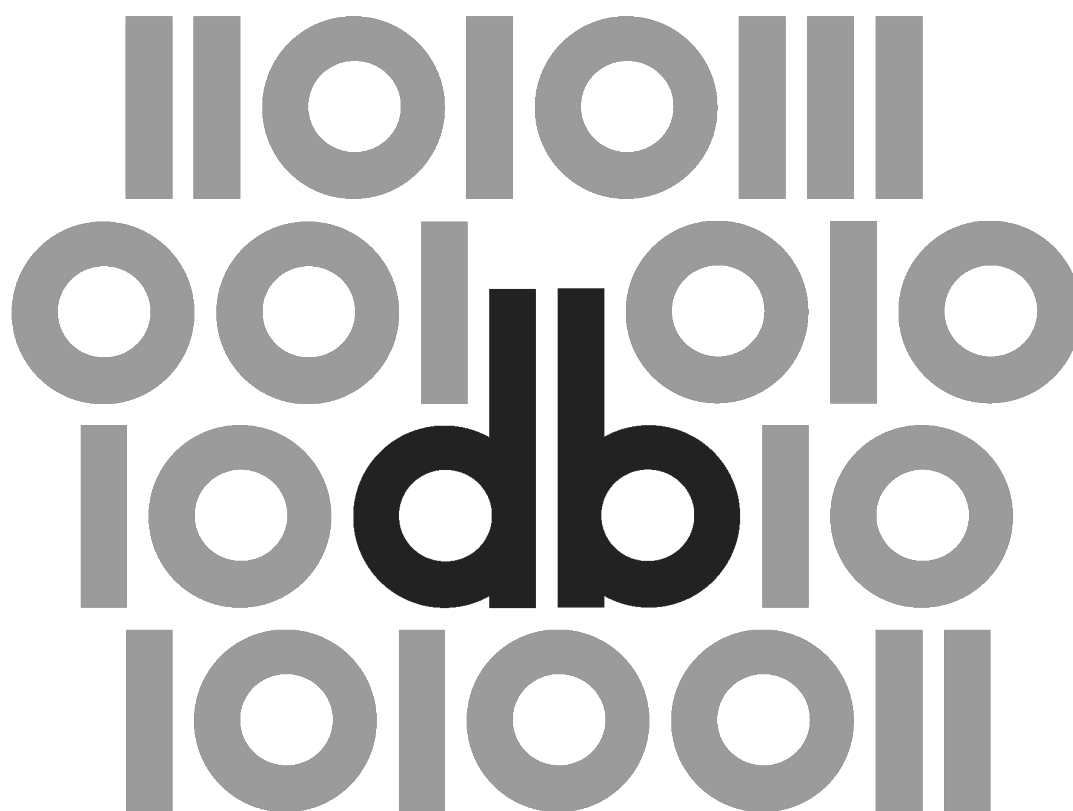


5 (2021) <DIGITÁLIS BÖLCSÉSZET>
A krakkói Computational Stylistics Group
(Különszám)

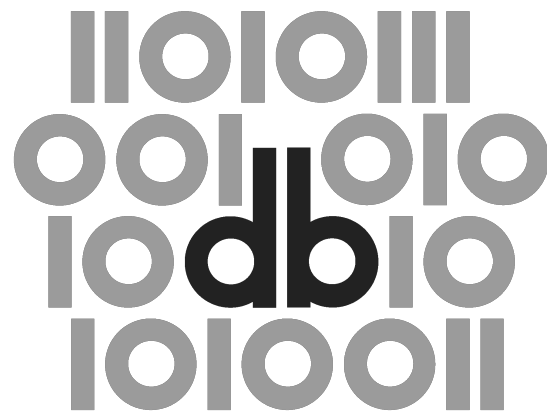


5 (2021) </DIGITÁLIS BÖLCSÉSZET>

Digitális Bölcsészet
2021., ötödik szám

A krakkói Computational Stylistics Group
(Különszám)

<DIGITÁLIS BÖLCSÉSZET>



5 (2021)

A krakkói
Computational Stylistics Group

(Különszám)

A különszámot Szemes Botond szerkesztette.

Felelős szerkesztő:

Maróthy Szilvia

Szerkesztőség:

Kokas Károly, Parádi Andrea

Rovatvezetők:

Tanulmányok: Kiss Margit

Műhely: Péter Róbert

Kritika: Almási Zsolt

Labor: Mártonfi Attila

Tanácsadó testület:

Bartók István, Fazekas István, Golden Dániel, Horváth Iván, Palkó Gábor, Pap Balázs, Sass Bálint, Seláf Levente

Korábbi munkatársaink:

Bartók Zsófia Ágnes (szerkesztő, rovatvezető), Fodor János (szerkesztő),

†Labádi Gergely (szerkesztő, rovatvezető), †Orlovsky Géza (tanácsadó testület)

ISSN 2630-9696

DOI 10.31400/dh-hun.2021.5

Kiadja a Bakonyi Géza Alapítvány és az ELTE BTK Régi Magyar Irodalom Tanszéke (1088 Budapest, Múzeum krt. 4/A).

Felelős kiadó az ELTE BTK Régi Magyar Irodalom Tanszék vezetője.

Megjelenik az Open Journal Systems (OJS) v. 3. platformon, melynek működtetését az ELTE Egyetemi Könyvtár- és Levéltár biztosítja.

A különszám megjelenését a Wacław Felczak Alapítvány támogatta.



WACŁAW
FELCZAK
ALAPÍTVÁNY

FUNDACJA
IM. WACŁAWA
FELCZAKA



Ez a mű a Creative Commons *Nevezd meg! – Ne add el! – Így add tovább! 2.5 Magyarország Licenc* (<http://creativecommons.org/licenses/by-nc-sa/2.5/hu/>) feltételeinek megfelelően felhasználható.

Honlap: <http://ojs.elte.hu/digitalisbolcseszett>

Email cím: dbfolyoirat@gmail.com

Olvasószerkesztő: Bucsics Katalin

Tördelés: Hegedüs Béla

Grafika: Hegyi Gábor

Tartalom

ELŐSZÓ	1
Joanna Byszuk – Szemes Botond <i>A krakkói Computational Stylistics Group bemutatkozása</i> <i>Előszó a Digitális Bölcsészet folyóirat tematikus lapszámához</i>	3
TANULMÁNYOK	1
Maciej Eder <i>Elena Ferrante: Egy „virtuális” szerző</i>	3
Jan Rybicki <i>Vive la différence!</i> <i>Írók nemének azonosítása többváltozós szógyakorisági elemzések során</i>	19
Greta Franzini – Mike Kestemont – Gabriela Rotari – Melina Jander – Jeremi K. Ochab – Emily Franzini – Joanna Byszuk – Jan Rybicki <i>Szerzőazonosítás Jacob és Wilhelm Grimm zajos, digitalizált</i> <i>levelezésében</i>	39
Artjoms Šeļa – Boris Orekhov – Roman Leibov <i>Gyenge műfajok</i> <i>A költői versmérték és a jelentés közötti kapcsolat modellálása</i> <i>az orosz költészetben</i>	69
Albert Leśniak– Zbigniew Pasek <i>Neoprotesztáns és katolikus tanúságtételek a korpuszalapú</i> <i>diskurzuselemzés perspektívájából</i>	91
Helena Grochola-Szczepanek – Ruprecht Von Waldenfels – Rafał L. Górski – Michał Woźniak <i>A szepességi lengyel nyelvjárás korpusznyelvészeti elemzése</i>	113

Albert Leśniak  0000-0002-7141-576X

Institutu Języka Polskiego PAN

albert.lesniak@ijp.pan.pl

Zbigniew Pasek  0000-0003-2580-4366

Akademia Górniczo-Hutnicza w Krakowie

pasek@agh.edu.pl

Neoprotestáns és katolikus tanúságtételek a korpuszalapú diskurzuselemzés perspektívájából

Jelen tanulmány tárgya evangelizáló tanúságtételek komparatív elemzése római katolikus (a Deon.pl internetes portál) és protestáns – pünkösdi-karizmatikus (a *Chrześcijanin* folyóirat stb.) források alapján. A szerzők a korpusznyelvészet eszközeinek felhasználásával rávilágítanak a két korpusz közötti eltérésekre. Érdekesnek bizonyult a „bűn területeinek” rekonstruálása – a rossznak, amit az Isten segítségével a tanúságot tevő emberek legyőztek. Ez két eltérő terület: a katolikus tanúságtételek esetén a szerelem–szex–aszkézis fogalmak köré épül a szemantikai mező. A pünkösdi tanúságtételek az élvezeti cikkek–függőség–patológia fogalmak köré összpontosulnak. A két részből álló kutatás (a korpuszon végzett gyakoriság- és a kulcsszóelemzés, valamint a leglényegesebb kollokációk vizuális bemutatása) eredményeképpen következtetések vonhatók le azokról a domináns jelentéstartományokról, amik meghatározzák a keresztény egyház e két ágához tartozó hívek vallásos életének mentális térképét.

Kulcsszavak:

korpusznyelvészet, keresztény tanúságtételek, lengyel vallási kultúra, korpuszalapú diskurzuselemzés



A modern vallási nyelvben egyre fontosabb helyet foglal el az úgynevezett tanúságtétel, amely Małgorzata D. Nowak definíciója szerint „egy személyes vallási élményről

* Eredeti megjelenés: Albert Leśniak i Zbigniew Pasek, „Świadectwa ewangelikalne i katolickie w perspektywie korpusowej analizy dyskursu,” *Socjolingwistyka* 34 (2020), 57–75, <https://doi.org/10.17651/SOCJOLING.34.4>. Hálásan köszönjük a szerzőknek a magyar nyelvű ábrák létrehozását! – a szerk.)

(az Istennel való »találkozásról«) szóló beszámoló”.¹ A szerző hangsúlyozza, hogy a tanúságtételek általában az „azelőtt (az élmény előtt) – akkor (az élmény) – most (az élmény után)” séma alapján jönnek létre.² Grzegorz Pełczyński megállapítja, hogy a tanúságtétel „az orális (kvázifolklorisztikus) és irodalmi nyelvhasználatban funkcionáló önálló epikai műfajjá” vált.³ Az amúgy igen változatos keresztény vallásos nyelvben a 20. század második felétől nő meg a tanúságtétel mint vallási kifejezés-mód szerepe. A protestáns világból származó terminust (valamint a tanúságtevés gyakorlatát) számos kortárs katolikus megújulási mozgalom is átvette, kicsit másképp hangsúlyozva a tanúságtétel szerepét saját lelkiségükben. A neoprotesztantizmus számára a tanúságtétel továbbra is az evangelizáció egyik fő formája, míg a katolikus vallásosságban ez inkább a számos gyakorlat egyike. A püünkösdi tanúságtétel fő célja az, hogy a hallgatókat ösztönözze életük radikális megváltoztatására vagy megerősítse hitükben a már hívőket. Erőteljesebben megnyilvánul benne a megtérés előtti és utáni helyzet ellentétes bemutatásán alapuló séma. A katolikus gyakorlatban a tanúságtétel nemcsak egy gyökeres lelki fordulatról számol be, hanem arról is, hogyan avatkozik be Isten hétköznapi, kevésbé látványos módon az ember életébe. A keresztény tanúságtételeket vizsgálva érdemes megjegyezni továbbá, hogy domináns nyelvi funkciójuk a meggyőzés.

Jelen tanulmány célja a katolikus és protestáns – püünkösdi – tanúságtételek komparatív elemzése, oly módon, hogy a tanúságtételekre az egyes felekezetek – azaz katolikus és püünkösdi – diskurzusának részeként tekint. Ennek során arra vállalkozunk, hogy körülírjuk a vallási diskurzus e két típusának bizonyos sajátosságait. Aleksandra Pawlikowska nyomán a vallási diskurzust következőképpen értjük:

[...] egy adott felekezeti közösség szövegeinek és megnyilvánulásainak összessége, amely tükrözi azt, hogyan tesz állításokat a világról ez a közösség vallási szempontból. A diskurzus a valóság ábrázolásának meghatározott rendszerét tartalmazza, valamint sajátos ontológiát, axiológiát és kommunikációs stratégiák, illetve szabályok rendszerét.⁴

Cikkünkben a következtetéseket a kulcsszavak és kollokációk hasonlóságainak, illetve eltéréseinek elemzésére alapoztuk. Szándékunk az volt, hogy a korpusz-nyelvészet eszközeinek felhasználásával bemutassuk, miként jelennek meg vallási eszmék a lengyel római katolikusok és a püünkösdi-karizmatikusok (*protestanckich-zielonościowych*, a továbbiakban: püünkösdi) vallási nyelvezetében. Figyelmünket különösen is a bűn és a bűnösség forrásainak kérdésére fordítottuk, mivel a kutatás folyamán nyilvánvalóvá vált, hogy a két korpusz másképpen alakítja ki a „bűn területeinek” profilját.

¹ Małgorzata Danuta Nowak, *Świadectwo religijne: Gatunek – język – styl* (Lublin: Towarzystwo Naukowe KUL. „Nowenna pompejańska”, 2005), hozzáférés: 2021.07.26, <http://mysliborska28.pl/nowenna-pompejanska>.

² Uo., 22.

³ Grzegorz Pełczyński, „«Świadectwa». Opowieści o własnym nawróceniu,” *Lud* 90 (2006): 37–52, 43.

⁴ Aleksandra Pawlikowska, *Zróźnicowanie leksykalnosemantyczne dyskursów wyznaniowych na materiale polskich tekstów ewangelickich, katolickich i prawosławnych* (Wrocław: Quaestio, 2015), 20.

Kutatási módszertan

Kutatásunk során egy, a római katolikus és a pütkösi egyházak híveinek tanúságtételeit tartalmazó szövegkorpuszt vizsgáltunk, az ismétlődő kulcsszavakra és azok kollokációira összpontosítva. A vizsgálat alapjául szolgáló adatokat programozási eszközökkel, illetve a korpusznyelvészet és a kvantitatív nyelvészet eredményeire támaszkodva kidolgozott statisztikai mérőszámok segítségével nyertük ki és különítettük el. Az ilyen típusú nyelvi korpuszok, azaz reprezentatív, digitális formátumú, véges méretű szöveges adatok⁵ ugyanis kiindulópontként szolgálhatnak egy adott nyelvváltozatról (pl. a vallásos nyelvről) szóló állításoknak. Emellett a visszatérő nyelvi minták (nyelvi szokások) feltárására is alkalmasak, így segítségükkel következtethetünk arra, hogy a nyelvhasználók hogyan használják az adott nyelvet a diszkurzív világ konstruálására.⁶ Mellőzve a részletes okfejtéseket magával a diskurzus fogalmával kapcsolatban, szeretnénk kiemelni annak dinamikus jellegét. Ez bizonyos értelemben Michel Foucault meghatározásából fakad: számára a diskurzus nem más, mint „gyakorlatok, amik rendszeresen létrehozzák tárgyuk értelmezését”,⁷ vagyis esetében a „nyelv a cselekvésben” szemlélet érvényesül. Ez a megközelítés pedig – Paul Baker és Tony McEnery megfigyelése szerint – tökéletesen illeszkedik a korpusznyelvészethez, hiszen ez a tudományterület egy adott nyelv nagy szövegtárainak az elemzésére vállalkozik, sőt „legmélyebb értelmében az egész korpusznyelvészet tulajdonképpen diskurzuselemzés”.⁸

A tanulmányban használt két elemzési módszer a lexémák előfordulásának és együttes előfordulásának gyakoriságelemzésén alapul: ez a kulcsszavak és a kollokációk kinyerésének a művelete. A korpuszvizsgálatokban a kulcsszavakra két fő feltevés vonatkozik: először is a jelentés nem oszlik szét egyenletesen a szövegben, azaz vannak szavak, amelyek szemantikai potenciálja nagyobb, „jelentésteljesebbek”. Másodszor, a tágabb jelentéssel bíró szavak megkülönböztetése mérhető, ami alapjául szolgál annak, hogy azokat a szavakat tekintsük kulcsszavaknak, amelyek gyakorisága a korpuszban a referenciakorpuszhoz képest statisztikailag jelentős. Amíg a szavak gyakorisági listája azt írja le, hogy a szavak milyen (relatív vagy kumulatív) gyakorisággal fordulnak elő a korpuszban, a kulcsszavak listája azt mutatja meg, hogy mennyire lényegesek ezek a szavak a vizsgált korpusz szempontjából.⁹ Mike Scott szerint a kulcsszavak olyan „szavak, amik egy adott szövegben váratlanul gyakran fordulnak elő”,¹⁰ és Christopher Tribble-lel együtt hozzáteszi:

⁵ Vö. Tony McEnery and Andrew Wilson, *Corpus Linguistics: An Introduction* (Edinburgh: Edinburgh University Press, 2001).

⁶ Paul Baker, *Using Corpora in Discourse Analysis* (London, New York: Continuum, 2006), 1, <https://doi.org/10.5040/9781350933996>.

⁷ Michael Foucault, *A tudás archeológiája*, ford. Perczel István (Budapest: Atlantisz Könyvkiadó, 1972), 49.

⁸ Paul Baker and Tony McEnery, eds., *Corpora and Discourse Studies: Integrating Discourse and Corpora* (London: Palgrave Macmillan, 2015), 4, <https://doi.org/10.1057/9781137431738>.

⁹ Baker, *Using Corpora*, 125.

¹⁰ Mike Scott, „PC Analysis of Key Words – And Key Key Words,” *System: An International Journal of Educational Technology and Applied Linguistics* 25, 2. sz. (1997): 233–245, 236, [https://doi.org/10.1016/S0346-251X\(97\)00011-0](https://doi.org/10.1016/S0346-251X(97)00011-0).

[...] a kulcsfontosság olyan tulajdonság, amivel szavak rendelkezhetnek egy adott szöveg vagy szöveggyűjtemény keretei között és aminek megléte azt sugallja, hogy az érintett szavak fontosak és visszatükrözik, hogy miről szól a szöveg valójában, kihagyva az apróságokat és a lényegtelen részleteket.¹¹

Nyelvstatisztikai szempontból ezek tehát olyan lexémák, amelyek gyakorisága egy adott korpuszban váratlanul magas (pozitív kulcsszavak) vagy alacsony (negatív kulcsszavak), emellett pedig olyan szemantikai potenciállal rendelkeznek, ami fontos információkat hordozhat magáról a szövegről vagy a korpuszról.

A diskurzuselemzés számára a kulcsszavak sajátos nyelvi esetek, amelyekre alapozva történik a társadalmi és kulturális valóság konstrukciója. Ennek a kultúraalkotó potenciált figyelembe vevő megközelítésnek előfutára William Raymond volt, aki a mára már híressé vált *Keywords* című munkájában felhívta a figyelmet arra, hogy a kulturális kulcsszavak nemcsak koruk kultúráját írják le, hanem azok is használják őket, akik erről a kultúráról folyó diskurzusban részt vesznek. Williams szerint a kulturális kulcsszavak „szavak és jelentések közös gyűjteménye a legáltalánosabb angol nyelvű diskurzusban azokról a gyakorlatokról és intézményekről, amiket együttesen kultúrának és társadalomnak nevezünk”.¹² Williams nem használt statisztikai módszereket a kultúra szempontjából jelentős szemantikai potenciállal rendelkező szavak kinyerésére, főleg azért nem, mert még nem állt rendelkezésére olyan módszer, ami lehetővé tette volna megfelelő nagyságrendben a szövegek vizsgálatát. Alan Duran úgy vélte:

A nagy nyelvi korpuszokhoz használt digitális keresési lehetőségek fejlődése (a módszerek, amiket együttesen korpusznyelvészetnek hívunk) arra ösztönöz, hogy figyelmünket újra a kultúrával kapcsolatos kulcsszavak felé irányítsuk.¹³

Hangsúlyozzuk, hogy egy kulcsszólista létrehozása a kultúrával foglalkozó tudományok esetében csak egy bevezető szakasz lehet, amely a vizsgált szavak kontextusának mélyebb elemzéséhez vezet el. Ahogy Victoria Kamasa írja: „a kulcsszavak fontos útmutatók, amik a kutatók figyelmét ráirányítják azokra a jelenségekre, amik az általuk vizsgált diskurzusra jellemzőek. Egyfajta kiindulási pontot jelenthetnek a gyűjtött adatokhoz.”¹⁴

A korpuszvizsgálatok legtöbbször emergens és disztributív jellegűek.¹⁵ Anna Bączkowska szerint az emergencia, azaz amikor a következtetések levonása előzetes feltevések nélkül, kizárólag a korpuszelemzésből fakadó megfigyelések alapján történik,

¹¹ Mike Scott and Christopher Tribble, *Textual Patterns: Keyword and Corpus Analysis in Language Education* (Amsterdam: John Benjamins Publishing, 2006), 55–56, <https://doi.org/10.1075/sc1.22>.

¹² Raymond Williams, *Keywords: A Vocabulary of Culture and Society* (London: Fontana, 1983), 15.

¹³ Alan Durant, „Raymond Williams’s *Keywords*: Investigating Meanings ‘Offered, Felt for, Tested, Confirmed, Asserted, Qualified, Changed,’” *Critical Quarterly* 48, 4. sz. (2006): 1–26, 16, <https://doi.org/10.1111/j.1467-8705.2006.00743.x>.

¹⁴ Victoria Kamasa, „Techniki językoznawstwa korpusowego wykorzystywane w krytycznej analizie dyskursu. Przegląd,” *Przegląd Socjologii Jakościowej* 10, 2. sz. (2014): 100–117, 107.

¹⁵ Barbara Lewandowska-Tomaszczyk, *Podstawy językoznawstwa korpusowego* (Łódź: Wydawnictwo Uniwersytetu Łódzkiego, 2005).

lényegesen felértékeli a kontextus szerepét, mert feltételezi, hogy a jelentés nemcsak egyes szavak statikus definíciójában rejlik, hanem azok kontextusában is, amit a vizsgált szót körbevevő szavak csoportja kódol.¹⁶ Bączkowska szerint a kontextus részt vesz a kulcsszavak jelentésének megalkotásában, profilálja és pontosítja őket és rámutat a rájuk jellemző szemantikai preferenciákra. Érdemes itt megfigyelni, hogy a kontextus fogalmát szélesebben is lehet érteni, nemcsak mint az adott kulcsszót körbevevő lexikai egységek csoportját, hanem arra a kommunikációs helyzetre is vonatkoztatva, amiben az adott megnyilatkozás elhangzott.¹⁷ Ehhez kapcsolódóan a kontextust a következő értelemben is használni fogjuk: „egy struktúra, ami tartalmazza a társadalmi helyzet összes olyan tulajdonságát, ami lényeges a diskurzus létrehozása és befogadása szempontjából”.¹⁸

Szorosabb értelemben a kontextus azáltal is rekonstruálható, hogy azonosítjuk az úgynevezett kollokációkat, vagyis azokat a szavakat, amelyek hajlamosak az együttes előfordulásra, gyakrabban, mint ahogy ez a valószínűség-elmélet szerint következne – feltéve, hogy a szavak eloszlása a korpuszban véletlenszerű. John R. Firth szintén úgy vélte, hogy a szó jelentéseinek egy része meghatározható kollokációi alapján, például az *éjszaka* (angolul: *night*) szó egyik jelentése megfogalmazható a *sötét* (angolul: *dark*) szón keresztül, ami gyakran kapcsolódik hozzá. Az állítását ebben a híres mondatban foglalta össze: „egy szót megismerhetsz a környezetéből”.¹⁹ Baker számára azonban a kollokációk elemzése egy „módszer, amin keresztül megérthetjük a szavak jelentéseit és a szavak közti kapcsolatokat, amiket kisebb skálán, egyetlen szöveg elemzése alapján nehéz lenne meghatározni”.²⁰ A diskurzuselemzésben az effajta kollokációs profilokat összetettebb diszkurzív struktúrák azonosítására használjuk:

[...] legtöbbször azért használjuk őket, hogy részletesebb információkat szerezünk az adott szavak működéséről a vizsgált szövegekben. Ilyen információ alapján a kutató azután azonosíthatja a diszkurzív struktúrákat.²¹

A statisztikai jelentőség mérése a két szó együttes előfordulása esetén sok szempontból hasonló lehet a korpuszok lexikális eltéréseinek kulcsszavak alapján történő méréséhez, sőt gyakran használják ebből a célból mindkét esetben ugyanazt az asszociációs mértéket. A lexikális egységeket (mind a kulcsszavakat, mind a kollokációkat), amelyek magas gyakorisága a korpuszban statisztikailag jelentős, azaz a véletlenszerűnél jóval magasabb, a jelen vizsgálatban a *log-likelihood* módszer segítségével azonosítottuk, amit először Ted Dunning javasolt a Vaclav Brezina kézikönyvében leírt eljárás szerint.²² Ez az egyik leggyakrabban használt módszer, amivel meg lehet határozni, hogy az esemény előfordulása (azaz a szó előfordulása az A korpuszban a B korpuszhoz

¹⁶ Anna Bączkowska, „Korpusowa analiza dyskursu związanego z tematyką imigracji w brytyjskiej prasie opiniotwórczej,” *Conversatoria Linguistica* 10 (2016): 8–9.

¹⁷ Vö. Barbara Boniecka, „Tekst w kontekście (problemy metodologiczne),” *Polonica* 16 (1994): 43–67.

¹⁸ Katarzyna Skowronek, *Między sacrum a profanum: Studium językoznawcze listów pasterskich Konferencji Episkopatu Polski (1945–2005)* (Kraków: Wydawnictwo Lexis, 2006), 26–27.

¹⁹ John Rupert Firth, *Papers in Linguistics 1934–1951* (London: Oxford University Press, 1957), 179.

²⁰ Bake, *Using Corpora*, 196.

²¹ Kamasa, *Techniki językoznawstwa*, 108.

²² Lásd Ted Dunning, „Accurate Methods for the Statistics of Surprise and Coincidence,” *Computational Linguistics* 19, 1. sz. (1993): 61–74, illetve: Vaclav Brezina, *Statistics in Corpus Linguistics: A Practical*

képest, vagy két szó együttes előfordulása egy korpuszban) statisztikailag jelentős-e, vagy véletlenszerű. Az összes számítást lemmatizált formátumok alapján végeztük, ingyenes licenc alapján használható *Python 3.7* programozási környezetben. A számítások (a generált adatokkal együtt) elérhetők a <http://github.com/alblesniak/swiadectwa> oldalon.

A kollokációs eredmények egyik lehetséges bemutatási formája a gráf és a kollokációs hálózat.²³ Ez a vizuális módszer lehetővé teszi bonyolult összefüggések elemzését a nyelvi adatkészletben, ami hagyományos táblázatos formában jóval nehezebben volna vizsgálható. Míg a kollokációs gráfokban a csomópontok a leglényegesebb kollokációkat mutatják (egy definiált távolság alapján, amivel meghatározzuk a vizsgált szó előtt és után álló szavaknak azt a lehetséges számát, amit a számítás során figyelembe veszünk, illetve az asszociációs mérték egy szintén előre meghatározott minimális értéke esetén), a kollokációs hálózatok még egy lépést tesznek előre, és a hálózat csomópontjaiban több (esetünkben 50) kulcsszó kapcsolódásait is mutatják, illetve figyelembe veszik a másodrendű kollokációkat (a kollokációk kollokációit) is. A két csomópontot összekötő élek szélessége a kollokációs erőnek felel meg, a csomópontok nagysága pedig a szavaknak a referenciakorpuszhoz képest mutatott kulcsfontosságát ábrázolja.

Munkánk során azt is be akartuk mutatni, hogyan nyitnak utat a korpusznyelvészeti eszközök a digitális bölcsészet irányába, és hogy nemcsak hatékonyabbá teszik a hatalmas méretű szövegtörzsekkel (adatgyűjteményekkel) folytatott munkát, hanem annak köszönhetően, hogy alkalmasak tendenciák (vagy azok hiányának) bizonyítására, lehetővé teszik a korábbi megfigyelések megerősítését vagy elutasítását, ebből fakadóan pedig megbízhatóbb következtetéseket tesznek megfogalmazhatóvá egy adott csoport modern vallási nyelve és vallási kultúrája terén is.

Korpuszok

A vizsgált korpusz összesen 281746 tokent tartalmaz (373 dokumentumot). Két kisebb alkorpuszból áll: a Pütkösdi Egyház hívei által leírt tanúságtételek (73554 token, 78 dokumentum) és a Római Katolikus Egyház híveinek tanúságtételei (208192 token, 295 dokumentum). A protestáns tanúságtételek korpuszában található szövegek két forrásból származnak: a *Chrześcijanin* folyóiratból – a Pütkösdi Egyház lapjából – és az egyház gyülekezeteinek hivatalos weboldalairól. A referenciakorpusz alapját képező katolikus tanúságtételeket a Deon.pl katolikus internetes portálról gyűjtöttük. Az adatok gyűjtésére automatizált módszert alkalmaztunk, a szövegeket a Scrapy keresőrobot-rendszer felhasználásával, úgynevezett *web scraping* módszerrel nyertük

Guide (New York: Cambridge University Press, 2018), <https://doi.org/10.1017/9781316410899>; Wit Piotr Chlondowski, „Pentekostalizacja w ujęciu kardynała J. Ratzingera oraz dokumentów episkopatu – dar czy zagrożenie?”, hozzáférés: 2021.07.30, <https://deon.pl/magazyn/pentekostalizacja-w-ujeciu-kard-j-ratzingera-oraz-dokumentow-episkopatu-dar-czy-zagrozenie,481475>.

²³ Vaclav Brezina, „Collocation Graphs and Networks: Selected Applications,” in Pascual Cantos-Gómez and Moisés Almela-Sánchez, eds., *Lexical Collocation Analysis: Quantitative Methods in the Humanities and Social Sciences*, 59–83 (Cham: Springer, 2018), https://doi.org/10.1007/978-3-319-92582-0_4.

ki.²⁴ A protestáns alkorpuszhoz tartozó szövegek túlnyomó többsége a 20–21. század fordulójáról származik, a katolikus tanúságtételek pedig a 2010–2019 közötti évtizedben keletkeztek.

Annak ellenére, hogy a forrásszövegeket különféle médiumokból gyűjtöttük, kommunikációs csatornájuk homogén, ugyanis mindegyik írott szöveg. A szövegekhez való szerkesztői hozzájárulás mértéke nehezen megállapítható, némi óvatossággal mégis kijelenthetjük, hogy szerzői jellegűek, keletkezésük ideje pedig hasonló, így területi, kronológiai és szociológiai-kulturális szempontból is egységesek.²⁵ A vizsgált korpuszok közötti méretbeli különbség abból fakad, hogy a katolikus forrásszövegek könnyebben elérhetőek digitális formátumban. Érdeemes emellett szem előtt tartani azt is, hogy a lengyelországi pütkösi egyház híveinek száma (kb. 25 ezer fő) jóval kisebb a katolikus egyháznál, ami nem teszi lehetővé egymásnak teljesen megfelelő korpuszok létrehozását. Bár az alkorpuszok a dokumentumok számában eltérnek, a szövegek átlaghosszúsága hasonló, a méretbeli különbségekből eredő esetleges problémákat pedig olyan módon minimalizáltuk, hogy az analízisben a tokenek relatív előfordulási számát (a korpusz méretével normalizált számot), valamint megfelelően megválasztott statisztikai mérőszámot használunk. Ennek fényében az a véleményünk, hogy a korpuszok reprezentatív és összehasonlítható jellegűek.

1. táblázat. A vizsgált korpuszok statisztikai jellemzői.

Tanúságtételek	PROTESTÁNS	KATOLIKUS
Szövegek száma	78	295
Tokenek száma	73 554	208 192
Lemmák száma	6209	11 448
Átlagos tokenszám egy dokumentumban	943	705,8
Minimális tokenszám egy dokumentumban	193	28
Maximális tokenszám egy dokumentumban	5201	3299

A korpuszban szereplő összes dokumentumot tokenekre bontottuk, majd ezeket (a központozás törlése után) lemmatizáltuk, azaz nyelvtani alapalakjukra, úgynevezett lemmákba csoportosítottuk. Ez a folyamat lehetővé teszi minden lexéma számszerűsítését, függetlenül attól, hogy milyen ragozott alakban szerepeltek a szövegben. Az így előkészített korpuszon különféle statisztikai számítási sorozatok végezhetőek el az úgynevezett asszociációs mértékek felhasználásával; ezeket két elem közti korreláció vizsgálatában alkalmazzuk, mint például egy lexéma gyakorisága a két különböző korpuszban vagy két lexéma együttes előfordulása ugyanabban a korpuszban. Az első

²⁴ Lásd Ryan Mitchell, *Ekstrakcja danych z językiem Python: Pozyskiwanie danych z Internetu*, tłum Krzysztof Sawka (Gliwice: Helion, 2019).

²⁵ Jadwiga Sambor, *Słowa i liczby: Zagadnienia językoznawstwa statystycznego* (Wrocław: Zakład Narodowy im. Ossolińskich, 1972), 24.

említett eset lehetővé teszi egy adott korpuszra jellemző lexémák, azaz a kulcsszavak meghatározását, a második segítségével pedig kinyerhetjük azokat a szókapcsolatokat, amelyek hajlamosak az együttes előfordulásra – ezek az úgynevezett kollokációk.

A kutatás eredményei 1.

Kulcsszavak

Az itt bemutatott 2. és 3. táblázatban összesítve mutatjuk a vizsgálat eredményeit. A jelen dolgozatban csak az 50 leggyakoribb kulcsszót vettük figyelembe a két korpuszból. Következtetéseink levonásához felhasználtuk a táblázatokban található további adatokat is; ezeket elérhetővé tettük a Github portálon.

2. táblázat. Az 50 leggyakoribb kulcsszó a katolikus tanúságtételek korpuszában, a *log-likelihood* érték alapján számolva.

	Kulcsszó	log-likelihood érték	Előfordulások száma 1000 szóra
1	rekolekcja – 'lelkigyakorlat'	101,5917	1,124
2	spowiedź – 'gyónás'	80,5576	0,8598
3	świadek – 'tanú'	71,0707	1,4266
4	chłopak – 'srác'	70,6767	1,0231
5	związki – 'kapcsolat'	65,8248	0,9126
6	czystość – 1. 'tisztaság', 2. 'szüzesség'	56,9192	0,6148
7	różaniec – 'rózsafüzér'	53,0052	0,538
8	cięża – 'terhesség'	50,2416	0,5572
9	odmawiać – 1. '[imát] mondani', 2. 'visszautasítani'	46,1537	0,5956
10	Maryja – 'Mária'	43,5417	0,6052
11	modlitwa – 'ima'	36,388	2,6802
12	ślub – 1. 'esküvő', 2. 'fogadalom'	36,2593	0,8358
13	badanie – 'vizsgálat'	35,5301	0,3891
14	sakrament – 'szentség'	34,9741	0,3843
15	intencja – 'intenció, szándék'	34,4187	0,3795
16	tydzień – 'hét'	34,1442	1,1096
17	matka – 'anya'	33,8066	0,7829
18	małżeństwo – 'házasság'	29,4044	0,8021
19	ksiądz – 'pap'	27,0809	0,8406
20	ciężki – 'nehéz'	24,4665	0,3602
21	miłosierdzie – 'irgalom'	24,3482	0,3266
22	cierpienie – 'szenvedés'	23,3219	0,3795

23	kolejny – 'következő'	23,2697	1,3305
24	siebie – 'magát'	23,2233	5,2404
25	miesiąc – 'hónap'	21,8457	1,0951
26	operacja – 'műtét'	20,9935	0,3266
27	poród – 'szülés'	19,719	0,2113
28	przypadek – 1. 'eset', 2. 'véletlenség'	19,6632	0,4227
29	dług – 'adósság'	19,1594	0,2065
30	dziewczyna – 'lány'	18,7765	0,5091
31	uzdrowienie – 1. '(meg)gyógyítás', 2. 'gyógyulás'	18,271	0,3554
32	walczyć – 'küzdeni'	18,1164	0,3266
33	pojawiać – 'megjelenni'	17,1941	0,317
34	dobra – 'jó'	16,7309	0,2546
35	konkretny – 'konkrét'	16,3781	0,1825
36	choroba – 'betegség'	16,3686	0,5236
37	zaufanie – 'bizalom'	16,0594	0,2161
38	przyczyna – 'ok'	15,8257	0,1777
39	pierwsza – 'első'	15,681	1,1192
40	walka – 'küzdalem'	15,6145	0,3266
41	seks – 'szex'	15,6145	0,3266
42	warto – 'érdemes'	15,2374	0,269
43	lekarz – 'orvos'	14,979	0,61
44	oglądać – 'néz'	14,7715	0,2642
46	obój – 'oboa'	14,7251	0,1681
45	owoc – 'gyümölcs'	14,7681	0,2354
47	wspólny – 'közös'	14,5784	0,4995
48	kryzys – 'válság'	14,1772	0,1633
49	znak – 'jel'	13,8904	0,3074
50	doświadczenie – 'tapasztalat'	13,5202	0,6677

3. táblázat. Az 50 leggyakoribb kulcsszó a protestáns tanúságtételek korpuszában, a *log-likelihood* érték alapján számolva.

	Kulcsszó	log-likelihood érték	Előfordulások száma 1000 szóra
1	bóg – 'isten'	189,9654	12,5894
2	Jezus – 'Jézus'	173,5969	6,3627
3	Biblia – 'Biblia'	164,6221	1,278
4	życie – 'élet'	161,3172	9,8295
5	zbór 'gyülekezet'	154,2113	0,8293
6	alkohol – 'alkohol'	133,1858	1,3595
7	Chrystus – 'Krisztus'	112,4584	1,7402
8	czytać – 'olvasni'	96,0239	1,645

9	chrzest –'keresztelés, keresztelő'	86,2822	0,6798
10	społeczność –'közösség'	85,4751	0,4758
11	pastor –'lelkész'	77,085	0,4622
12	kościół – 1. 'templom', 2. 'egyház'	73,506	2,5831
13	nabożeństwo –'istentisztelet'	72,8846	0,7613
14	papieros –'cigaretta'	67,4022	0,5166
15	ośrodek –'központ'	64,5074	0,6526
16	więzienie –'börtön'	57,3104	0,3807
17	przyjąć –'elfogadni'	56,7842	1,0876
18	pić – 'inni'	54,6864	0,6254
19	zbawienie 'megváltás'	54,1362	0,5302
20	słowo –'szó'	52,7294	2,8414
21	chrześci- jański –'keresztény'		
		52,0894	0,503
22	zbawić –'megváltani'	50,7596	0,4351
23	testament –'testamentum'	49,8966	0,3399
24	wieczny –'örök'	48,5263	0,4215
25	człowiek –'ember'	46,4807	4,133
26	Basia (női becenév)	43,8891	0,4079
27	uczęszczać –'jár, eljár [pl. közösségbe, iskolába]'	43,1411	0,3399
28	pustka –'üresség'	42,3944	0,503
29	pieśń –'ének'	39,5986	0,3807
30	wodny –'vízi'	38,5425	0,2311
31	dom – 'ház'	38,3874	2,488
32	pismo – 'írás'	37,2055	0,5846
33	wieczność – 'örökkévalóság'	35,3232	0,2583
34	narkotyki – 'kábitószer'	33,65	0,4622
35	zbawiciel – 'megváltó'	33,3054	0,4486
36	żyć – 'él'	32,722	1,7266
37	religia – 'vallás'	31,4581	0,3535
38	boży – 'isten'	30,1777	2,2432
39	grzesznik –'bűnös'	29,2594	0,2719
40	serce – 'szív'	28,6687	2,8686
41	swój – 'saját'	28,4267	2,2432
42	Wojtek (férfi becenév)	28,3178	0,1767
43	werset – '(szentírási) vers'	28,3178	0,1767
44	wołać – 'hívni, kiáltani'	27,3432	0,3127
45	duch – 'lélek'	27,2701	1,4547

46	palić – 1. 'éget', 2. 'cigaret-tázik'	27,084	0,2855
47	śpiewać – 'énekelni'	23,7291	0,3943
48	syn – 'fiú(gyermek)'	23,7145	1,1284
49	kolega – 'kolléga'	23,6056	0,5982
50	przyjść – 'jönni'	23,6048	0,9789

A 2. és 3. táblázatban foglalt adatok elemzése alapján az alábbi következtetések fogalmazhatók meg. Először is fontos megjegyezni, hogy a katolikus korpuszban erőteljesen jelen van az általunk ideiglenesen szexuális-aszketikusnak elnevezett szemantikai mező (szerelem–szex–aszkezis). A listán szereplő első 1000 kulcsszó elemzése kimutatta, hogy ez a szemantikai szerkezet olyan magas gyakoriságú lexémákból áll, mint: *chłopak* 'fiú, barát', *związki* 'kapcsolat', *czystość* 'tisztaság, szüzesség', *cięża* 'terhesség', *małżeństwo* 'házasság', *poród* 'szülés', *dziewczyna* 'lány, barátnő', *seks* 'szex', *małżeński* 'házassági', *ukochana* 'kedves[em][nőnemű]', *łóżko* 'ágy', *prytulic* 'ölel', *zakochać* 'beleszeret', *rodzinny* 'családi', *kobieta* 'nő'. A protestáns (pünkösdi) korpuszban azonban egy másik szemantikai mező figyelhető meg, amit a következőképpen lehet megfogalmazni: élvezeti cikkek–függőség–patológia. Olyan szavakból áll, mint: *alkohol* 'alkohol', *papieros* 'cigaretta', *pić* 'inni', *więzienie* 'börtön' és ehhez hasonlók. Szeretnénk hangsúlyozni, hogy tudatosan kihagytunk olyan nyilvánvaló jelenségeket, mint például hogy a protestáns korpuszban olyan szavak dominálnak, mint: *czytać* 'olvasni', *Biblia* 'Biblia', *pismo* 'írás', *testament* 'testamentum', *pastor* 'lelkész', a katolikus korpuszban pedig: *różaniec* 'rózsafüzér', *Maryja* 'Mária', *sakrament* 'szentség', *ksiądz* 'pap', *eucharystia* 'eucharisztia'.

Az elemzésben azokra a szavakra összpontosítottunk, amelyek többé-kevésbe kapcsolatban állnak a kiemelt fogalmi tartományokkal. A két tartományt alkotó lexémák relatív és kumulatív gyakoriságára vonatkozó adatok a 4. táblázatban találhatóak. A komparatív elemzés eredményeként képet kaptunk az általunk kulcsszóként meghatározott szavak gyakoriságában előforduló aránytalanságokról.

4. táblázat. A szerelem–szex–aszkezis szemantikai szerkezetre jellemző lexémák.

Kulcsszó	Előfordulások száma a katolikus korpuszban	Előfordulások száma a katolikus korpuszban 1000 szóra	Előfordulások száma a protestáns korpuszban	Előfordulások száma a protestáns korpuszban 1000 szóra	Különbség
chłopak – 'fiú barát'	213	1,02	11	0,15	0,87
związki – 'kapcsolat'	190	0,91	9	0,12	0,79
czystość – 1. 'tisztaság', 2. 'szüzesség'	128	0,61	3	0,04	0,57
małżeństwo – 'házasság'	167	0,8	19	0,26	0,54

cięża – 'terhesség'	116	0,56	3	0,04	0,52
dziewczyna – 'lány'	106	0,51	12	0,16	0,35
seks 'szex'	68	0,33	6	0,08	0,25
poród – 'szülés'	44	0,21	1	0,01	0,2
małżeński – 'házassági, házas'	44	0,21	5	0,07	0,14
łóżko – 'ágy'	54	0,26	9	0,12	0,14
kobieta – 'nő'	102	0,49	26	0,35	0,14
rodzinny – 'családi'	47	0,23	9	0,12	0,1
ukochana – 'kedves[em] [nőnemű]'	21	0,1	1	0,01	0,09
zakochać – 'beleszeret'	20	0,1	2	0,03	0,07
przytulić – 'átölelni'	15	0,07	1	0,01	0,06

5. táblázat. Az élvezeti cikkek–függőség–patológia szemantikai szerkezetre jellemző lexémák.

Kulcsszó	Előfordulások száma a katolikus korpuszban	Előfordulások száma a katolikus korpuszban 1000 szóra	Előfordulások száma a protestáns korpuszban	Előfordulások száma a protestáns korpuszban 1000 szóra	Különbség
alkohol – 'alkohol'	36	0,17	100	1,36	1,19
pić – 'inni'	20	0,1	46	0,63	0,53
papieros – 'cigaretta'	7	0,03	38	0,52	0,48
narkotyki – 'kábitószer'	19	0,09	34	0,46	0,37
więzienie – 'börtön'	3	0,01	28	0,38	0,37
palenie – 1. 'égetés', 2. 'cigaretttázas'	4	0,02	15	0,2	0,18
picie – 'ivás'	4	0,02	15	0,2	0,18
wyrok – 'ítélet'	4	0,02	15	0,2	0,18

alkoholik – 'alkoholista'	8	0,04	14	0,19	0,15
karny – 'bűntető'	1	0	8	0,11	0,1
kac – 'másnaposság'	1	0	6	0,08	0,08

Az előfordulások számát 1000 szóra vetítve megalapozottnak bizonyult kiemelni a katolikus korpuszban a férfi és nő, fiú és lány közötti kapcsolatra vonatkozó fogalmi tartományt, melynek hangsúlya erősen szexuális és aszketikus (e tekintetben lásd a *czystość* 'szüzesség' szó magas pozícióját, ami a szexuális önmegtartóztatásra vonatkozik). A gondolatmenetet folytatva feltehetjük a kérdést: lehet-e ezeket a statisztikai jellegű tendenciákat úgy értelmezni, mint annak a tézisnek a megerősítését, hogy a Lengyelországban ma élő katolikusok alapvető problémaként tapasztalják meg a (szexuális) kapcsolatok erotikus dimenziójában a két partner között fellépő feszültségeket. A tanúságtételek elemzése „Isten működésének erejét” ezeknek a kapcsolatoknak a rendezésében mutatja ki. A kapcsolatok ezen szférája kiemelkedik mint a vallásos nyugtalanság, a problémák, vagy – keresztény nyelvezettel – a bűn fő forrása. Az elemzés kimutatja, milyen intenzitással és milyen mértékben van jelen a szexualitás komplexuma ezekben a tanúságtételekben, ami alátámasztja azt az tézist, hogy ez a szféra egyike azoknak, amik a katolikusok mindennapi lelki küzdelmeit dominálják. A probléma mérlegelése során érdemes felidézni azt a vallásszociológusok által már számtalanszor kiemelt ellentmondást, ami az Egyház erre vonatkozó tanítása és a lengyel katolikusoknak ezzel szembenő hozzáállása és gyakorlata között feszül (például a házasság előtti szex vagy a fogamzásgátló szerek alkalmazásának tiltása). Vallásszociológiai kutatások szerint a katolikus egyház által hirdetett etikai szabályokat minden ötödik lengyel fogadja el.²⁶

A Deon.pl portál szerzői (a *Chrześcijanin* szerkesztőihez hasonlóan) a publikáció előtt válogatnak a beérkezett tanúságtételek között. Feltételezhetjük, hogy a szövegek közzététele során figyelembe veszik azok aktualitását, vagyis hogy kapcsolódnak-e a hívők mindennapi problémáihoz vagy az evangelizáció szempontjából potenciálisan hatékonyak-e. A közzétett tanúságtételek azokra a problémákra reflektálnak, amikkel a hívők együttélnek, ezért úgy tűnik, hogy az ezek alapján felállított diagnózis érvényes. A szerelem szexuális dimenziójával kapcsolatos fogalmi komplexum jelenlétét csak részben magyarázza az elsősorban a fiatalokhoz szóló Deon.pl portál sajátos célcsoportja (bár a korpuszban számos idősebbektől származó tanúságtétel található, a szerzők életkorának megállapítása sajnos nem mindig lehetséges). Így hát, némileg szűkítve korábbi megállapításunkat, felállíthatjuk azt a tézist, hogy a fiatal lengyel katolikusok számára jelentős, sőt kulcsfontosságú problémát jelent az intim élet területén összeegyeztetni az egyház tanítását és személyes szükségleteiket.

Érdemes azt is megjegyezni, hogy a katolikus korpuszban a tanúságtételek 65%-a női szerzőtől származik (6. táblázat). Ez a besorolás automatikusan történt, az egyes vagy

²⁶ Többek között: Janusz Mariański, „Religia i moralność w świadomości Polaków: zależność czy autonomia?” *Konteksty Społeczne* 3, 1. sz. (2015): 12–13.

többes számú, múlt idejű igék ragozott alakjainak gyakorisága alapján – ezek a lengyel nyelvben rendelkeznek olyan tulajdonságokkal, amelyekből megállapítható a beszélő neme. Ez a tény részben megmagyarázza például azt a jelenséget, hogy a *chłopak* 'fiú, barát' szó kétszer gyakrabban fordul elő, mint a *dziewczyna* 'lány, barátnő'.

6. táblázat. A tanúságtételek megoszlása a szerző neme szerint.

Tanúságtételek	férfiként besorolt szerző	nőként besorolt szerző	nem megállapítható
katolikus	89 (30,07%)	194 (65,54%)	13 (4,39%)
protestáns	37 (47,44%)	41 (52,56%)	0 (0,0%)

Az erotikára vonatkozó fogalmi mező jelenlétét kimutató adatok olyan általánosabb trendeket is megvilágítanak, mint például azt, miért növekedett meg az utóbbi tíz évben azoknak a lelki útmutatóknak (könyveknek, közösségi médiában látható filmeknek, szemináriumoknak, workshopoknak és lelkigyakorlatoknak) a népszerűsége, amelyek témájukban egyre gyakrabban érintik az intim természetű kérdéseket és problémákat keresztény perspektívából. Elegendő itt megemlíteni például a kapucinus szerzetes, Ksawery Knotz katolikus kiadói tevékenységét²⁷ vagy a rendszeres médiaszereplő domonkos szerzetest, Adam Szustakot, akinek YouTube-csatornáját több mint 500000 felhasználó követi.

A pütkösdí tanúságtételek szerzői vegyes társadalmi háttérrel rendelkeznek, az általuk leírtakat pedig a lehető legszélesebb közönségnek szánják. A „bűn és bűnösség” hasonló tartományát keresve ebben a korpuszban érdemes megfigyelni, milyen mértékben felülreprezentáltak (statisztikailag gyakoribbak) azok a szavak, amelyekből kirajzolódik az élvezeti cikkek – függőség – patológia tartományra vonatkozó szemantikai mező. A pütkösdista tanúságtételek magas szintű sematizációja (a szavak/témák változatosságának alacsonyabb foka) alapján ez a fogalmi mező elsősorban a megtérés, a neoprotestáns hívő (ebbe az irányzatba tartozik a pütkösdí mozgalom) életének fordulópontja előtti időszakra vonatkozik. Annak érdekében, hogy a tanúságtétel megmutassa az egyén bensőjét megváltoztató Isten mindenhatóságát, szuggesztív módon írja le a korábbi életszakaszokat. A vizsgált tanúságtételek által érintett területek lefedik a modern lengyel patológia térképét (főként alkoholizmus és kábítószerhasználat).

A kutatás eredményei 2.

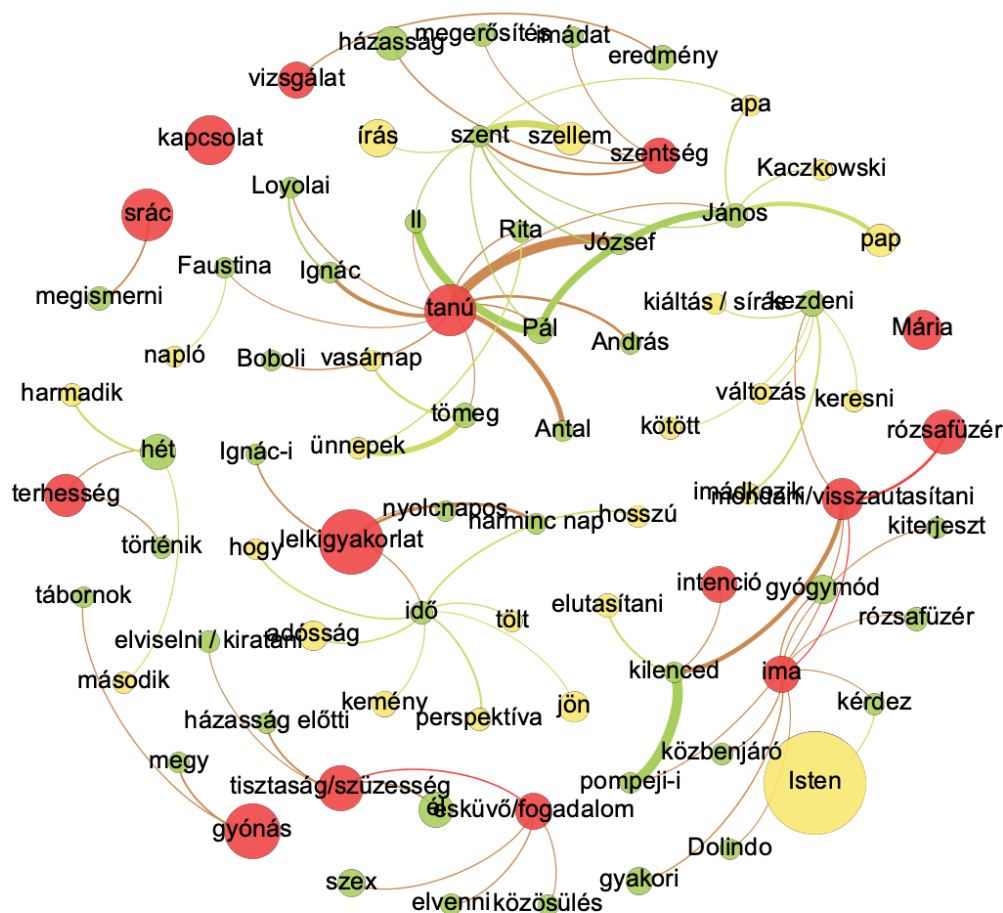
Kollokációk és kollokációs hálók

Az analízis következő szakaszai lehetővé tették új adatok begyűjtését. Az eddig megfogalmazott tételeket követve megvizsgáltuk a kulcsszavak kollokációját mindkét korpuszban – amint a módszertani fejezetben már kifejtettük, a nyelvi nyilatkozatok

²⁷ Ksawery Knotz i Krystyna Strączek, *Seks jest boski, czyli erotyka katolika* (Kraków: Znak, 2010).

kontextusa meghatározza a szavak jelentését, így az elemzésnek ez a fázisa elengedhetetlenül szükséges a kezdeti hipotézisek alátámasztásához.

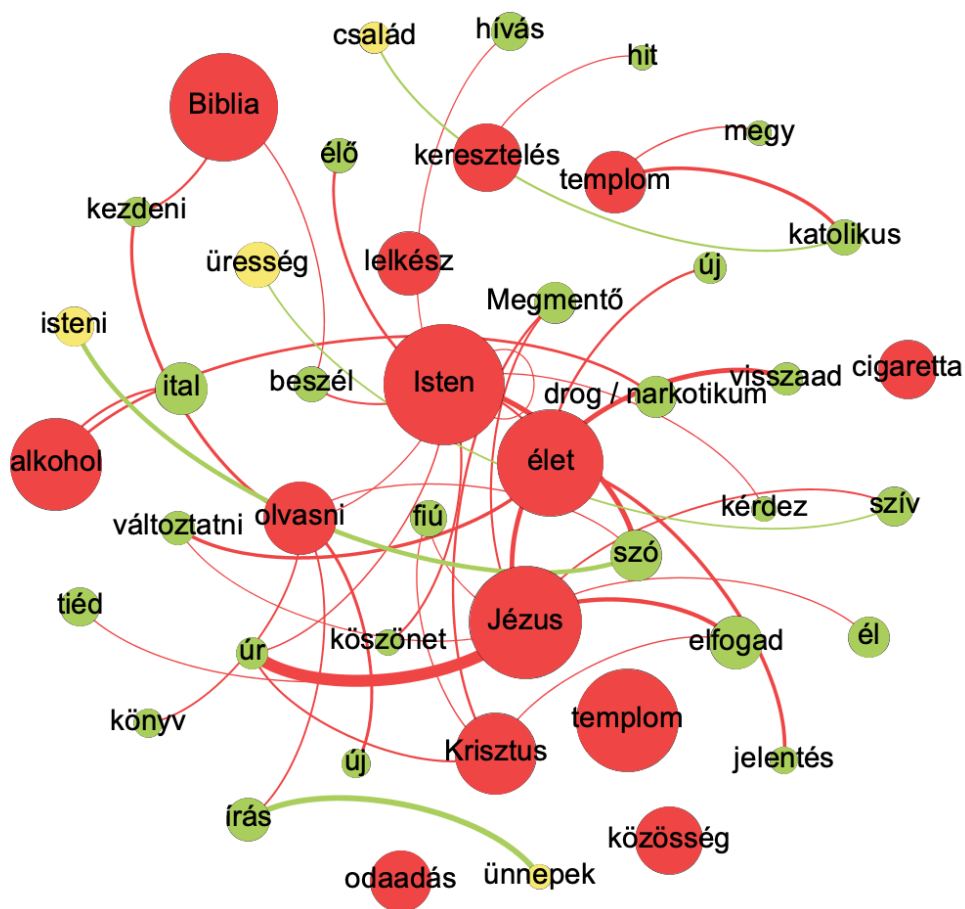
A digitális bölcsészet a vizualizációt mind a vizsgálat alkotóelemeként, mind az eredmények bemutatásának módszereként felhasználja. A továbbiakban bemutatott két háló és két gráf segítségével vizuálisan jelenítjük meg a korpuszok kulcsszavainak kollokációs rendszerét. A pontok mérete megfelel a kulcsszó diskurzív mértékének, amelyet az előfordulásoknak a korpusz méretéhez viszonyított száma alapján állapítottunk meg. Piros színnel a korpuszból kinyert kulcsszavakat jelöltük, zöld színnel ezek kollokációit, sárgával pedig a másodrendű kollokációs kapcsolatokat.



1. ábra. A katolikus korpusz 15 leggyakoribb kulcsszavának vizualizált kollokációs hálózata.

A vizualizált kollokációs háló láthatóvá teszi a protestáns tanúságtételek már említett szemantikai homogenitását (2. ábra). A protestáns korpuszt ábrázoló gráf feltárja, milyen domináns szerepet töltenek be a szemantikai kapcsolatok felépítésében az olyan fogalmak, mint *Bóg* 'Isten', *Jesus*, *Chrystus* 'Jézus Krisztus', *Biblia* 'Biblia', és közvetlenül utánuk: *społeczność* 'közösség' (a neoprotestáns mozgalmak egyik kulcsszava), *alkohol* 'alkohol', *czytać* 'olvasni', *zbor* 'gyülekezet'.

A katolikus kollokációs hálóból (1. ábra) a tanúságtételek nagyobb tematikai változatossága olvasható ki: ezek nem mutatnak olyan erős lexikális és strukturális egységességet, mint a protestáns szövegek. A gráf elemzése során látható, hogy olyan szavak foglalnak el központi helyet, mint: *świadek* 'tanú' (ez a *św.* – *święty* 'szent' – rövidítés hibás lemmatizálásának következménye), *rekolekcja* 'lelkigyakorlat', *różaniec* 'rózsafüzér', *Maryja* 'Mária', *sakrament* 'szentség', *spowiedź* 'gyónás', *modlitwa* 'ima'. Nem bizonyultak azonban kevésbé fontosnak a *chłopak* 'fiú barát', *związki* 'kapcsolat', *czystość* 'szüzesség', *cięża* 'terhesség', *ślub* 'esküvő' szavak sem. Figyeljük meg, hogy a *ślub* 'esküvő' szó az egyértelmű *wziąć* 'megkötni' kifejezésen kívül kapcsolatot alkot a *współzycie* 'együttélés' és *seks* 'szex' szavakkal is.



2. ábra. A protestáns korpusz 15 leggyakoribb kulcsszavának vizualizált kollokációs hálózata.

A katolikus tanúságok „hagyományos” olvasása, amelyet a „digitális” elemzéssel együtt végeztünk, még nagyobb változatosságot mutat a hívők életébe történő „isteni beavatkozás” leírásában. Az Istennel való kapcsolat elérésében hatékonynak bizonyult Galgani Szent Gemma (40 – a zárójelben itt és a továbbiakban a tanúságtételek sorszámja szerepel) vagy Szent József (38) imája, illetve a záródoklat Jasna Górára

(54). A katolikus korpuszban található olyan tanúságok is, amelyek struktúrájukban hasonlítanak a pünkösdistá tanúságtételekre – ezek az emberi életben történő radikális változásokról számolnak be és a protestánsokhoz hasonló szókinccset használnak, mint például *msza o uwolnienie* 'a szabadulásért [mondott] mise' (90), *osobisty egzorcyzm* 'személyes ördögűzés' (100). Ezek mellett feltűnnek olyan tanúságtételek is, mint például az, amelyikben a szerző visszaemlékezik gyermeke csodás megmenekülésére Bobola Szent András képének segítségével a varsói felkelés idején (118), vagy olyan tanúságtételek, amelyekben Isten segítségével megkerül az elveszett kulcsosomó.

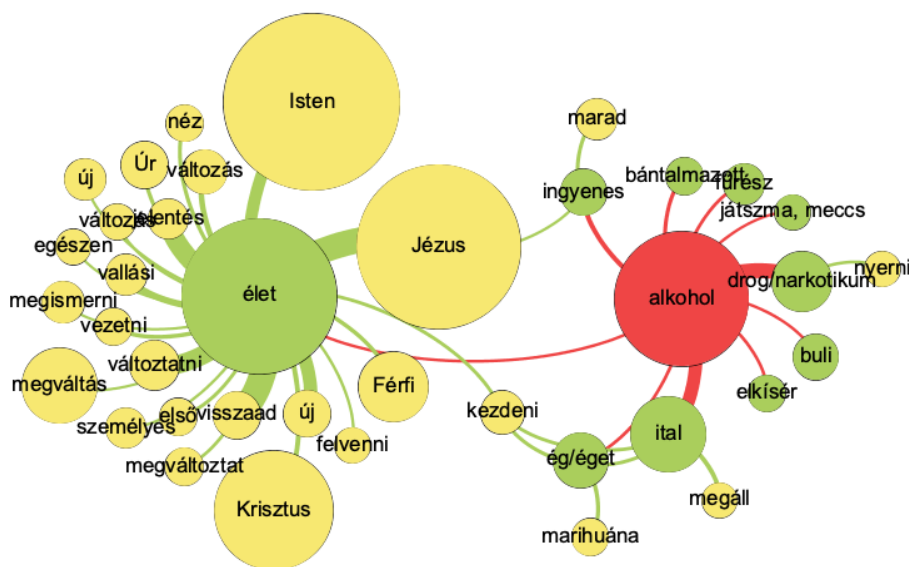
Ahogy korábban is említettük, statisztikai szempontból kiemelkedő csoportot alkotnak a tanúságtételek között a házasság előtti tisztaságról szólók, vagy azok, ahol az isteni beavatkozás segít az életre szóló társ megtalálásában és megtartásában (301 és mások). Figyeljük meg, hogy például a *cięża* 'terhesség' szó nagyon ritkán fordul elő a protestáns tanúságtételekben (*log-likelihood* értéke 10 vagy kevesebb), míg a katolikus korpuszban számos kollokációt alkot, a *być w ciąży* 'terhesnek lenni' pedig a korpusz egyik kulcskifejezése.



3. ábra. A *czystość* 'tisztaság, szüzesség' kulcsszó vizualizált kollokációs hálózata a katolikus korpusz alapján.

A katolikus korpuszban figyelemre méltó a két szóból álló *nowenna pompejańska* 'pompeji kilenced' kollokáció magas gyakorisága is. Ez az ima a Fortunata Agrelli által megtapasztalt Mária-jelenések hatására született 1884-ben, Olaszországban. A kilencedet Bartolo Longo domonkos harmadrendi szerzetes tette népszerűvé. Az ezzel kapcsolatos tanúságok elemzése kimutatja, hogy azok, akik ezt az imát mondják, mélyen hisznek annak hatékonyságában és azt „ellenállhatatlan”-ként írják le.

A következő szakaszban megvizsgáltuk a kulcsszavakat és felállítottuk a kollokációs hálózatokat, hogy verifikáljuk az eddigi következtetéseket. A jelen tanulmány céljaira kiválasztottunk két kifejezést: a *czystość* 'tisztaság, szüzesség' (katolikus korpusz) – 3. ábra – és az alkohol (protestáns korpusz) – 4. ábra – szavakat. Ezek vizualizációja azzal a céllal készült, hogy a szövegek mélyebb rétegeibe nyújtsanak bepillantást. A 3. ábrán például megfigyelhetjük a *czystość* 'tisztaság' erős kapcsolatát a *żyć z Bogiem* 'Istennel élni' frazémával, de ugyanez a szó erős kapcsolatot alkot a következőkkel is: *przedmałżeński* 'házasság előtti', *współżycie* 'együttélés', *seks* 'szex', valamint *dochować* 'megtartani', *walka* 'küzdelem', *trwanie/wytrwać* 'kitartás/kitartani', *zachować/zachowanie* 'fenntartani'. A katolikus diskurzusban lexikalizálódott *dochować* vagy *zachować czystość* 'megtartani a tisztaságot' fordulat hídként tűnik fel két erősen központi komponens, az 'Istennel való élet' és a házasság előtti szexuális együttélés között.



4. ábra. Az alkohol 'alkohol' kulcsszó vizualizált kollokációs hálózata a protestáns korpusz alapján.

A kollokációk elemzése rendkívül érdekes fogalmi metaforákat is feltár, amelyek funkciója a katolikusok házasság előtti szexuális önmegtartóztatásával kapcsolatos nehézségek belső tapasztalatának konceptualizálása. A korábban már említett (részben

várható) szóhasználat mellett, mint amilyen a *zachowanie czystości* 'a tisztaság megtartása', megjelennek kevésbé egyértelmű, de a megtett erőfeszítések mértékéről sokat eláruló, a KÜZDELEM forrástartományból merített metaforikus kifejezések is. A korpuszból olyan, ehhez a fogalmi tartományhoz tartozó mondatokat tudunk példaként kiemelni, mint: „Ne féljetez harcolni a tisztaságért, mert a várakozás ugyan keserű, de gyümölcssei annál édesebbek” (114), „[...] a fiatal nők átlagos életét élem. Minden nap küzdök a tisztaságért” (159), „Máig gondban vagyok ezzel, de mégis harcolok – sikeresen. Már régóta tartom magam a tisztaságban” (183), „Mi magunk küzdünk saját magunkkal, hogy minden ellenére ne lépjük át a határt” (123), „Több olyan emberre van szükség, aki akar harcolni a tisztaságért” (286) stb. A szexuális önmegtartóztatás szünet nélküli *küzdelemként* történő konceptualizációja egy folyamatos feszültséggel teli alak nyelvi képét rajzolja meg, akinek célja az, hogy uralja szexuális ösztönét az általa vallott értékekhez való hűség nevében, ennek érdekében pedig aktív tevékenységet fejt ki.

Az *alkohol* szónak a protestáns korpusz alapján vizualizált hálózata egyfelől bemutatja a kifejezés erős kapcsolódásait az élvezeti cikkek és patológia fogalmaival: *pić/piła* 'ital/ivott', *palić* 'dohányozni', *narkotyki* 'drog', *nadużywać* 'túlzott mértékben használni'. Másfelől megfigyelhető az erős kapcsolódás a *życie* 'élet' szóval és a *zmienić* és *odmienić* 'megváltoztatni', *zmiana* 'változás', *nowe* 'új' szemantikai szerkezettel, amely olyan főkulcsszavak köré épül (lásd a gráfon a pontok méretét), mint például *Bóg* 'Isten', *Jezus* 'Jézus', *Chrystus* 'Krisztus'. A diagram egyrészt ábrázolja a pünkösdi hívő Krisztus-központúságát (az élet változásának okozója Jézus), másrészt láthatóvá válnak rajta a gonosz és a fenyegetés fő területei is, amelyekkel a pünkösdi-karizmatikusok harcban állnak.

Összegzés

Vizsgálatunk lehetővé tette, hogy bepillantást nyerjünk a katolikus és pünkösdi vallási diskurzus egy szeletébe, a vallási tanúságtételekbe. A két korpusz közötti különbséget csak részben magyarázza, hogy azokat különböző médiumokból (internetes portálról és újságból) gyűjtöttük. A megnyilatkozók a két diskurzusban különböző módon közelítik meg magát a keresztény tanúságtételt. A pünkösdié számára a tanúság beszámoló az egész életet megváltoztató megtérésről, míg a katolikusok számára Istennek a mindennapi életben megnyilvánuló folyamatos segítségéről és jelenlétéről szóló elbeszélés. Az evangélikál életmodell drasztikus megtérést feltételez, a katolikus minta viszont a tökéletesedés irányába kifejtett állandó és fokozatos erőfeszítés modelljéhez áll közelebb. A tanúságtételek az ezeken a területeken kapott isteni segítség különféle formáiról számolnak be.

A különbség azt eredményezi, hogy a pünkösdi tanúságtételek sematikusabbak, kisebb szókinccset használnak, így az egyes kulcsszavak hangsúlyosabbak ebben az alkörpuszban. Az élethelyzetek leírása erősebben standardizált, ugyanis a megtérés előtti sötét, gonosz élet (mely leggyakrabban valamilyen függőséghez kapcsolódik) és a lelki fordulat hatására bekövetkező pozitív, gyökeres változás közötti kontrasztra összpontosítanak. A megtérés – a protestáns teológiai nézőpontnak megfelelően – kegyelmi ajándék, Isten adománya. A katolikus tanúságok lexikálisan jóval sokszí-

nőbbek, mert az isteni beavatkozás és segítség jóval változatosabb következményeiről számolnak be. Bár találhatunk kivételeket, fő szabályként mégis hiányzik belőlük az áttörésszerű lelki megtérés motívuma. A katolikus diskurzusban a hitet annak a folyamatos támogatásnak és segítségnek tulajdonítják, amit Isten nyújt az ember mindennapjaiban. A protestáns diskurzusban Isten úgyszólván „pontoszerűen” működik, és teljes mélységében megváltoztatja az egyén életét.

Elemzésünk feltárta, hogy a korpuszokat erősen megkülönböztető elem a gonosz problematikája. A katolikus szövegekben a gonosz szemantikai jelentésszerkezete a barát-barátnő, férfi-nő kapcsolatok kontextusában bukkan fel és nagyon gyakran érinti az erotika témáját. A szavak gyakoriságára vonatkozó statisztikai elemzés kimutatja a terület állandó jelenlétét a partneri kapcsolataikat építő fiatalok tudatában. Felállítható az a hipotézis, hogy a kapcsolatok szexuális (testi) dimenziójának hangsúlyosságát a fiatal katolikusokat feszítő dilemmák okozzák. Ezek a fiatalok szeretnének hűek maradni az egyház támasztotta követelményekhez (a házasság előtti szex tilalmához és a közösülést illető szabályozáshoz), egyúttal viszont a kortárs nyugati kultúra liberális szabályaival hasonlítják össze őket. A katolikus tanúságtételek a két értékrend konfliktusából származó feszültségről is tanúskodnak.

Háromfokozatú (a lexikálistól a kollokációkig terjedő) és több száz tanúságtételt figyelembe vevő analízisünkből látható, hogy a lengyel katolikusok életében az isteni jelenlét fő területe az erotika és az érzések szférájának „rendezése”. A protestánsok számára az isteni beavatkozás mindenekelőtt a függő ember hívő kereszténnyé változtatásában nyilvánul meg (a függőség palettája rendkívül széles, az alkoholtól a szexen át a kábítószerrekig terjed).

A fiatal katolikusokat kínzó lelkiismereti válságok elemzése rámutat, milyen dilemmákkal szembesülnek, miközben vallási identitásukat építik. A kutatás alátámasztja azt a feltevést, hogy az elemzett tanúságtételek nyelvi világa a helyes szexualitás problematikájára koncentrál, mintha katolikusnak lenni egyet jelentene bizonyos konkrét szabályokkal, amelyekben a moralitás elválaszthatatlanul összekapcsolódik a helyes módon folytatott szexuális élettel. A pütkösi diskurzusban a gonosz valójában a múlt része. A megtérés „megszabadít a gonosztól” és azt a múltba száműzi.

A jelen dolgozatban bemutatott kísérleti kutatás szemlélteti, milyen lehetőségeket kínál a bölcsészettudományok számára a kvantitatív és kvalitatív módszerek összekapcsolása, példánkban a korpuszalapú diskurzuselemzés. Megjeleníti, hogy a korpusznyelvészeti eszközök (amelyeket a digitális bölcsészet egyik elemeként kezelünk) hogyan lehetnek segítségünkre nagy terjedelmű szöveges adatok interpretálásában. Szándékunk szerint demonstráltuk, miképpen egészíthetik ki ezek az eszközök a modern vallásos életről és a lengyel vallásos kultúráról szóló ismereteinket.

Fordította: Zöldy Anna

Evangelical and Catholic Testimonies in the Perspective of Corpus-based Discourse Analysis

The article presents the results of a comparative analysis of Roman Catholic (from the portal Deon.pl) and Protestant, Pentecostal (from the magazine *Chrześcijanin* et al.) Christian testimonies. Using the tools of corpus linguistics, the authors show the differences between both collections of texts. Especially interesting was the reconstruction of the “areas of sin”, the evils the testimony authors overcame thanks to God’s help. These are different spheres for both types of texts, in the case of Catholic testimonies it is formed by a semantic complex built around the terms of “love – sex – asceticism”, and Pentecostal testimonies are focused on the “stimulants – addictions – antisocial behaviour” complex. The analysis consisting of two stages (frequency, describing the role of keywords in the corpus, and visual recognition of the most important collocations) allowed us to formulate conclusions on the dominant areas of meaning which define the mental map of the religious life of the followers of both the Christian denominations.

Keywords:

corpus linguistics, Christian testimonies, Polish religious culture, corpus-based discourse analysis