek.

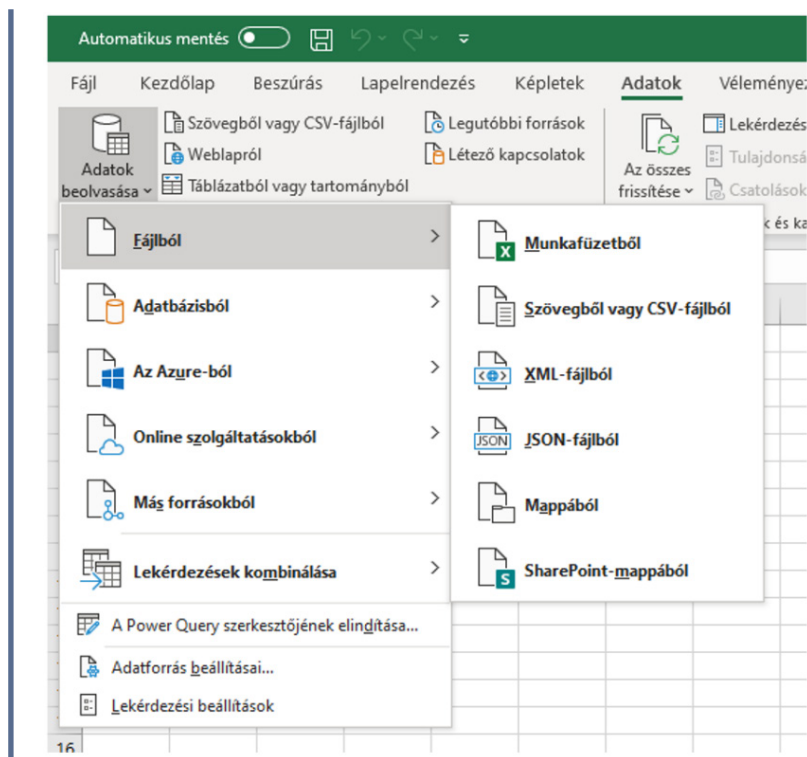
Az Excel központi szerepét az indokolja, hogy egyszerű kezelni, kisebb adathalmazokkal jól működik és széles körben elérhető. Az Excelnek online verziója is van, amely mindenki számára ingyenes. Ugyan a lehetőségek korlátozottak az online verzióban, a csoportmunkát megkönnyítheti, hiszen egyszerre többen is szerkeszthetik ugyanazt a fájlt. Az online verzió az <https://office.live.com/start/excel.aspx> oldalon érhető el.

A fejezet röviden bemutatja az adatok importálásának, a feltételes formázásnak és a függvények használatának módjait, illetve kitér az adatok szűrésére is.

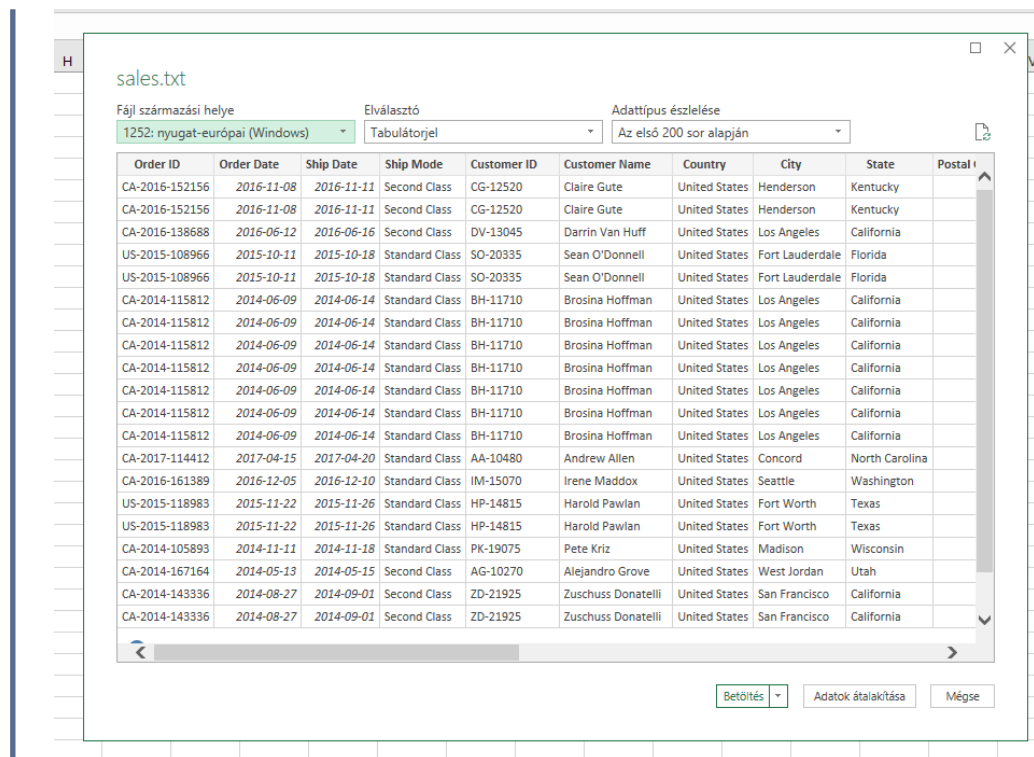
4.3.1. Adatok importálása

Minden adatelemzés adatgyűjtéssel kezdődik. Adatainkat többféle formátumban gyűjthetjük: az online kérdőívekből, korpuszokból kinyert adatok általában már táblázatkezelőben megnyitható formátumban kerülnek elmentésre, azonban előfordul, hogy valamilyen szöveges formátumban (txt, csv stb.) kapjuk ezeket az adatokat. Ebben az esetben a táblázatkezelőbe importálhatjuk ezeket.

Excelben az adatok importálását az **Adatok** fülön tudjuk elvégezni⁹. Itt az **'Adatok beolvasása'** gombra kattintva számos lehetőség közül választhatunk. Ha a **'Fájlból'**-t választjuk, megadhatjuk, hogy milyen típusú fájlból szeretnénk az adatokat importálni (27. ábra). Ezután a fájl kiválasztásával megjelenik egy ablak, amelyben táblázatos formában láthatjuk az adatainkat. Itt még végezhetünk rajtuk módosításokat, majd a betöltés gombra kattintva az adatok megjelennek az Excel-táblázatban (28. ábra).



27. ábra. Adatok importálása Excelben



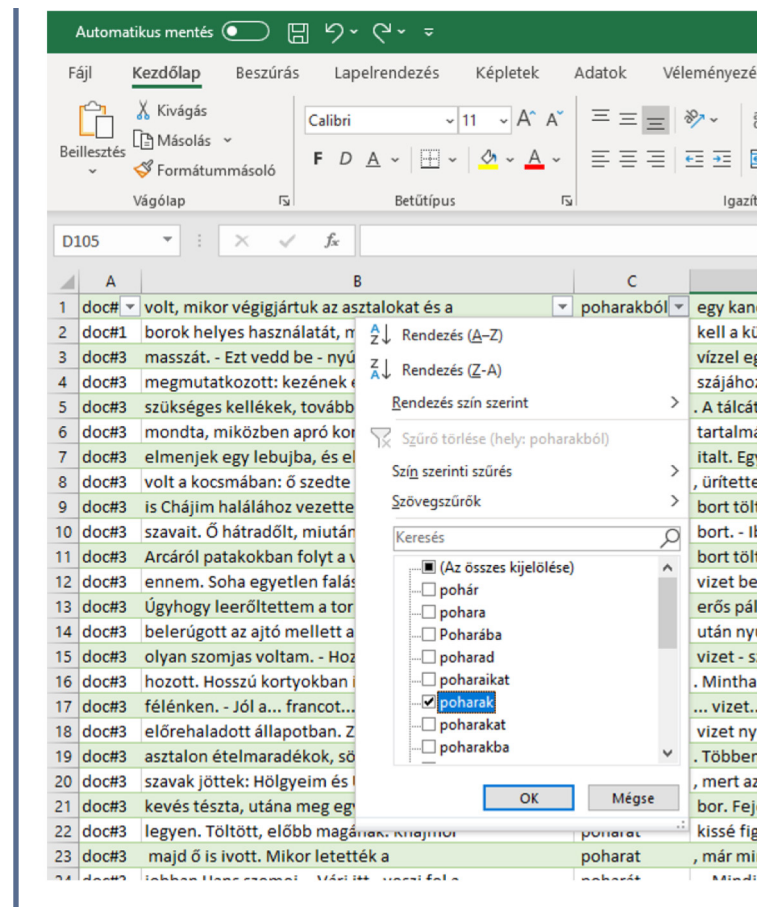
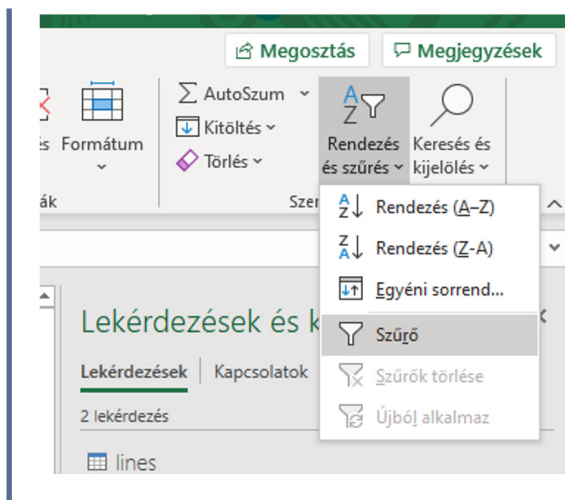
28. ábra. Adatok importálása fájlból

⁹ Fontos: az importálás az Excel különböző verzióiban eltérően működik, itt a 2020-as verziót mutatom be.

4.3.2. Szűrők alkalmazása

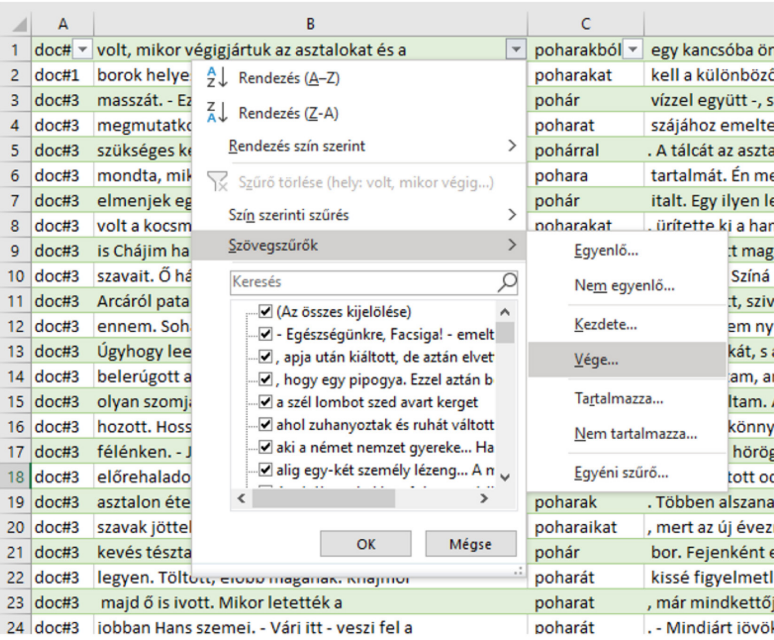
Az adataink jobb áttekinthetősége érdekében alkalmazhatunk szűrőket, amelyekkel bizonyos feltételek alapján szűrhetjük adatainkat. A szűrni kívánt adattábla kijelölése után szűrőt adhatunk az egyes oszlopokhoz. Ezt a szalag **Kezdőlap** fülén tehetjük meg a **'Rendezés és szűrés'** menüt lenyitva a **'Szűrő'** kiválasztásával (29. ábra). A szűrő segítségével ezután szűrhetjük az adatainkat. A 30. ábrán egy MNSz-ből exportált, majd Excelbe importált konkordanciát láthatunk a *pohár* szó és ragozott alakjainak előfordulásairól. A szűrő hozzáadásával az oszlopokon megjelenik egy lenyitható menü. Ha a *pohár* adatokat tartalmazó oszlopon ezt lenyitjuk, kiválaszthatjuk, hogy mely adatokat szeretnénk látni. Például kijelölhetjük a *poharak* szót, így csak az ehhez kapcsolódó konkordanciákat fogja megjeleníteni a táblázat.

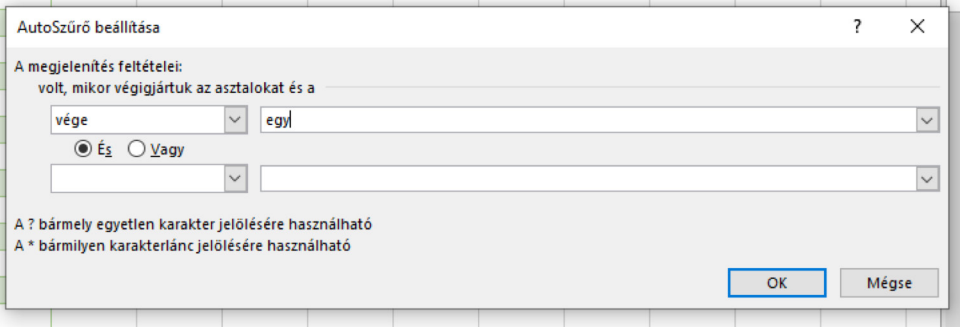
29. ábra. Szűrő hozzáadása az adattáblához



30. ábra. A szűrő alkalmazása

A szűrésnél egyéni feltételeket is megadhatunk, vagy szín és formátum szerint is szűrhetjük a cellákat. Például megadhatjuk, hogy a *pohár* előtti szövegrészlet az *egy* szóra végződjön, így szűrhetünk azokra az adatokra, amelyeknél a *pohár* és ragozott alakjai határozatlan névelőt kaptak (31. ábra).





	A	B	C	D
1	doc#	volt, mikor végigjártuk az asztalokat és a	poharakból	egy kancsóba öntöttük a maradékot. Ezt
3	doc#3	masszát. - Ezt vedd be - nyújtotta nekem egy	pohár	vízzel együtt -, segít elaludnod, segít
10	doc#3	szavait. Ő hátradől, miután kiivott egy	pohár	bort. - Ibn Színá halála után szokásommá
12	doc#3	ennem. Soha egyetlen falás kenyeret vagy egy	pohár	vizet be nem nyújtottatok. Semmit, semmit
15	doc#3	olyan szomjas voltam. - Hozzon egy	pohár	vizet - szóltam. A tag kiment, visszajött
17	doc#3	félénken. - Jól a... francot... hozz egy	pohár	... vizet... - hörögte az öreg. Kezével
18	doc#3	előrehaladott állapotban. Zizmalki egy	pohár	vizet nyújtott oda neki, de az öreg nem
21	doc#3	keves tészta, utána meg egy üveg sör és egy	pohár	bor. Fejenként egy pakli dohány. Khajmor
30	doc#3	olyan, mint meginni apró cseppekben egy	pohár	húszesztendős bort, gondolta Khajmor, miközben
32	doc#3	fel. Most ingatom a fejem, iszom egy	pohár	vizet. A hőség őrjási, szerencsére nekem
33	doc#3	keves szeszt vett ki. Szirupot töltött egy	pohárba	, úgy egyharmadáig, a többi meg feltöltötte
39	doc#5	érzelmes vagyok jól tudom Uram add hogy egy	pohár	bor egyhítsen végre sorsomon italok mágiák
41	doc#6	szerette volna, hogy a nagy melegben egy-egy	pohár	jó hideg vizet adjon az embereknek, és
45	doc#31	hátulról, diszkrétén: " Még parancsol egy	pohárral	? " Ő bólintott, s már nyúlt is a
52	doc#31	vigasztaltuk, kínáltuk étellel-itallal, de csak egy	pohár	vizet kért, abba is csak belekóstolt. Megnéztem
54	doc#31	nyugalommal, aztán zsíros kenyeret és egy	pohár	bort kért. Koccintottunk egészségünkre...
72	doc#63	pálinkát. Mikor az ember éhes, inna egy	pohár	pálinkát. S az vót a baj. Reggel felkelt
73	doc#64	vasárnapom - mondtam a pófának, és meghívtam egy	pohár	italra. Egészen jó kedvem lett. Úgy éjfél
75	doc#66	csoportot az üzemi klubba. - Megiszunk még egy	pohárral	- folytatta mosolyogva Tisztafej -, mindazok
76	doc#66	bekente gépszírral. - És megiszunk még egy	pohárral	... Együtt... - Háháhu! Háháhu!
79	doc#66	földbegyökerezetteket Dugigháj. Valaki eltört egy	poharat	, a fényes cipők talpa alatt morzsolódott
91	doc#66	egyedül... " Mit tehetnék? Megiszom egy	pohárral	. Néha többel is, mert iktóztatlan hosszúak

31. ábra. Az "egy" végződésű szövegek szűrése

Sok esetben hasznos, ha kiszűrjük azokat az adatokat, amelyek többször is szerepelnek valamelyik oszlopban. Így például táblázatot készíthetünk arról, hogy milyen alakokban fordul elő a *pohár*. Ezt úgy tehetjük meg, hogy az oszlop tartalmát kimásoljuk, majd beillesztjük a táblázat egy másik pontjára. Ezután az **Adatok** fülön lévő **'Ismétlődések eltávolítása'** menüt választjuk (32. ábra).

32. ábra. Ismétlődő adatok szűrése

The screenshot shows the Microsoft Excel interface with the 'Adatok' (Data) ribbon selected. The 'Ismétlődések eltávolítása' (Remove Duplicates) button is highlighted in the 'Adatkezelés' (Data Management) group. Below the ribbon, a table is displayed with two columns, C and F, containing the word 'pohár' in various contexts.

	C	D	E	F	G	H	I	J
	poharakból	egy kancsóba öntögettük a maradékot. Ezt		poharakból				
	poharakat	kell a különböző borfajtákhoz használni		poharakat				
	pohár	vízzel együtt -, segít elaludnod, segít		pohár				
or a	poharat	szájához emelte, sápadt arca, csaknem mélabús		poharat				
	pohárral	. A tálcát az asztalra helyezte, majd szó		pohárral				
	pohara	tartalmát. Én meglehetősen hamar		pohara				
	pohár	italt. Egy ilyen lebujsban láttam meg Chájimot		pohár				
	poharakat	. üritette ki a hamutartókat, mosta fel		poharakat				
	pohár	bort töltött magának, majd felállt, odament		pohár				
	pohár	bort. - Ibn Színá halála után szokásommá		pohár				
	pohár	bort töltött, szivarából mélyet szívott		pohár				
	pohár	vizet be nem nyújtottatok. Semmit, semmit		pohár				
	pohár	erős pálinkát, s az isteni folyadék el		pohár				
	pohár	után nyúltam, amikor visszafordult, s azt		pohár				
	pohár	vizet - szóltam. A tag kiment, visszajött		pohár				
	pohárral	. Mintha a könnyeket akartam volna pótolni		pohárral				
	pohár	... vizet... - hörögte az öreg. Kezével		pohár				
	pohár	vizet nyújtott oda neki, de az öreg nem		pohár				
	poharak	. Többen alszanak, az asztalra borulva,		poharak				
	poharaikat	, mert az új évezred első pillanata mindjárt		poharaikat				
	pohár	bor. Fejenként egy pakli dohány. Khajmor		pohár				
	poharát	kissé figyelmetlenül töltötte meg, melléje		poharát				
	poharat	, már mindkettőjük szeme csillogott.		poharat				
	poharát	. - Mindjárt jövök. Hova is mennék		poharát				
	poharát	. - A háború minél hamarabbi végére		poharát				
	poharat	. A két lány nevetett. De Ingrid arcán		poharat				
	poharát	Hans, és hatalmasat kacagott. Ivott, majd		poharát				
	poharát	Cseh, s tette vissza az asztalra anélkül		poharát				
	pohara	fenekét bámulta. Nagyon, nagyon szomorú		pohara				
	pohár	húszesztendő bort, gondolta Khajmor, miközben		pohár				
	poharát	? Bocsásson meg, Khajmor, de szokásommal		poharát				
	pohár	vizet. A hőség óriási, szerencsére nekem		pohár				
	pohárba	, úgy egyharmadáig, a többi meg feltöltötte		pohárba				
	poharat	. Mikor inni akart belőle, a leány kikapta		poharat				
	poharat	de az első korty után erős köhögés kapta		poharat				
	poharat	. - Idd ki! És aztán csend legyen,		poharat				
	poharat	és a leány kezébe adta. Az meg letette		poharat				
	pohár	aljáról visszaneéz szemed tükrében vodka-könnycseppek		pohár				

4.3.3. Feltételes formázás

Az adatokat látványosan szűrhetjük a *'feltételes formázás'* lehetőséggel, amely lehetővé teszi, hogy bizonyos feltételek alapján automatikusan megváltozzon a leszűrt adatok formátuma. Például, ha a konkordanciánkban szeretnénk kiemelni a *pohár* tövű alakokat, akkor azt úgy tehetjük meg, hogy kijelöljük az adatokat tartalmazó oszlopot, majd a **Kezdőlap** fülön a **'Feltételes formázás'**-t választjuk (33. ábra).

33. ábra.
Feltételes
formázás

The screenshot shows the Microsoft Excel interface with the 'Kezdőlap' (Home) tab selected. The 'Feltételes formázás' (Conditional Formatting) button is visible in the 'Stílus' (Style) group. A dropdown menu is open, showing various formatting options. The 'Cellakiemelési szabályok' (Cell Style Rules) option is selected, leading to a list of rules. One rule is highlighted, showing the formula '=poharakból' and the format 'Normal'. The spreadsheet data is visible in the background, with the word 'pohár' appearing in several cells, which are highlighted in green.

	A	B	C	D
1	doc#1	volt, mikor végigjártuk az asztalokat és a	poharakból	egy kancsóba öntögettük a maradékot. Ezt
2	doc#1	borok helyes használatát, még azt is, milyen	poharakat	kell a különböző borfajtákhoz használni
3	doc#3	masszát. - Ezt vedd be - nyújtotta nekem egy	pohár	vízzel együtt -, segít elaludnod, segít
4	doc#3	megmutatkozott: kezének enyhe remegése, mikor a	poharat	szájához emelte, sápadt arca, csaknem mélabús
5	doc#3	szükséges kellékek, továbbá egy palack víz,	pohárral	. A tálcát az asztalra helyezte, majd szó
6	doc#3	mondta, miközben apró kortyokban ürítette ki	pohara	tartalmát. Én meglehetősen hamar
7	doc#3	elmenjek egy lebuja, és elfogyasszak néhány	pohár	italt. Egy ilyen lebujaiban láttam meg Chájimot
8	doc#3	volt a kocsmában: ő szedte össze az üres	poharakat	, ürítette ki a hamutartókat, mosta fel
9	doc#3	is Chájim halálához vezettek. Újabb	pohár	bort töltött magának, majd felállt, odament
10	doc#3	szavait. Ő hátradőlt, miután kiivott egy	pohár	bort. - Ibn Színá halála után szokásom
11	doc#3	Arcáról patakokban folyt a veríték. Újabb	pohár	bort töltött, szivarából mélyet szívott
12	doc#3	ennem. Soha egyetlen falás kenyeret vagy egy	pohár	vizet be nem nyújtottatok. Semmit, semmit
13	doc#3	Úgyhogy leeresztettem a torkán vagy három	pohár	erős pálinkát, s az isteni folyadék el
14	doc#3	belerúgott az ajtó mellett alvó Jenőbe. Éppen a	pohár	után nyúltam, amikor visszafordult, s azt
15	doc#3	olyan szomjas voltam. - Hozzon egy	pohár	vizet - szóltam. A tag kiment, visszajött
16	doc#3	hozott. Hosszú kortyokban ittam. Négy nagy	pohárral	. Mintha a könnyeket akartam volna pótolni
17	doc#3	félénken. - Jól a... francot... hozz egy	pohár	... vizet... - hörögte az öreg. Kezével
18	doc#3	előrehaladott állapotban. Zismalki egy	pohár	vizet nyújtott oda neki, de az öreg nem
19	doc#3	asztalon ételmaradékok, sörösüvegek, kiömlött	poharak	. Többen alszanak, az asztalra borulva,

Itt az 'Új szabály...' választásával adhatunk meg szűrőfeltételeket a felugró 'Új formázási szabály' ablakban (34. ábra). Mivel csak az adott tartalmú cellákat akarjuk formázni, a szabálytípusnál a 'Csak az adott tartalmú cellák formázása' lehetőséget kell választani. Ezután a szabály megadásánál meg kell adnunk, hogy mi alapján szeretnénk formázni. Jelen esetben megadott szöveggel kell egyeznie a cella tartalmának, ez a megadott szöveg pedig a *pohár*. A **Formátum** gombra kattintva beállíthatjuk a kívánt formátumot. Ezután, ha az **OK** gombra kattintunk, a táblázatunk a feltétel szerint formázva lesz, vagyis a *pohár* tövű cellák ki lesznek emelve pirossal (35. ábra).

34. ábra. Formázási szabály megadása

35. ábra. Feltételesen formázott táblázat

sztalokat és a	poharakból	egy kancsóba öntögettük a m
nég azt is, milyen	poharakat	kell a különböző borfajtákhoz
jtotta nekem egy	pohár	vízzel együtt -, segít elaludno
enyhe remegése, mikor a	poharat	szájához emelte, sápadt arca,
á egy palack víz,	pohárral	. A tálcát az asztalra helyezte,
tyokban ürítette ki	pohara	tartalmát. Én meglehetősen f
lfogyasszak néhány	pohár	italt. Egy ilyen lebujsban láttar
össze az üres	poharakat	, ürítette ki a hamutartókat, m
k. Újabb	pohár	bort töltött magának, majd fe
i kiivott egy	pohár	bort. - Ibn Színá halála után sz
veríték. Újabb	pohár	bort töltött, szivarából mélye
s kenyeret vagy egy	pohár	vizet be nem nyújtottatok. Se
kán vagy három	pohár	erős pálinkát, s az isteni folya
lvó Jenőbe. Éppen a	pohár	után nyúltam, amikor visszaf
zon egy	pohár	vizet - szóltam. A tag kiment,
ittam. Négy nagy	pohárral	. Mintha a könnyeket akartam
hozz egy	pohár	... vizet... - hörögte az öreg. Ke
izmalki egy	pohár	vizet nyújtott oda neki, de az
rösüvegek, kiömlött	poharak	. Többen alszanak, az asztalra
Uraim, készítsék	poharaikat	, mert az új évezred első pillar
y üveg sör és egy	pohár	bor. Fejenként egy pakli doh
inak. Khajmor	poharát	kissé figyelmetlenül töltötte
ték a	poharat	, már mindkettőjük szeme csil
itt - veszi fel a	poharát	. - Mindjárt jövök. Hova is mer
rte, majd felemelte	poharát	: - A háború minél hamarabbi

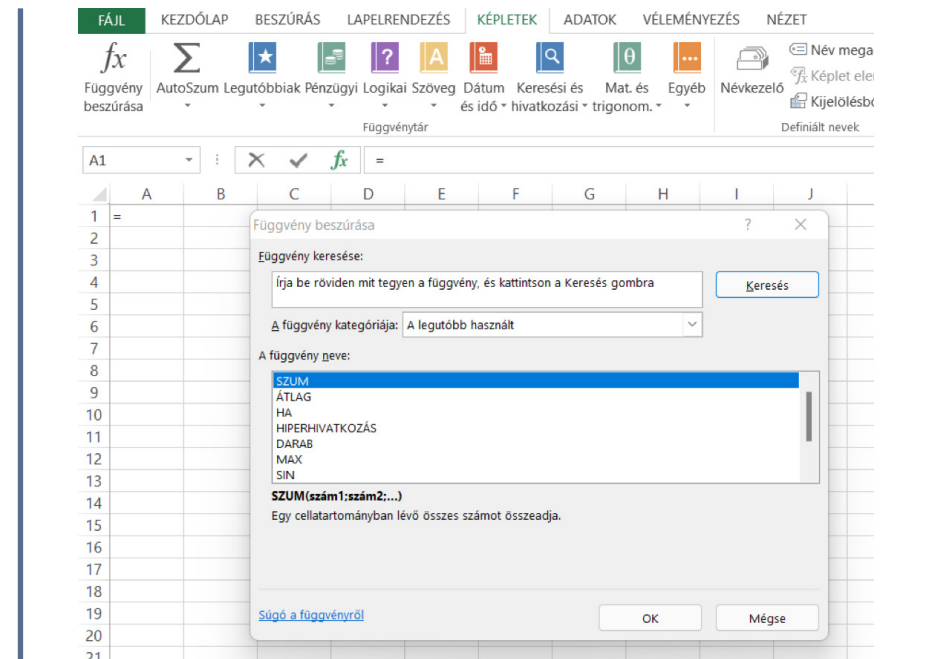
4.3.4. Függvényhasználat

A függvényhasználat talán az egyik leghasznosabb funkciója a táblázatkezelőknek. A függvények segítségével gyorsan és automatikusan számíthatunk statisztikai adatokat, amelyek lényegesen megkönnyíthetik és meggyorsíthatják a munkánkat.

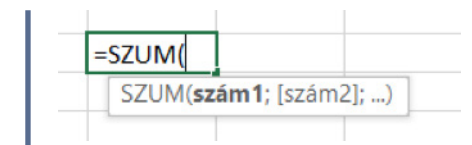
Függvényeket kétféle módon aktiválhatunk. Beszúrhatjuk őket menüből vagy begépelhetjük őket kézzel. A függvénybeszúrást a menüszalag **Képletek** fülén található **'Függvény beszúrása'** menüvel végezhetjük el (36. ábra). Ilyenkor egyszerű dolgunk van, mert az Excel rendelkezésünkre bocsátja a függvények listáját, és ezek között egyszerűen kereshetünk. Kijelölés után rövid leírást is ad a program a függvényről.

Ha már jobban ismerjük a függvényeket, és pontosan tudjuk, melyiket szeretnénk használni, akár be is gépelhetjük azt egy üres cellába. Ilyenkor egy egyenlőségjellel (=) jelezhetjük a cellában, hogy függvény fog következni. Amint elkezdjük begépelni a függvény nevét, az Excel kiegészíti azt, és azt is mutatja segítségül, hogy hány argumentumot kell megadnunk a függvénynek (37. ábra).

Jelen segédanyag csak a kutatások során is használható legfontosabb függvényeket sorolja fel. A függvényeket összefoglaló táblázatok első oszlopa a függvények nevét tartalmazza a függvény argumentumaival együtt, a második oszlopa a függvény funkcióját írja le.



36. ábra. Függvény beszúrása menüből



37. ábra. Függvény kézi bevitele

4.3.4.1. Matematikai függvények

A matematikai függvények matematikai műveletek, számítások elvégzésére szolgálnak. Ide sorolhatjuk még a véletlenszám-generálást (VÉL, VÉL.KÖZÖTT függvény) is.

SZUM(tartomány)	összeadja a számokat
SZUMHA(tartomány, kritérium, összeg_tartomány)	összeadja a számokat, ha megfelelnek egy bizonyos kritériumnak
INT(szám)	lekerekíti a számot a legközelebbi egész számra
KEREK(szám, számjegyek)	a <i>számjegyek</i> paraméterben megadott számú tizedesre kerekíti a számot
KEREK.FEL(szám, számjegyek)	felfele kerekít
KEREK.LE(szám, számjegyek)	lefele kerekít
CSONK(szám, számjegyek)	a <i>szám</i> paraméterben megadott számot a <i>számjegyek</i> paraméterben megadott számú tizedesjegyre jeleníti meg, nem kerekítés, csak elhagyja a felesleges tizedesértékeket
GYÖK(szám)	megadja a szám gyökét
SZORZAT(tartomány)	összeszorozza a számokat
ABS(szám)	megadja egy szám abszolútértékét
VÉL(); VÉL.KÖZÖTT()	véletlenszámot generál

4.3.4.2. Statisztikai függvények

A statisztikai függvények az alapvető statisztikai számítások gyors és pontos elvégzését teszik lehetővé. Gyakori statisztikai számítás az átlag, medián, maximum és minimum kiszámítása. Ezen kívül a megszámláló függvények (DARAB) is ide tartoznak, amelyek az értékek gyakoriságát adják vissza.

ÁTLAG(tartomány)	átlag
MAX(tartomány)	legnagyobb érték
MIN(tartomány)	legkisebb érték
MEDIÁN(tartomány)	kiszámítja az adott halmaz mediánját
MÓDUSZ.EGY(tartomány)	egy tömbből vagy adattartományból kiválasztja a leggyakrabban előforduló számot
DARAB(tartomány)	számot tartalmazó cellák darabszáma
DARAB2(tartomány)	kitöltött (vagyis nem üres) cellák darabszáma
DARABÜRES(tartomány)	üres cellák mennyisége
DARABTELI(tartomány, kritérium)	a kritériumnak megfelelő cellák mennyisége
DARABHATÖBB(kritériumtartomány, kritérium1, ...)	több adott feltétel alapján számolja meg a kijelölt cellákat
KICSI(tartomány,k)	k legkisebb számérték
NAGY(tartomány,k)	k legnagyobb számérték

4.3.4.3. Keresőfüggvények

A keresőfüggvények segítségével megadott feltételek szerint adatokat kereshetünk az adathalmazunkban. Ezek közül a legnehezebb talán az INDEX és HOL.VAN függvények kombinációja, amelyet arra használhatunk, hogy egy bizonyos adathoz hozzá tudjunk rendelni egy tőle balra eső értéket. A példában (38. ábra) a D oszlopban levő azonosítók alapján szeretnénk megtalálni az adatközlők nevét, és azt beírni az F oszlopba. Ezt úgy tehetjük meg, hogy az F2 cellába kattintunk, majd az INDEX és a HOL.VAN függvényeket alkalmazzuk. Az INDEX két koordináta (egy sor- és egy oszlopszám) alapján meg tudja mondani, hogy milyen érték van a koordináták metszéspontjában levő cellában. A HOL.VAN egy sorszámot tud visszaadni, ezt a sorszámot fogja az INDEX függvény felhasználni az egyik paramétereként. A helyes megoldás tehát ez: =INDEX(A\$2:B\$24, HOL.VAN(D2,B\$2:B\$24,0),1).

Az INDEX függvény tartomány része az A2:B24-es tartomány, amelyet \$ jelekkel tudunk rögzíteni, hiszen mindig ebből a tartományból kell, hogy keressen a függvény, így a függvény másolásakor nem csúszik el a tartomány. Ugyanez vonatkozik a HOL.VAN keresési tartományára.

FKERES(keresett_érték, tartomány, oszlop, közelítés)	csak balról jobbra keres
VKERES(keresett_érték; tartomány; sor, közelítés)	csak fentről lefele keres
INDEX(tartomány, sor, oszlop)	egy sor- és oszlopszám metszéspontjában álló értéket adja vissza
HOL.VAN(keresett_érték, keresési_tartomány, közelítési_mód)	megadja egy érték pozícióját
KERES(keresési_érték, keresési_vektor, eredmény_tartomány)	minden irányban keres

	A	B	C	D	E	F
1	Név	Azonosító		Adatközlő	Adat	Név
2	Gipsz Jakab	mugwW		NRZxZ	olyan szomjas voltam. - Hozzon egy	
3	Heu Réka	hJyE1		zYd9j	hozott. Hosszú kortyokban ittam. Négy nagy	
4	Meg Győző	mwyXU		DyF8W	félénken. - Jól a... francot... hozz egy	
5	Major Anna	Dmr89		NVCU0	előrehaladott állapotban. Zizmalki egy	
6	Riz Ottó	1VWzP		9nIWR	asztalon ételmaradékok, sörösüvegek, kiömlött	
7	Nemer Eszti	NRZxZ		AkIVf	szavak jöttek: Hölgyeim és Uraim, készítsék	
8	Ne Pál	zYd9j		fpi42	kevés tészta, utána meg egy üveg sör és egy	
9	Ka Pál	DyF8W		yE6hP	legyen. Töltött, előbb magának. Khajmor	
10	Elektrom Ágnes	NVCU0		P0OEI	majd ő is ivott. Mikor letették a	
11	Hú Zóra	9nIWR		Y2V8N	jobban Hans szemei. - Várj itt - veszi fel a	
12	Raj Zóra	AkIVf		WAjQr	lányok sörét vodkával keverte, majd felemelte	
13	Dia Dóra	fpi42		WgdSF	ital, gondolta, miközben szájához emelte a	
14	Tüdő R. Ákos	yE6hP		hJyE1	kérsz, akkor sokat is várok el - emelte fel	
15	Ria Dóra	P0OEI		mwyXU	Legalábbis én ezt hittem... - emelte meg	
16	Lev Elek	Y2V8N		Dmr89	hallgatás következett. Khajmor maga elé nézett,	
17	Hot Elek	WAjQr		mugwW	olyan, mint meginni apró cseppekben egy	

SZUM				=INDEX(A\$2:B\$24, HOL.VAN(D2,B\$2:B\$24,0),1)	
	A	B	C	D	F
1	Név	Azonosító		Adatközlő	Adat
2	Gipsz Jakab	mugwW		NRZxZ	olyan szomjas voltam. - Hozzon egy
3	Heu Réka	hJyE1		zYd9j	hozott. Hosszú kortyokban ittam. Négy nagy
4	Meg Győző	mwyXU		DyF8W	félénken. - Jól a... francot... hozz egy
5	Major Anna	Dmr89		NVCU0	előrehaladott állapotban. Zizmalki egy
6	Riz Ottó	1VWzP		9nIWR	asztalon ételmaradékok, sörösüvegek, kiömlött
7	Nemer Eszti	NRZxZ		AkIVf	szavak jöttek: Hölgyeim és Uraim, készítsék
8	Ne Pál	zYd9j		fpi42	kevés tészta, utána meg egy üveg sör és egy
9	Ka Pál	DyF8W		yE6hP	legyen. Töltött, előbb magának. Khajmor
10	Elektrom Ágnes	NVCU0		P0OEI	majd ő is ivott. Mikor letették a
11	Hú Zóra	9nIWR		Y2V8N	jobban Hans szemei. - Várj itt - veszi fel a
12	Raj Zóra	AkIVf		WAjQr	lányok sörét vodkával keverte, majd felemelte
13	Dia Dóra	fpi42		WgdSF	ital, gondolta, miközben szájához emelte a
14	Tüdő R. Ákos	yE6hP		hJyE1	kérsz, akkor sokat is várok el - emelte fel
15	Ria Dóra	P0OEI		mwyXU	Legalábbis én ezt hittem... - emelte meg
16	Lev Elek	Y2V8N		Dmr89	hallgatás következett. Khajmor maga elé nézett,
17	Hot Elek	WAjQr		mugwW	olyan, mint meginni apró cseppekben egy

38. ábra. Példa az INDEX + HOL.VAN függvények alkalmazására

4.3.4.4. Szövegfüggvények

A szövegfüggvények lehetővé teszik, hogy szöveges adatainkat manipuláljuk. Például létrehozhatunk egyedi azonosítókat adatközlőink vezetéknév- és keresztnévének plusz egy random sorszámnak a kombinációjából. Ezt mutatja be a 39. ábra. Az azonosítót úgy nyerjük, hogy egy 25 és 50 közötti véletlen számot összefűzünk a vezetéknév első két karakterével, valamint a keresztnév első két karakterével.

ÖSSZEFŰZ(szöveg1, szöveg2)	összefűz két szöveget
BAL(szöveg, karatkerszám)	a szöveg első <i>k</i> darab karakterét adja eredményül
JOBB(szöveg, karakterszám)	a szöveg utolsó <i>k</i> darab karakterét adja eredményül
KÖZÉP(szöveg, kezdet, karakterszám)	a <i>kezdet</i> paraméterben megadott sorszámú karaktertől számítva annyi karaktert ad vissza, amennyit a <i>karakterszám</i> paraméterben meghatározunk
HOSSZ(szöveg)	megadja a szöveg hosszát
SZÖVEG.KERES(keresett_szöveg; szöveg; kezdet)	azt a karaktersorszámot adja meg, ahol egy adott karakter vagy karaktersorozat kezdődik balról jobbra haladva, nem különböztet meg kis- és nagybetűket

39. ábra. Példa a szövegfüggvények használatára

C25 ✕ ✓ <i>f</i> =ÖSSZEFŰZ(VÉLETLEN.KÖZÖTT(25,50),BAL(A25,2),BAL(B25,2))			
	A	B	C
22			
23			
24	Vezetéknév	Keresztnév	Azonosító
25	Gipsz	Jakab	41GiJa
26	Kiss	Imre	32KiIm
27	Nagy	Mária	34NaMá
28	Rabsz	Olga	46RaOl
29	Trap	Pista	43TrPi
30	Csin	Csilla	33CsCs
31	Riz	Ottó	34RiOt
32	Negye	Dóra	47NeDó

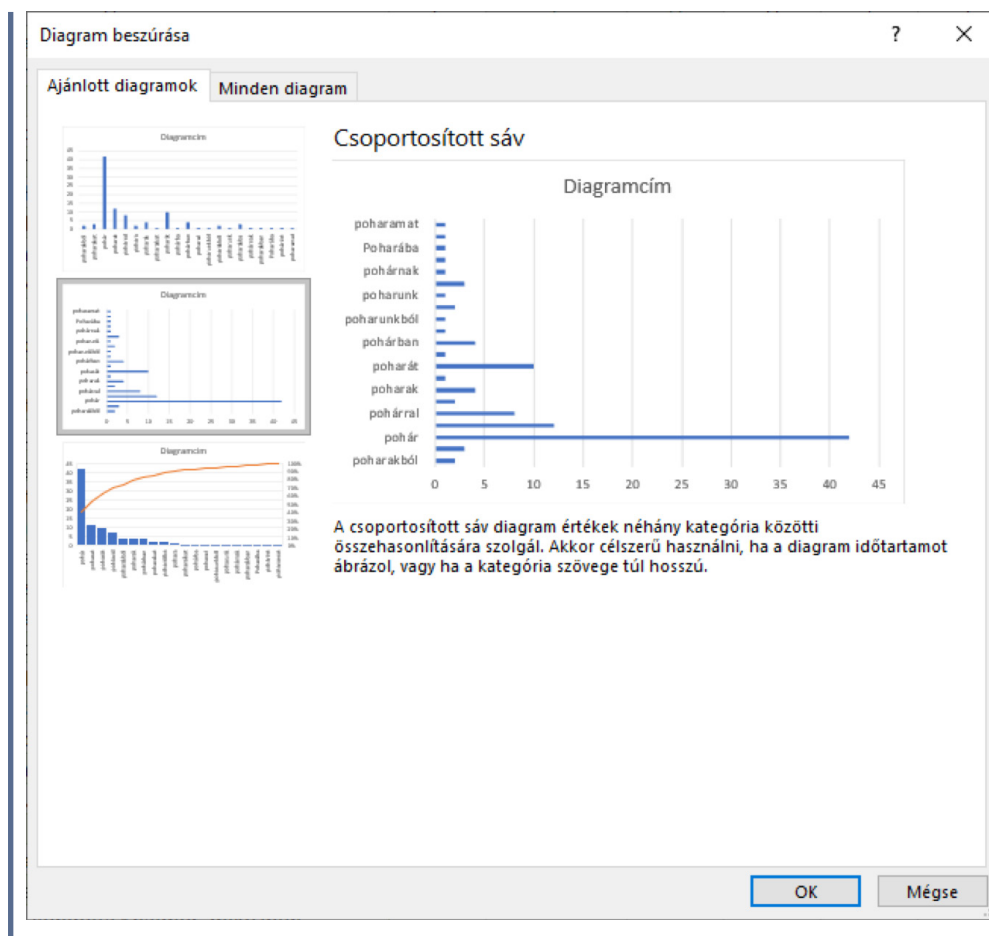
4.3.5. Diagramkészítés

Az Excel táblázatkezelőben egyszerűen készíthetünk diagramokat az adatainkból és számításainkból. Például kiszámolhatjuk a DARABTELI függvény segítségével, hogy a **32. ábra** szűrési eredményeinél az egyes adatok közül mennyi található meg a teljes halmazban (pl. hány *poharat* előfordulás van). Ezután pedig diagramot készíthetünk ezekről a gyakoriságokról. A diagram készítéséhez először ki kell jelölnünk azt a tartományt, amely adatállományként fog szolgálni a diagramhoz. A **Beszűrés** fül **Diagramok** részén találjuk az egyes diagramtípusokat, illetve, ha a jobb alsó sarokban lévő nyilat „lenyitjuk”, további paramétereket adhatunk meg (40. ábra).

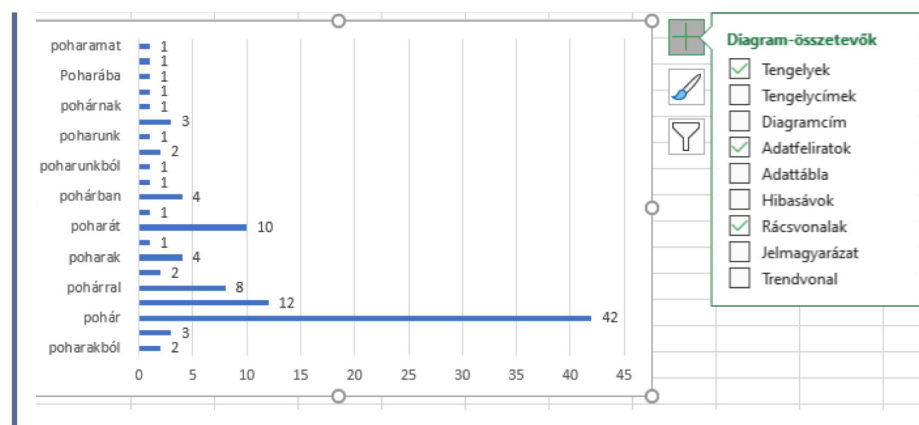
40. ábra. Diagram beszurása

B	C	D	E	F	G	H
or végigjártuk az asztalokat és a	poharakból	egy kancsóba öntögettük a maradékot. Ezt		poharakból	2	
lyes használatát, még azt is, milyen	poharakat	kell a különböző borfajtákhoz használni		poharakat	3	
- Ezt vedd be - nyújtotta nekem egy	pohár	vízzel együtt -, segít elaludnod, segít		pohár	42	
tkozott: kezének enyhe remegése, mikor a	poharat	szájához emelte, sápadt arca, csaknem mélabús		poharat	12	
s kellékek, továbbá egy palack víz,	pohárral	. A tálcát az asztalra helyezte, majd szó		pohárral	8	
miközben apró kortyokban ürítette ki	pohara	tartalmát. Én meglehetősen hamar		pohara	2	
egy lebuja, és elfogyasszak néhány	pohár	italt. Egy ilyen lebujaiban láttam meg Chájimot		poharak	4	
smában: ő szedte össze az üres	poharakat	, ürítette ki a hamutartókat, mosta fel		poharaikat	1	
halálához vezettek. Újabb	pohár	bort töltött magának, majd felállt, odament		poharat	10	
ő hátradől, miután kiivott egy	pohár	bort. - Ibn Színá halála után szokásommá		pohárba	1	
atkokban folyt a veríték. Újabb	pohár	bort töltött, szivarából mélyet szívott		pohárban	4	
ioha egyetlen falás kenyeret vagy egy	pohár	vizet be nem nyújtottatok. Semmit, semmit		poharad	1	
leeröltettem a torkán vagy három	pohár	erős pálinkát, s az isteni folyadék el		poharunkból	1	
tt az ajtó mellett alvó Jenőbe. Éppen a	pohár	után nyúltam, amikor visszafordult, s azt		poharakból	2	
imjas voltam. - Hozzon egy	pohár	vizet - szólám. A tag kiment, visszajött		poharunk	1	
összú kortyokban ittam. Négy nagy	pohárral	. Mintha a könnyeket akartam volna pótolni		poharakba	3	
. - Jól a... francot... hozz egy	pohár	... vizet... - hörögte az öreg. Kezével		pohárnak	1	
adott állapotban. Zismalki egy	pohár	vizet nyújtott oda neki, de az öreg nem		poharakban	1	
ételmaradékok, sörösüvegek, kiömlött	poharak	. Többen alszanak, az asztalra borulva,		Poharába	1	
ttek: Hölgyeim és Uraim, készítsék	poharaikat	, mert az új évezred első pillanata mindjárt		pohárért	1	
izta, utána meg egy üveg sör és egy	pohár	bor. Fejenként egy pakli dohány. Khajmor		poharamat	1	
öltött, előbb magának. Khajmor	poharat	kissé figyelmen kívül hagyta, melléje				
ivott. Mikor letették a	poharat	, már mindkettőjük szeme csillogott.				

A nyíl lenyitásával egy „varázslót” kapunk, amely segít kiválasztani a megfelelő diagramtípust (41. ábra). A diagram elkészültével különféle formázásokat végezhetünk (42. ábra).



41. ábra. DiagramVarázsló



42. ábra. A kész diagram testre szabása

4.4. Adatvizualizáció

Eredményeink közlésekor kiemelkedően fontosak a megértést segítő ábrák, az adatok vizuális ábrázolása. Ugyan a táblázatkezelőkben is tudunk egyszerűbb diagramokat készíteni, a lehetőségek nagyon limitáltak. Ebben a fejezetben olyan, többnyire webes eszközöket mutatok be, amelyek segítségével közleményeinkben, előadásainkban átláthatóbbá és rendszerezettebbé tehetjük az adatainkat. Az adatok diagramokon, gráfokon történő ábrázolása után a színesebb, inkább prezentációkhoz és posztterekhez használható info-grafikák és idővonalak szerkesztésére láthatunk lehetőségeket.

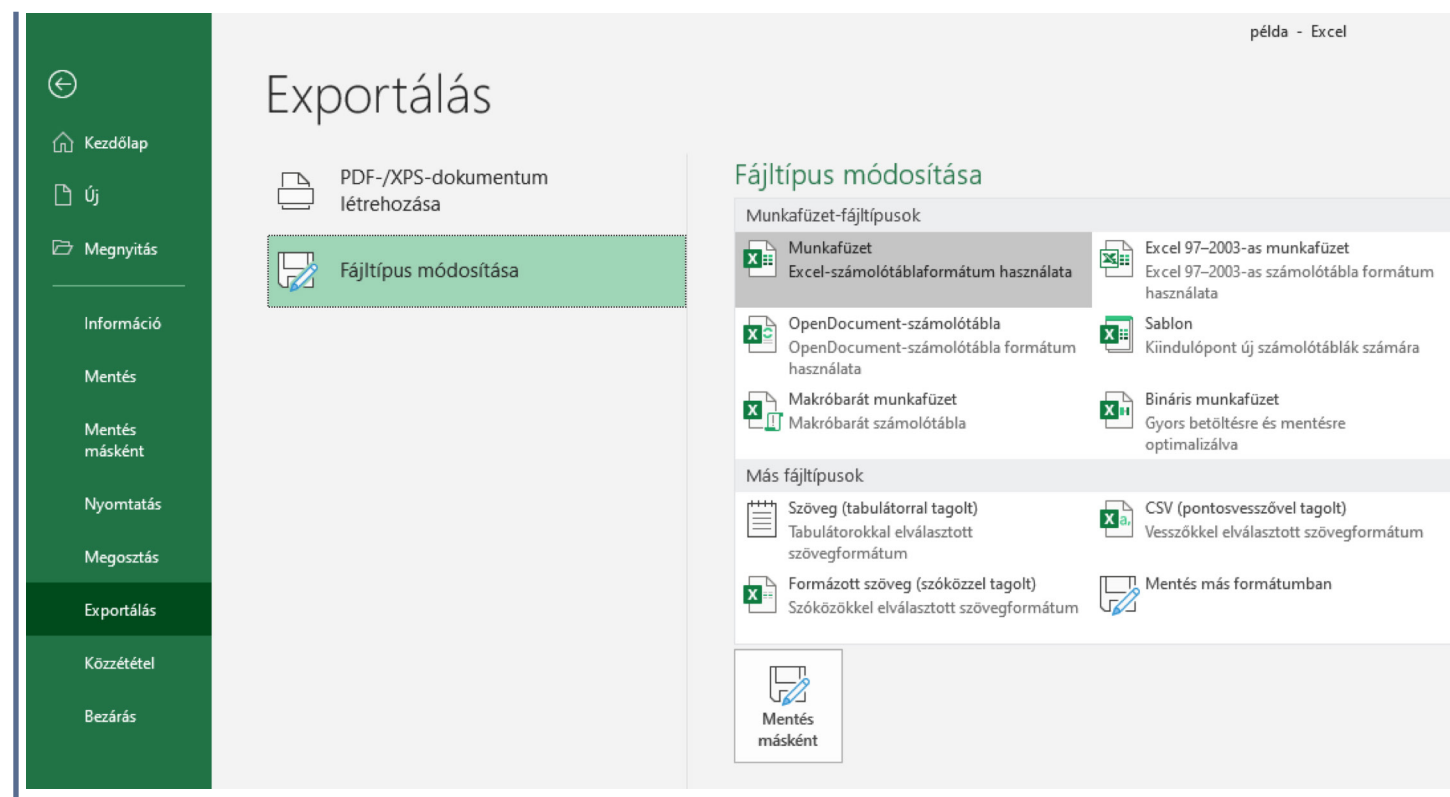
4.4.1. Diagramok, gráfok

A diagramok és gráfok kvantitatív adatok vizuális megjelenítésére szolgálnak. Az Excelben látott diagramokhoz hasonlóakat állíthatunk elő az alábbi webes eszközökkel. Ezeknek előnye, hogy többféle ábrázolási lehetőséget tartalmaznak és a már meglevő Excel-fájljainkat használhatjuk létrehozásukhoz.

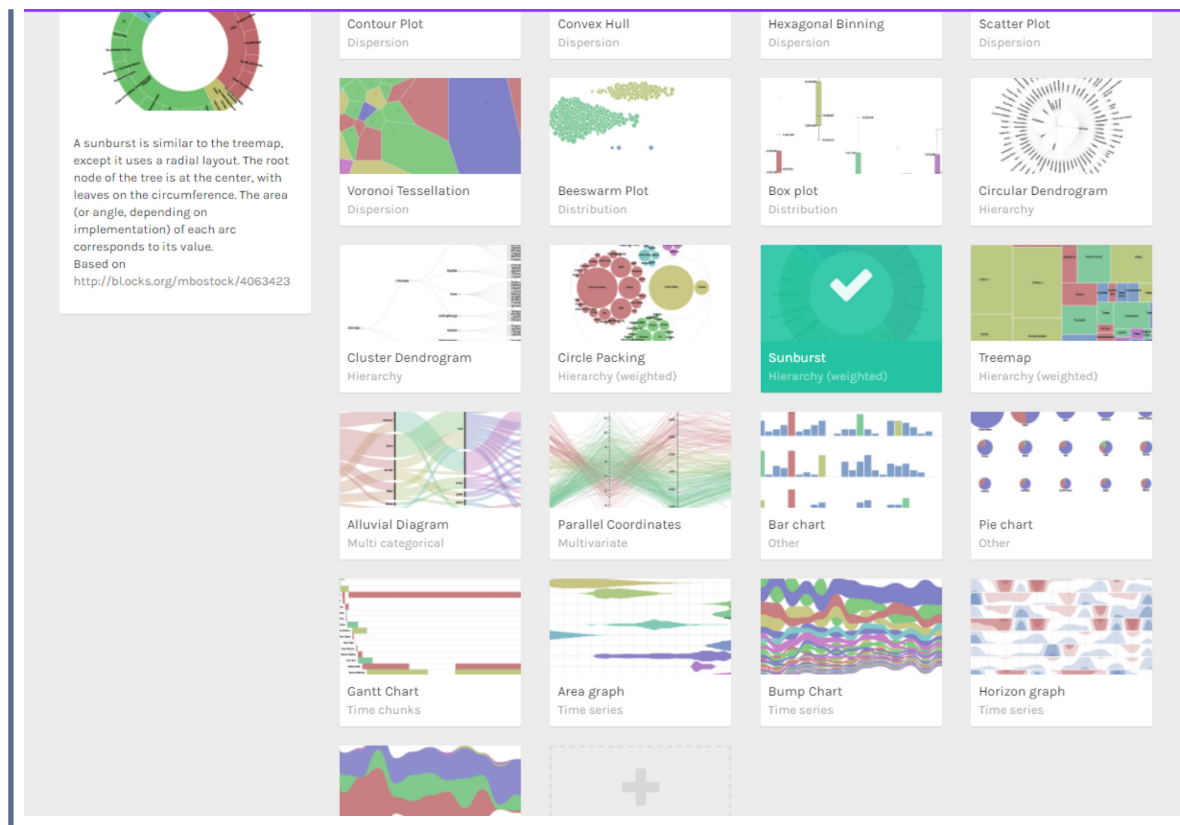
4.4.1.1. RAW

A RAW nagyon egyszerűen, regisztráció nélkül használható adatvizualizáló alkalmazás, amely a <https://app.rawgraphs.io/> oldalon érhető el. Az adatainkat .csv vagy .tsv formátumban célszerű a szövegmezőbe beilleszteni. Excelből tudunk ezekbe a formátumokba exportálni a **Fájl > Exportálás > Fájl típus** módosítása lehetőségeket választva (43. ábra).

A szöveg bemásolása után válogathatunk különféle diagramok közül (44. ábra), így akár a 45. ábrához hasonló diagramokat is létrehozhatunk. Ezek a diagramok elmenthetők képként és beilleszthetők weboldalakba is.

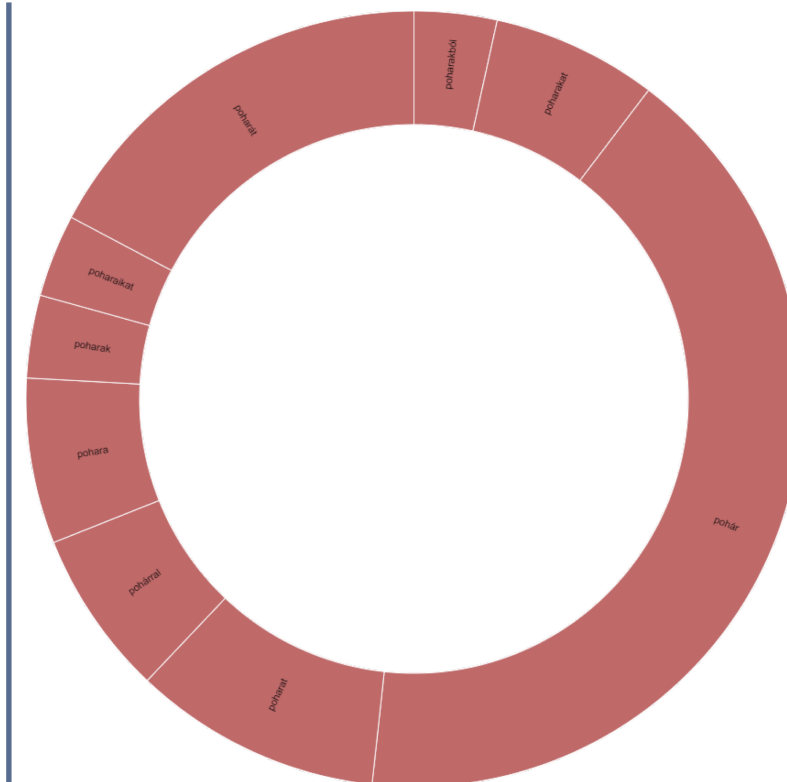


43. ábra. Exportálás Excelben más fájltípusokba



44. ábra. A RAW diagramválasztója

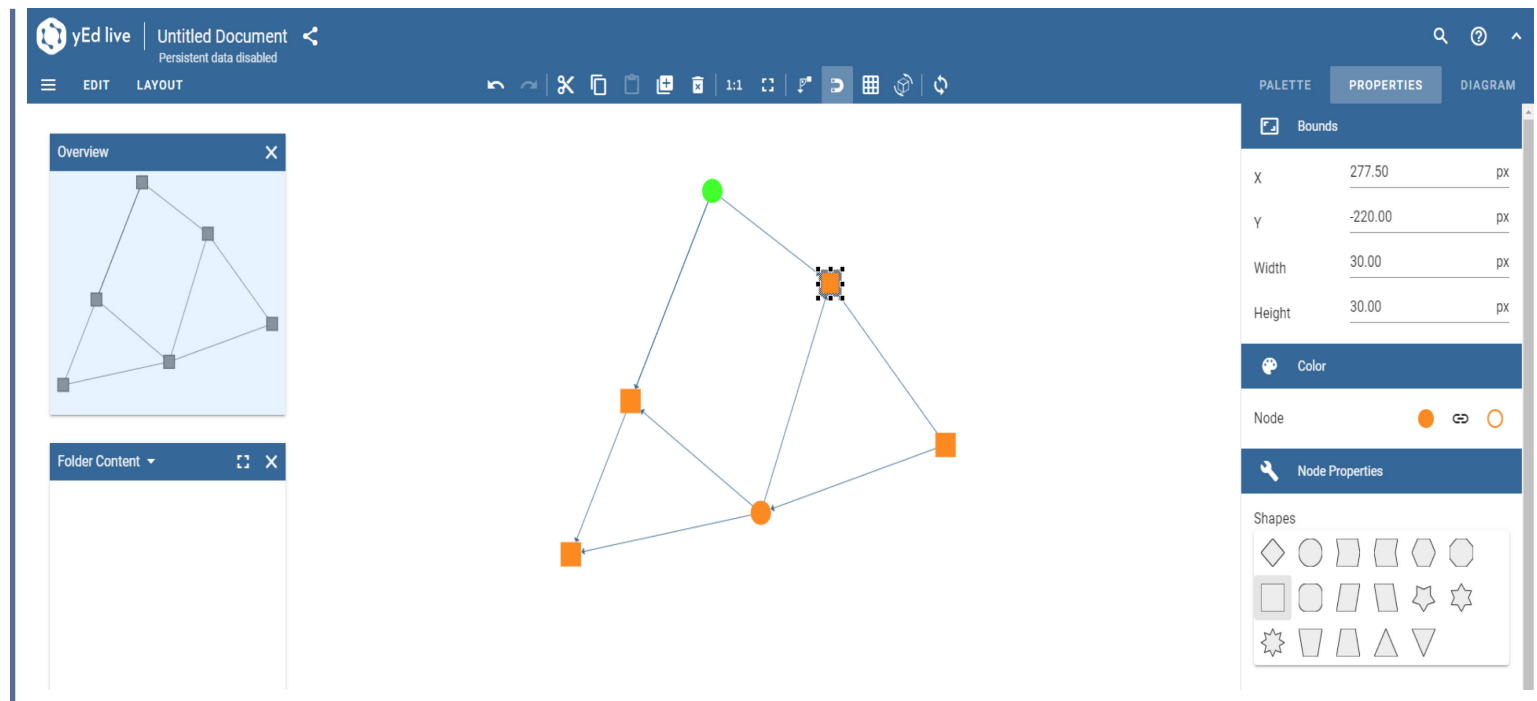
45. ábra. RAW-ban készített diagram



4.4.1.2. yED

A yED nem kifejezetten diagramok készítésére szolgál, inkább gráfokat, folyamatábrákat készíthetünk vele. Az alkalmazásnak létezik asztali és webes változata is, itt részletesebben a *yED Live* nevet viselő webes alkalmazás kerül bemutatásra, amely a <https://www.yworks.com/yed-live/> oldalon érhető el. Ingyenesen és regisztráció nélkül használható.

A *yED live* segítségével különböző formájú és színű nódusokat helyezhetünk el a szerkesztőfelületen, majd ezeket egyszerűen összeköthetjük (46. ábra). A kutatások során számos alkalommal van szükségünk hasonló gráfok elkészítésére. Az eszközzel akár mondattani ágrajz, akár a kutatás menetét bemutató folyamatábra, akár nagyobb terjedelmű publikáció tartalmának hálózatát bemutató ábra, infografika is készíthető.

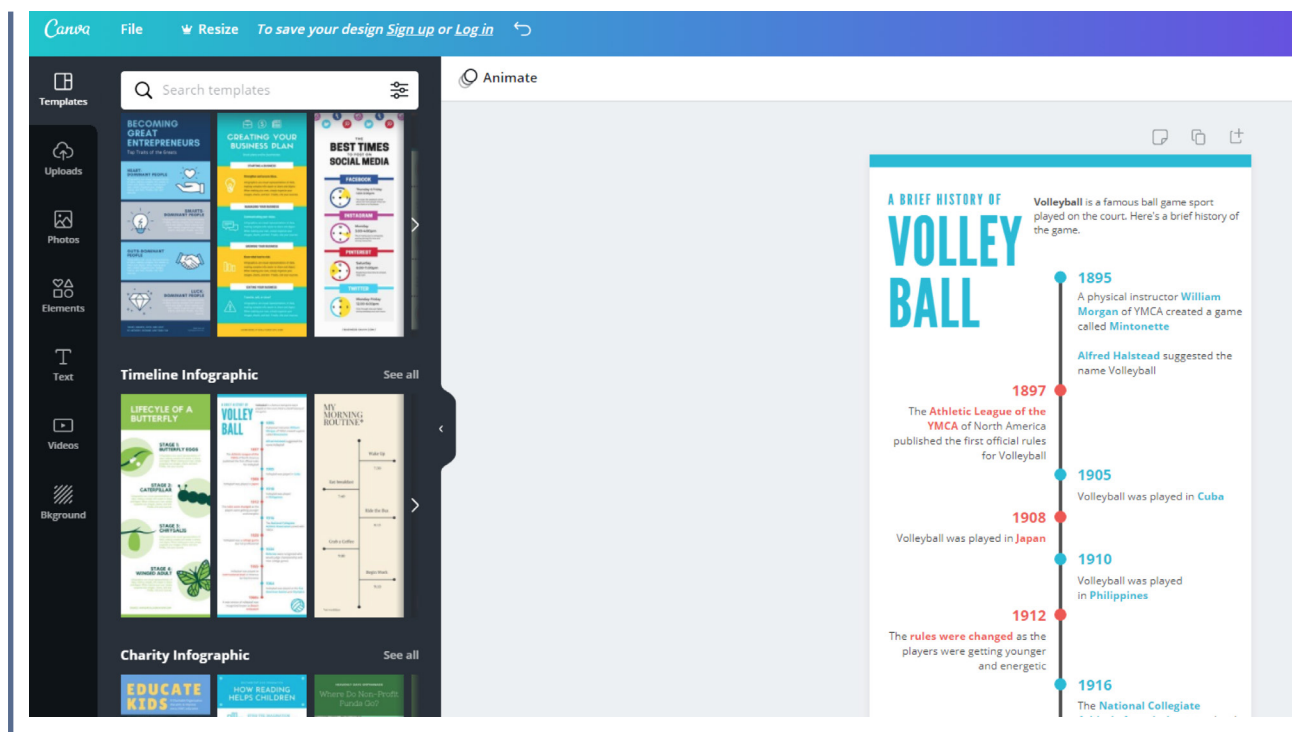


46. ábra. A yED Live felülete

4.4.2. Infografikák, idővonalak

Az infografikák és idővonalak azt a célt szolgálják, hogy színesen, vizuálisan mutassanak be kvantitatív és kvalitatív adatokat, az azokból kinyert összefüggéseket. Ez a fejezet csupán egy alkalmazást ismertet, amely jó kiindulási alap lehet infografikák készítéséhez. Ezen kívül a **Melléklet**ben található Programgyűjteményben felsorolok további lehetőségeket.

Az egyik leguniverzálisabb vizualizációs alkalmazás a *Canva*. Segítségével nemcsak diagramokat, infografikákat hozhatunk létre, hanem egyéb ábrákat, könyvborítót, prezentációt is. Az alkalmazást a <https://www.canva.com/> oldalon találjuk, regisztráció nélkül is használható, azonban a regisztrált verzióban elmenthetjük és később is szerkeszthetjük a munkánkat. A regisztráció és a használat ingyenes. Az oldalon választhatunk sablonokból, vagy készíthetünk saját alakzatokat a grafikai segédeszközök segítségével.



47. ábra. Idővonal-készítés
Canvá-val

5. Kutatás és eszközválasztás

Ahogy minden módszernek megvan a maga helye a kutatásban, úgy az egyes módszerekhez felhasználható eszközöknek is. Az előző fejezetben bemutatott eszközök, eljárások több különböző módszer használatakor, illetve több különböző kutatási szituációban alkalmazhatók. Nem szorítkozhatunk csupán egy eszköz használatára a kutatás során, ahogyan nem szorítkozunk egyetlen módszer alkalmazására sem, hanem több módszert együtt vagy összeépítve alkalmazunk. Ennek megfelelően az eszközök adekvát kiválasztásának több szempontja lehet:

- ¶ Fontos, hogy az eszköz alkalmas legyen annak a feladatnak az ellátására, amire kiválasztottuk.
- ¶ Az eszköz használata csökkentse a vizsgálat idejét, vagy ha növeli az időtartamot, akkor a kutatás minőségén javítson.
- ¶ Ne akarjunk mindenképpen minden helyzetben eszközt alkalmazni. Mérlegeljük, hogy valóban szükség van-e informatikai eljárás vagy eszköz bevonására.

Amennyiben kiválasztottunk egy eszközt, érdemes az alábbi kérdéseket feltenni, és mérlegelni, hogy valóban ez az eszköz-e a legalkalmasabb a kutatási feladatra:

- ¶ Valóban ez az eszköz a legmegfelelőbb? – Ilyenkor érdemes tájékozódni más hasonló eszközökről, esetleges hatékonyabb eljárásokról, hogy megbizonyosodjunk róla, melyik a legalkalmasabb.
- ¶ Az eszköz valóban azokat az információkat, eredményeket mutatja meg, amelyekre szükség van a kutatás során?

¶ Az eszköz hogyan és mennyire torzítja az eredményeket? A torzítás szignifikáns-e? Megengedhető-e ez a fajta torzítás a kutatás céljainak szempontjából? (Lásd a **we** szó keresési eredményeiről szóló kutatást a **3.1.1. fejezetben**)

Ez a fejezet annak alapján rendszerezi az eszközöket, hogy miként lehetnek segítségünkre a kutatás egyes fázisaiban, illetve hogyan járulnak hozzá az egyes kutatási módszerek sikerességéhez.

Webes keresést alkalmazhatunk a kutatás korai szakaszában a szakirodalom és egyéb hasznos információk felkutatására. Ehhez a Google, a Google Scholar és a *Duck Duck Go* felületét is használhatjuk. Az adatgyűjtési fázisban is hasznos lehet, ha a keresési feltételeinket jól tudjuk formába önteni, könnyebben tudunk bizonyos nyelvi jelenségekről információt gyűjteni az interneten. Fontos azonban kiemelni, hogy az internetes keresés nem mindig adekvát módja a nyelvi jelenségekről való információszerzésnek (lásd: **3.1.1. fejezet**). Ilyen módszerrel nem nyelvi adatokhoz juthatunk, csupán csak információt kaphatunk bizonyos trendekről.

Reguláris kifejezéseket alkalmazhatunk az adatgyűjtés során korpuszadatok lekérdezésére. Sokkal megbízhatóbb konkordanciákat kaphatunk, ha egy jólformált reguláris kifejezéssel keresünk, mintha csupán kulcsszavas keresést végzünk. Ezen kívül hasznát vehetjük a reguláris kifejezéseknek az adatok elemzésénél. Az adatokban mintázatokat kereshetünk és mutathatunk ki reguláris kifejezések segítségével vagy esetleg részekre bonthatjuk az adatokat írásjelek és szóközök mentén.

Az **online kérdőívek** a kérdőíves kutatások adatgyűjtési fázisában játszanak jelentős szerepet. Segítségükkel egyszerűen kitölthető kérdőíveket készíthetünk, amelyeket könnyen eljuttathatunk az adatközlőkhöz. Ezen felül az adatelemzést is segítik, a válaszok alapján kimutatásokat, diagramokat készítenek. A válaszokat letöltve magunk is végezhetünk számításokat az adatokon, és vizualizálhatjuk az eredményeket.

Táblázatkezelőt szinte minden itt felsorolt kutatási módszernél alkalmazhatunk az adatok elemzésének szakaszában. Mind korpuszadatokat, mind kérdőívekből és interjúkból származó adatokat hatékonyan kezelhetünk vele. A korpuszadatok feldolgozására láthatunk példát a **4.3. fejezet**ben. A táblázatkezelő elsősorban statisztikai számítások elvégzésére szolgál, egyszerre nagyszámú adatot tudunk vele feldolgozni, a szűrők segítségével szűrhetjük az adatokat bizonyos szempontok alapján, illetve kiemelhetjük a számunkra relevánsakat. Az eredmények szemléltetésére diagramokat hozhatunk létre a segítségével.

A táblázatkezelő diagramjain túl egyéb webes **adatvizualizáló eszközök**et is kipróbálhatunk. Ezek lehetővé teszik, hogy dinamikusan és kreatívan rendezzük el az adatainkat, így nem kell csupán a táblázatkezelő nyújtotta lehetőségekre hagyatkoznunk. A webes alkalmazásokkal készített ábrák általában letölthetők, megoszthatók, weboldalba illeszthetők.

6. Összegzés, kitekintés

Jelen kutatásmódszertani segédanyaggal az volt a célom, hogy olyan informatikai eszközöket és megoldásokat mutassak be nyelvészek és nyelvészhallgatók számára, amelyekből inspirálódhatnak, ötleteket meríthetnek a kutatásaikhoz. A segédanyag első fele a funkcionális kognitív nyelvészet adatokhoz és kutatási módszerekhez fűződő viszonyát mutatja be, részletezi a leggyakrabban alkalmazott adatgyűjtési és adatelemzési módszerek tulajdonságait, esetleges hiányosságait, buktatóit. Ezt követően olyan informatikai eszközöket sorakoztat fel, amelyek nagymértékben segíthetik az egyes kutatási módszerek kivitelezését. A felsorolt eszközök nem kifejezetten tudományos kutatáshoz készültek, inkább könnyen és sokoldalúan használható mindennapi alkalmazások. Célzott felhasználásukat a kutatások során is hasznosítható példákon keresztül mutatja be a segédanyag. A módszerek és az eszközök áttekintése után az eddigieket összekapcsoló rész következik, amely összefoglalja, hogy az egyes eszközök miként segíthetik a nyelvészek munkáját a tudományos kutatás egyes fázisaiban, illetve hogyan járulnak hozzá az egyes vizsgálati módszerek sikerességéhez.

Bár a segédanyag összefüggő egész mű, mégsem teljes körű, és nem is törekedett a teljességre. Célja az volt, hogy kezdeményezze a funkcionális kognitív nyelvészeti kutatási módszerek és a kutatásban felhasználható informatikai eszközök összegyűjtését. Végző célom az lenne, hogy ez a gyűjtemény közösségi legyen, bárki hozzáteheszen ötleteket, példákat, új eszközöket vagy a régi eszközök felhasználásának új lehetőségeit.

Az eszközyűjtemény kiszélesítése mellett a jövőben hasznos lehet a segédanyag videós anyagokkal való kiegészítése, illetve képzések, workshopok szervezése, amelyeken mind hallgatók, mind oktatók feltehetik kérdéseiket, megoszthatják tapasztalataikat a kutatási módszerekkel és a kutatás informatikai eszközeivel kapcsolatban. Ugyan a mű a funkcionális kognitív nyelvészek igényeit célozta meg, a legtöbb módszer és eszköz más elméleti keretben dolgozó empirikus kutatásokat folytató nyelvészek számára is hasznos lehet.

7. Bibliográfia

- ANTHONY, Laurence (2013): A critical look at software tools in corpus linguistics. *Linguistic Research* 30 (2). 141–161. DOI: 10.17250/khisli.30.2.201308.001
- BAKER, Paul (2010): Corpus Methods in Linguistics. In: LITOSSELITI, Lia (ed.): *Research Methods in Linguistics*. Continuum, London/New York. 93–113.
- BIBER, Douglas (1993): Representativeness in corpus design. *Literary and Linguistic Computing* 8 (4). 243–257.
- BORONKAI Dóra (2009): *Bevezetés a társalgáselemzésbe*. Ad Librum Kft., Budapest.
- CHOMSKY, Noam (1965): *Aspects of the Theory of Syntax*. The MIT Press, Massachusetts.
- CHOMSKY, Noam (1968): *Language and Mind*. Harcourt, Brace & World, New York.
- CSENDES Dóra – HATVANI Csaba – ALEXIN Zoltán, CSIRIK János – GYIMÓTHY Tibor – PRÓSZÉKY Gábor; VÁRADI Tamás (eds.) (2003): Kézzel annotált magyar nyelvi korpusz: a Szeged Korpusz. *I. Magyar Számítógépes Nyelvészeti Konferencia előadásai*: MSZNY 2003. 238–245.
- EDDINGTON, David (2015): *Statistics for Linguists: A Step-by-Step Guide for Novices*. Cambridge Scholars Publishing, Lady Stephenson Library, Newcastle upon Tyne.
- FAUCONNIER, Gilles – TURNER, Mark (2002): *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. Basic Books, New York.
- FÓRIS Ágota (2008): *Kutatásról nyelvészeknek. Bevezetés a tudományos kutatás módszertanába*. Nemzeti Tankönyvkiadó, Budapest.
- GEERAERTS, Dirk – CUYCKENS, Hubert (2007): Introducing Cognitive Linguistics. In: Geeraerts, Dirk – Cuyckens, Hubert (eds.) (2007): *The Oxford Handbook of Cognitive Linguistics*. Oxford University Press, Oxford.

- GIBBS, Raymond W. Jr (2006): Introspection and cognitive linguistics. Should we trust our own intuitions? *Annual Review of Cognitive Linguistics* 4 (1). 135–151.
- GIBBS, Raymond W. Jr. – O'BRIEN, Jennifer E. (1990): Idioms and mental imagery: The metaphorical motivation for idiomatic meaning. *Cognition* 36 (1). 35–68.
- GONZALEZ-MARQUEZ, Monica – MITTELBERG, Irene – COULSON, Seana – SPIVEY, Michael J. (eds.) (2007): *Methods in Cognitive Linguistics*. John Benjamins, Amsterdam/Philadelphia.
- GRONDELAERS, Stefan – GEERAERTS, Dirk – SPEELMAN, Dirk (2007): A case for a cognitive corpus linguistics. In: GONZALEZ-MARQUEZ, Monica – MITTELBERG, Irene – COULSON, Seana – SPIVEY, Michael J. (eds.) (2007): *Methods in Cognitive Linguistics*. John Benjamins, Amsterdam/Philadelphia. 149–169.
- HALÁCSY Péter – KORNAI András – NÉMETH László; SASS Bálint – VARGA Dániel – VÁRADI Tamás – VONYÓ Attila (eds.) (2005): A Hunglish korpusz és szótár. *III. Magyar Számítógépes Nyelvészeti Konferencia, Szeged*.
- KARTTUNEN, Lauri – CHANOD, Jean-Pierre – GREFENSTETTE, Gregory – SCHILLER, Anne (1997): Regular Expressions for Language Engineering. *Natural Language Engineering* 2 (4). 1–24.
- KEMMER, Suzanne – BARLOW, Michael (2000): Introduction: A usage-based conception of language. In: BARLOW, Michael – KEMMER, Suzanne (eds.): *Usage-Based Models of Language*. CSLI Publications, Stanford, California. vii–xxviii.
- KENNEDY, Graeme D. (1998): *An Introduction to Corpus Linguistics*. Longman, London/New York.
- KVALE, Steinar – BRINKMANN, Svend (2009): *InterViews: Learning the Craft of Qualitative Research Interviewing*. 2nd ed. Sage Publication, Los Angeles.
- LAKOFF, George – JOHNSON, Mark (1980): *Metaphors We Live By*. The University of Chicago Press, Chicago/London.
- LANGACKER, Ronald W. (1987): *Foundations of Cognitive Grammar. Volume I. Theoretical Prerequisites*. Stanford University Press, Stanford, California.

- LANGACKER, Ronald W. (1999): Assessing the cognitive linguistic enterprise. In: JANSSEN, Theo – REDEKER, Gisela (eds.): *Cognitive Linguistics: Foundations, Scope and Methodology*. Mouton de Gruyter, Berlin/New York. 13–60.
- LANGACKER, Ronald W. (2008): *Cognitive Grammar. A Basic Introduction*. Oxford University Press, Oxford.
- LEVON, Erez (2010): Organizing and Processing Your Data: The Nuts and Bolts of Quantitative Analyses. In: LITOSSELITI, Lia (ed.): *Research Methods in Linguistics*. Continuum, London/New York.
- LEWICKI, Pawel – HILL, Thomas – BIZOT, Elizabeth (1988): Acquisition of procedural knowledge about a pattern of stimuli that cannot be articulated. *Cognitive Psychology* 20 (1). 24–37.
- MCENERY, Tony – WILSON, Andrew (2001): *Corpus Linguistics. 2nd edition. Edinburgh Textbooks in Empirical Linguistics*. Edinburgh University Press, Edinburgh.
- MITTELBERG, Irene – FARMER, Thomas A. – WAUGH, Linda R. (2007): They actually said that? An introduction to working with usage data through discourse and corpus analysis. In: GONZALEZ-MARQUEZ, Monica – MITTELBERG, Irene – COULSON, Seana – SPIVEY, Michael J. (eds.) (2007): *Methods in Cognitive Linguistics*. John Benjamins, Amsterdam/Philadelphia. 19–52.
- ORAVECZ Csaba – VÁRADI Tamás – SASS Bálint (2014): The Hungarian Gigaword Corpus. In: *LREC 2014 Proceedings*. ELRA, Reykjavik. 1719–1723.
- PARTINGTON, Alan. (1998): *Patterns and Meanings: Using Corpora for English Language Research and Teaching*. John Benjamins, Amsterdam/New York.
- SASS Bálint (2009): „Mazsola” – eszköz a magyar igék bővítményszerkezetének vizsgálatára. In: VÁRADI Tamás (szerk.): *Válogatás az I. Alkalmazott Nyelvészeti Doktorandusz Konferencia előadásaiból*. MTA Nyelvtudományi Intézet, Budapest. 117–129.
- SASS Bálint (2017): Keresés korpuszban: a kibővített *Magyar történeti szövegtár* új keresőfelülete. In: FORGÁCS Tamás – NÉMETH Miklós – SINKOVICS Balázs (szerk.): *A nyelvtörténeti kutatások újabb eredményei*. Szegedi Tudományegyetem Magyar Nyelvészeti Tanszék, Szeged. 267–277.
- SINCLAIR, John (1991): *Corpus, Concordance, Collocation*. Oxford University Press, Oxford.

- TALMY, Leonard (2007a): Foreword. In: GONZALEZ-MARQUEZ, Monica – MITTELBERG, Irene – COULSON, Seana – SPIVEY, Michael J. (eds.) (2007): *Methods in Cognitive Linguistics*. John Benjamins, Amsterdam/Philadelphia.
- TALMY, Leonard (2007b): Introspection as a Methodology in Linguistics. Előadás handout: *International Cognitive Linguistics Conference*. Krakow, Poland. 16.07.2007.
- TÁTRAI Szilárd – BALLAGÓ Júlia (2020): A stílustulajdonítás szociokulturális szituáltsága. Funkcionális kognitív pragmatikai megközelítés. *Magyar Nyelvőr* 144 (1). 1–43.
- TOLCSVAI NAGY Gábor (2013): *Bevezetés a kognitív nyelvészetbe*. Osiris Kiadó, Budapest.
- VÁRADI et al. (2018): e-magyar – A Digital Language Processing System. In: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation*, Miyazaki. 1307–1312. <https://aclanthology.org/L18-1208.pdf> (Utolsó elérés: 2022. 09. 09.)
- WILLEMS, Klaas (2012): Intuition, introspection and observation in linguistic inquiry. *Language Sciences* 34 (6). 665–681.

8. Melléklet: Programgyűjtemény

Webes keresőmotorok, keresőfelületek

Google Tudós
Google Könyvek
Google Speciális kereső
Google Speciális képkeresés
Duck Duck Go!

Reguláris kifejezéseket feldolgozó programok

Regexr
RegEx Testing

Online kérdőívkészítők

Google Űrlapok
online-kerdoiv.com

Táblázatkezelők

Microsoft Excel
Open Office Calc
Libre Office Calc
Google Táblázatok

Korpuszok, korpuszlekérdezők

Mai magyar nyelvű szövegek
Magyar Nemzeti Szövegtár (MNSZ2)
Mazsola
Wortschatz Leipzig
Történeti korpuszok
Magyar történeti szövegtár (MTSZ)
Ómagyar korpusz
Történeti magánéleti korpusz (TMK)

Spontán beszéd korpuszok

Budapesti Szociolingvisztikai Interjú (BUSZI)
Magyar Óvodai Beszélt Nyelvi Korpusz (MONYEK)

Kétnyelvű korpuszok

Hunglish Corpus

Korpuszlekérdezők

Voyant Tools

Nyelvtani elemzők

E-magyar
UDPipe

Diagram ábrázolók

RAWGraphs
Draw.io
Datawrapper
Tableau

