

PORTA LINGUA – 2019

**Interdiszciplináris megközelítések
a szaknyelvoktatásban
és -kutatásban**

**cikkek, tanulmányok
a hazai szaknyelvoktatásról
és -kutatásról**

**BUDAPEST
2019**

Főszerkesztő:
Bocz Zsuzsanna

Szerkesztő:
Besznyák Rita

Lektorok:
**Aradi András, Bocz Zsuzsanna, Demeter Éva, Einhorn Ágnes,
Fischer Márta, Keresztes Csilla, Kiszely Zoltán,
Kovátsné Loch Ágnes, Kurtán Zsuzsa, Rébék-Nagy Gábor,
Sárvári Judit, Seidl-Pécs Olívia, Sturcz Zoltán, Szabó Csilla**

Idegen nyelvi lektor:
Jon Marquette

Szerkesztőbizottság tagjai:
**Bánhegyi Mátyás, Besznyák Rita, Bocz Zsuzsanna, Fischer Márta,
Kurtán Zsuzsa, Lakatos-Báldy Zsuzsanna, Loch Ágnes,
Mátyás Judit, Sárvári Judit, Sturcz Zoltán**

Tanácsadó Testület tagjai:
**Einhorn Ágnes, Hamp Gábor, Hidasi Judit, Károly Krisztina,
Stephen Patrick**

ISSN 1785-2420

Kiadja:
Szaknyelvoktatók és -Kutatók Országos Egyesülete

Budapest
2019

Tartalomjegyzék

INTERKULTURÁLIS SZAKNYELVI KOMMUNIKÁCIÓ, SZAKFORDÍTÁS.....	7
BENE KRISZTIÁN: <i>Egy francia-magyar nyelvű történettudományi adatbázis létrehozásának szaknyelvi és fordítási nehézségei.....</i>	9
FAJT BALÁZS: <i>Aspects of improving 21st century skills in tertiary education: cognitive flexibility and complex problem solving.....</i>	19
HILD GABRIELLA – CSONGOR ALEXANDRA – NÉMETH TIMEA: <i>A magyar orvosi szaknyelv elsajátítását befolyásoló tényezők: négy mobil, norvég orvostanhallgató longitudinális esettanulmánya.....</i>	29
JÁMBOR EMŐKE: <i>A spanyol idiomatikus kifejezések jellege, használata: spanyol anyanyelvű hallgatókkal készített felmérés tükrében.....</i>	45
KÁRPÁTI LÁSZLÓ: <i>The use of communication strategies by Hungarian Teachers and Students.....</i>	52
KÓRIS RITA: <i>Három egyetem – két kontinens – egy tanterem: Szaknyelvi kompetenciák fejlesztése nemzetközi interkulturális, interdiszciplináris projekt keretében</i>	69
SZAKNYELV ÉS TERMINOLÓGIA.....	79
ÉLTHES ÁGNES: <i>Az autós lexika kialakulásának kezdetei Franciaországban: neologizmusok ma és tegnap Mirbeau: La 628-E8 című regénye alapján.....</i>	81
KERESZTES CSILLA – BORDA BERNADETT – VINCZE VERONIKA: <i>Mi derül ki a SZEMEK-ből? Egy magyar nyelvű orvos korpusz morfológiai sajátosságai.....</i>	97
MÉSZÁROS ÁGNES: <i>Önálló szakszókincsfejlesztés idegen nyelven.....</i>	111
NAGANO, ROBIN LEE: <i>How research is presented: A look at wh - structures in titles of journal articles.....</i>	123
RODEK BEGELLA ANNAMÁRIA: <i>Műszaki terminusok diverzitásának vizsgálata.....</i>	137
STURCZ ZOLTÁN: <i>Szakmai hagyományaink: az 1918-as őszirózsás forradalom nyelvpolitikai törekvései.....</i>	149
SZAKNYELVOKTATÁS, SZAKNYELVI TANTERV- ÉS TANANYAGFEJLESZTÉS.....	165
CSONGOR ALEXANDRA – NÉMETH TIMEA – HILD GABRIELLA: <i>Új kihívások az orvosi szaknyelvoktatásban.....</i>	167
HALÁSZ RENÁTA – RÉBÉK-NAGY GÁBOR: <i>Az orvostanhallgatói magyar szaknyelv oktatásáról</i>	175

JAKUSNÉ HARNOS ÉVA: <i>Having a Word. Writing a role play based on concept analysis</i>	187
KOPPÁN ÁGNES – KÁRPÁTI ESZTER – EKLICS KATA: <i>Pragmatikai megfontolások az orvosi szaknyelv oktatásában</i>	197
SKADRA MARGIT: „Like-olod?” – <i>Nyelvtanulást támogató online oldalak és azok hatékonysága a szaknyelvvoktatásban</i>	211
TÖRÖK JUDIT – UGRAI ZSUZSANNA: <i>Projektalapú tananyag a BGE KKK Német szaknyelvi csoportjaiban</i>	223
ZABÓNÉ VARGA IRÉN: <i>A műszaki szaknyelv oktatásának interdiszciplináris megközelítése</i>	235

ÁLTALÁNOS NYELVI ÉS SZAKNYELVI KOMPETENCIÁK MÉRÉSE; ÉRTÉKELÉS ÉS SZÜKSÉGLETELEMZÉS 245

BAJZÁT TÜNDE: <i>Idegennyelv-használati szükségletelemzés klinikai kutatási munkatársak (CRA) körében</i>	247
BÁNHÉGYI MÁTYÁS: <i>Alternatív értékelés a felsőoktatásban és a szaknyelvvoktatásban: a portfóliómódszer</i>	255
FÜKÖH BORBÁLA: <i>Kutatáson alapuló tesztfelállítás – írásfeladatok egy angol tudományos szaknyelvi vizsga számára</i>	275
HEGEDŰS ANITA: <i>A háttérismeret szerepe az orvosi szaknyelvvizsgákon</i>	289
KISZELY ZOLTÁN: <i>Online szóbeli vizsgáztató-továbbképzés a vizsgáztatók szemével</i>	299
LESZNYÁK MÁRTA: <i>Társak fordításainak elemzése mint fordítás-oktatási módszer</i>	311
NÉMETH TIMEA – HILD GABRIELLA – CSONGOR ALEXANDRA: <i>Szaknyelv-oktatás 2.0</i>	323

A TUDOMÁNYTERÜLETEK NYELVHASZNÁLATÁNAK SAJÁTOSÁGAI, DISKURZUSELEMZÉS 335

BÉRCES EDIT: <i>A multilingual conversation booklet for ultramarathon-runners. The Romanian perspective</i>	337
CSÁK ÉVA: <i>Pragmatikai kontrasztok a német és a magyar nyelvű írásbeli üzleti kommunikációban</i>	349
GYURÓ MONIKA: <i>Figuratív nyelvhasználat az egészségügyi kommunikációban</i>	361
HAMBUCH ANIKÓ – SÁRKÁNYNÉ LŐRINC ANITA – EGYED CSILLA – KRÁNICZ RITA: <i>Ki dönt? Közös döntéshozatal interakcionális megvalósulása orvos-beteg konzultációkban</i>	373
MÁTYÁS JUDIT: <i>Fiatalok reklámfogyasztási szokásai a magyar és idegen nyelvű reklámüzenetek relációjában</i>	391
NEUMAYER DÉNES: <i>Interactional and interactive discourse in physics</i>	

<i>coursebooks – a case study across languages</i>	403
RÁDI ÉVA: <i>A pénzügyi kultúra és a szaknyelvkutatás</i>	415
IN MEMORIAM	425
BAGAMÉRI ZSUZSANNA – NÉMETH ESZTER: <i>Egy példaértékű életút – Erdős József szakmai munkássága</i>	427
UGRIN ZSUZSA: <i>e-gyütt „a jóért”? – Egy önkéntes hallgatói fordítási csoportprojekt techno-lingvisztikai tanulságai</i>	439
ABSTRACTS	449

Keresztes Csilla¹ – Borda Bernadett² – Vincze Veronika³

^{1,2} Szegedi Tudományegyetem, ÁOK

³ MTA-SZTE Mesterséges Intelligencia Kutatócsoport

Mi derül ki a SZEMEK-ből?

Egy magyar nyelvű orvosi korpusz morfológiai sajátosságai

Négy magyar orvostudományi szaklap (Bőrgyógyászati és Venerológiai Szemle, Cardiologia Hungarica, Magyar Onkológia, Orvosi Hetilap) több mint 100, 2015 és 2018 között megjelent cikkét vizsgáltuk meg. A Szegedi Medical Korpusz (SZEMEK) több mint 12 ezer mondatot, körülbelül 256 ezer szót (token) tartalmaz. Eddigi vizsgálatunk során statisztikai és korpusznyelvészeti megközelítést alkalmaztunk, így számszerűsíthető megállapításokat tehetünk az írott magyar orvosi nyelvre vonatkozóan. Bemutatjuk a SZEMEK létrehozását és főbb adatait, ezután részletesen elemezzük a korpusz morfológiai jellemzőit, érintőlegesen tárgyaljuk a szintaktikai jellegzetességeit, majd összevetjük a köznyelvi korpusz (Szeged Korpusz – SZK) vonatkozó adataival. A morfológiai vizsgálat során azonosított főbb különbségek a két korpuszban a következők: a SZEMEK arányaiban lényegesen több nominalizációt (főnevet és melléknévet) tartalmaz, mint az SZK, a SZEMEK-ben több a számnév is, viszont kevesebb az ige, a névmás és a hely- és időhatározó. A vizsgált orvosi szaknyelvi korpusz lexikai elemzése során gyűjtöttük a leggyakrabban előforduló főneveket, igéket és melléknéveket, majd szófelhőre képeztük le ezeket. A szógyakorisági vizsgálat eredményeit is összevetettük a köznyelvi korpusz adataival. Rávilágítunk a gyűjtött adatok és a kapott eredmények hasznosíthatóságára a szaknyelvoktatásban és a szakfordításban.

Kulcsszavak: köznyelvi korpusz, morfológiai vizsgálat, orvosi szaknyelvi korpusz, szintaktikai vizsgálat, szógyakorisági elemzés

Bevezetés

A több mint 440 éves⁴ múlttal rendelkező magyar orvosi nyelv vizsgálata, elemzése kihívást jelent a nyelvészek, nyelvapolók, orvostörténészek számára. Az elmúlt 40 évben számos alkalmazott nyelvészeti (szociolingvisztikai, kontaktusnyelvészeti, korpusznyelvészeti, konverzációelemzés, stb.) kutatást végeztek magyar orvos- és egészségtudományi szövegeken (vö. Kontra, 1981; Keresztes, 2003; Bösze, 2004; Kránicz, 2014). A meglévő orvosi korpuszok (PTE, PPKE, MTA–NYTI, SZTE magyar orvosi nyelvi korpuszai) mellett, a kortárs

¹ Orvosi Szaknyelvi Kommunikációs és Fordítóképző Csoport

² Szegedi Tudományegyetem, ÁOK, Sebészeti Klinika

³ MTA-SZTE Mesterséges Intelligencia Kutatócsoport és Szegedi Tudományegyetem, Nyelvtechnológiai Csoport

⁴ Váradi Lencsés György (1530–1593) *Egész orvosságról való könyv, azaz Ars Medica* című munkája 1577 körül jelent meg.

kutatóorvosok írásbeli tudományos nyelvhasználatát kívántuk tanulmányozni.

Jelen írásunkban a korpuszelemzés első eredményeit mutatjuk be. Mivel az orvostudomány igen szerteágazó tudományág, vizsgálatunk három szakterület: a bőrgyógyászat, kardiológia és az onkológia írott szaknyelvének tanulmányozására vállalkozik. Így három magyar nyelvű orvosi szaklap, a *Bőrgyógyászati és Venerológiai Szemle*, a *Cardiologia Hungarica* és a *Magyar Onkológia* 2015 és 2018 között megjelent 77 cikkét (eredeti közlemény vagy esettanulmány) vontuk be a szövegelemzésbe. Negyedik szaklapként az *Orvosi Hetilap*ot választottuk a korpusz alapjául (33 cikk), hiszen ez a legrégebbi (1857-ben alapított), ma is kiadott magyar nyelvű orvosi szaklap. Az *Orvosi Hetilap* változatos profillal jelentet meg orvosi közleményeket, nem csak egy adott szakterületen belül.

Dolgozatunkban részletesebben elemezzük a korpusz morfológiai jellemzőit és érintőlegesen tárgyaljuk a szintaktikai jellegzetességeit, majd összevetjük egy köznyelvi korpusz (Szeged Korpusz – SZK) vonatkozó adataival. Rávilágítunk a gyűjtött adatok és a kapott eredmények hasznosíthatóságára a szaknyelvoktatásban és a szakfordításban, valamint a lehetséges további kutatásra és más kutatócsoportokkal való együttműködésre.

Módszerek

Az általunk gyűjtött és vizsgált korpusz magyar, egynyelvű, írott nyelvi szövegek gyűjteménye: teljes, autentikus szövegek, amelyek elektronikusan elérhetőek, olvashatóak és elemezhetőek. A 110 cikk reprezentatív (Lüdeling – Kytö, 2008) az adott három évre (2015-2018) vonatkoztatva. A korpusz építésekor igyekeztünk úgy gyűjteni az alkorpuszokat, hogy azok lehetőleg arányosak legyenek. Ezt azonban a cikkek terjedelme, illetve az évente megjelenő közlemények száma nem tette teljesen lehetővé (1. táblázat).

1. táblázat. A szaknyelvi korpusz (SZEMEK) és az alkorpuszok statisztikai adatai

(al)korpusz	tanulmány (n)	mondat (n)	szó (n)
bsz	37	3236	67191
ch	22	2307	46 366
mo	25	3438	74 037
oh	26	3164	68 470
SZEMEK összesen	110	12 145	256 064

Rövidítések: bsz: Bőrgyógyászati és Venerológiai Szemle; ch: Cardiologia Hungarica; mo: Magyar Onkológia; oh: Orvosi Hetilap; SZEMEK: Szegedi Medical Korpusz.

A korpusz eredetileg statikus, de az évente megjelentetett alacsony közleményszám miatt monitorrá, folyamatosan bővülővé, azonban az eredeti arányokat megtartóvá alakítottuk át. A szaknyelvi korpuszunk annotált, s ezt összevetettük egy köznyelvi korpuszsal, a Szeged Korpuszsal (SZK). Az SZK magyar természetes nyelvi szövegadatbázis, kézzel annotált magyar nyelvű szövegállomány, morfológiai és szintaktikai információkkal ellátva (Csendes et al., 2005). A teljes SZK 82 000 mondatot és 1,5 millió szövegszót tartalmaz 6 különféle szövegtípusból: üzleti rövidhírek, újságcikkek, iskolás fogalmazások, irodalmi szövegek, számítástechnikai szövegek és jogi szövegek. A két korpusz (SZEMEK és SZK) összevethetősége érdekében morfológiai szinten az ún. univerzális morfológia szófaji kódjait és morfológiai jegyeit használtuk, minden egyes szóhoz hozzárendeltük a szótövét (Vincze et al., 2017). Szintaktikai elemzéskor a *Szeged Dependencia Treebank* szintaktikai címkekészletét alkalmaztuk (Vincze et al., 2010).

Az adatrögzítés az írott nyelvi korpuszok esetében leginkább elterjedt *Text Encoding Initiative* (TEI) módszerrel történt. A nyelvészeti elemzés számára szükséges információkat az úgynevezett jelölő (*mark-up*) nyelv (*XML–Extensible Markup Language*) segítségével kódoltuk platform függetlenül a TEI nemzetközi szövegmódolási ajánlásának megfelelően. A jelölőelemek (*tagok*) szigorú szintaxisát szintén a TEI ajánlásai szabályozták. Automatikus elemzést végeztük, a *magyarlanc*⁵ szoftver (Zsibrita et al., 2013) segítségével mondatokra, majd szavakra bontottuk a szövegeket, majd morfológiai és függőségi (szintaktikai) elemzéssel láttuk el a mondatokat. A *magyarlanc* tisztán JAVA nyelvű modulokat tartalmaz (nincsenek más programozási nyelvre wrapperek), ez biztosítja a platformfüggetlenséget és a nagyobb rendszerekbe történő integrálhatóságot.

Kizárólag a fő szófaji információkat (minden token esetében csak a fő szófajt: főnév, ige, melléknév stb.) adtuk meg. A morfológiai jegyek részletes vizsgálatától eltekintettünk és az írásjeleket sem vizsgáltuk (Vincze et al., 2013).

⁵ A magyarlanc csomag moduljai a mondatra bontó (magyarra adaptált, szabályokkal kiterjesztett), szavakra bontó (magyarra adaptált), POS-tagger és lemmatizáló, a Stanford POS-tagger módosított változata, amely az ismeretlen szavakra a morfológiai elemző által adott lehetséges elemzéseket használja fel. A morfológiai elemző Gyepesi György véges állapotú automatákon alapuló kódja, amely a morphdb.hu erőforrásra épül. A morfológiai elemzés eredményét (KR kód) átkonvertáltuk MSD kódra. A modell a Szeged TreeBank-en tanult egy csökkentett kódhalmazon.

Eredmények

A korpusz szófaji elemzésekor megnéztük, hogy melyik szófaj hányszor jelenik meg a szövegekben (SZEMEK vs. SZK). A szófajok előfordulási gyakoriságát az 1. ábra mutatja. Követtük az univerzális morfológia jelölését, így az alábbi rövidítéseket alkalmaztuk:

ADJ – melléknév

ADV – határozószó

ADP – névutó

AUX – segédige

CONJ – mellérendelő kötőszó

DET – névelő

INTJ – indulatszó

NOUN – főnév

NUM – számnév

PART – igekötő

PRON – névmás

PROPN – tulajdonnév

PUNCT – központozás

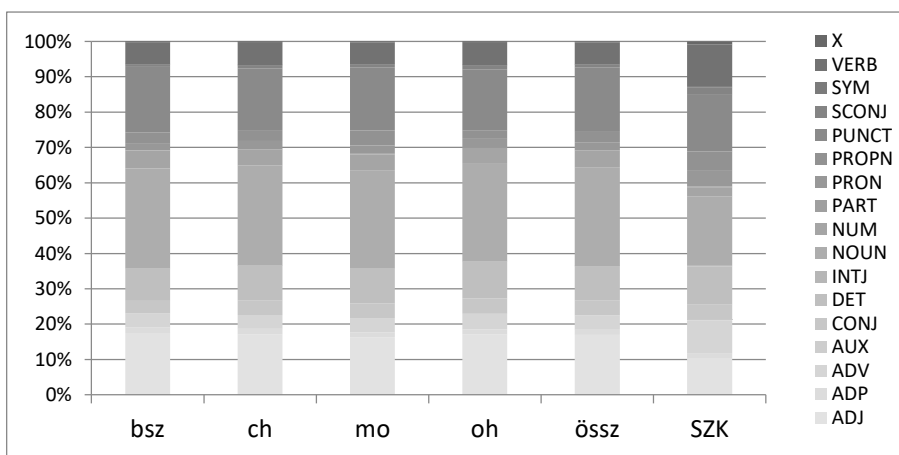
SCONJ – alárendelő kötőszó

SYM – szimbólum

VERB – ige

X – egyéb

1. ábra. A szófajok aránya a SZEMEK-ben, annak alkorpuszaiban és az SZK-ban



Rövidítések: bsz: Bőrgyógyászati és Venerológiai Szemle; ch: Cardiologia Hungarica; mo: Magyar Onkológia; oh: Orvosi Hetilap; össz: összesen a Szegedi Medical Korpuszban; SZK: Szeged Korpusz

Az egyes szaklapok között nem találtunk jelentős eltérést a morfológia terén, de a köznyelvel szemben (SZK) különbségek mutatkoztak. A szófajok gyakorisági eloszlását a 2. táblázat mutatja be.

2. táblázat. A SZEMEK-ben, alkorpuszaiban és az SZK-ban előforduló szófajok gyakorisága

	bsz (n=67 191)	ch (n=46 366)	mo (n=74 037)	oh (n=68 470)	Összesen SZEMEK (n=256 064)	SZK (n=1,5 millió)
ADJ	11 763 (17,5%)	7891 (17%)	12 034 (16,2%)	11 653 (17%)	43 341 (16,9%)	127 887 (8,5%)
NOUN	19 075 (28,4%)	13 102 (28,2%)	20 450 (27,6%)	19 025 (27,8%)	71 652 (28%)	240 456 (16%)
NUM	3385 (5%)	2065 (4,4%)	3376 (4,5%)	2927 (4,3%)	11 753 (4,6%)	31 641 (2,1%)
PRON	1165 (1,7%)	1088 (2,3%)	1718 (2,3%)	1715 (2,5%)	5686 (2,2%)	57 716 (3,8%)
VERB	4095 (6,1%)	3024 (6,5%)	4558 (6,1%)	4474 (6,5%)	16 151 (6,3%)	148 068 (9,9%)

Rövidítések: bsz: Bőrgyógyászati és Venerológiai Szemle; ch: Cardiologia Hungarica; mo: Magyar Onkológia; oh: Orvosi Hetilap; össz: összesen a Szegedi Medical Korpuszban; SZK: Szeged Korpusz; ADJ: melléknév; NOUN: főnév; NUM: számnév; PRON: névmás; VERB: ige.

Ezután részletesen elemeztük a főbb szófajokat (főnév, melléknév és ige), kigyűjtöttük az alkorpuszokból a leggyakoribbakat, majd szófelhőre képeztük le a fő korpuszban megjelenő adott szófajú tokeneket. Az elemzésnél mindig a szótövet vettük figyelembe. Mind a négy szaklap esetén a *beteg*, *eset* és *vizsgálat* főnevek fordultak elő a leggyakrabban, s mind a négy lapban a *kezelés* szó is az első kilenc leggyakoribb főnév közé tartozott (3. táblázat).

3. táblázat. A SZEMEK alkorpuszaiban megjelenő leggyakoribb főnevek

bsz alkorpusz főnevek	előfordulás (n)	ch alkorpusz főnevek	előfordulás (n)
eset	532	beteg	486
beteg	526	vizsgálat	271
vizsgálat	438	eset	256
kezelés	391	kezelés	166
sor	193	sor	152
tünet	177	eredmény	136
betegség	147	csoport	132
bőr	147	hatás	124
ábra	143	mg	116
terápia	141	ábra	111

eredmény	138	terápia	104
alap	131	idő	86
év	128	szívelégtelenség	69
bőrtünet	125	vérnyomás	54

mo alkorpusz főnevek	előfordulás (n)	oh alkorpusz főnevek	előfordulás (n)
eset	472	beteg	584
vizsgálat	376	eset	361
beteg	357	vizsgálat	280
tumor	334	eredmény	230
sejt	234	sor	192
kezelés	218	alap	187
eredmény	175	év	173
daganat	172	adat	166
alap	169	kezelés	146
minta	168	állapot	113
módszer	141	modell	104
sor	132	betegség	102

Rövidítések: bsz: Bőrgyógyászati és Venerológiai Szemle; ch: Cardiologia Hungarica; mo: Magyar Onkológia; oh: Orvosi Hetilap

Mind a négy szaknyelvi korpuszban a létige (*lesz* és *van*) fordult elő leggyakrabban, emellett főként történet kifejező igéket azonosítottunk magasabb arányban (*kerül, történik*) (4. táblázat).

4. táblázat. A SZEMEK alkorpuszaiban megjelenő igék gyakorisága

Alkorpusz:	Igék	Előfordulás (n)
bsz	van	398
	mutat	119
	igazol	99
	történik	99
	lesz	84
	észlel	82
	jelentkezik	80
	kerül	61
ch	van	340
	mutat	73
	csökken	72
	történik	67
	végez	67

	lesz	66
	vizsgál	54
	kerül	52
mo	van	404
	lesz	151
	mutat	135
	történik	87
oh	van	434
	lesz	102
	kerül	90
	mutat	89

Rövidítések: bsz: Bőrgyógyászati és Venerológiai Szemle; ch: Cardiologia Hungarica;
mo: Magyar Onkológia; oh: Orvosi Hetilap

Kigyűjtöttük a szaknyelvi korpuszból (SZEMEK) a leggyakoribb főneveket, igéket és mellékneveket, majd szófelhőre képeztük le ezeket. Ezután összevetettük a köznyelvi korpusz adott szófelhőivel (2-7. ábra).

2. ábra. A leggyakoribb főnevek az összesített orvosi szaknyelvi korpuszban (SZEMEK)



Az összesített szaknyelvi korpuszban¹ a *klinikai* és a *nagy* melléknevek fordultak elő leggyakrabban, a köznyelvi korpuszban a *nagy*, *első* és *jó* melléknevek.

6. ábra A leggyakoribb melléknevek az összesített szaknyelvi korpuszban (SZEMEK)



7. ábra A leggyakoribb melléknevek a köznyelvi korpuszban (SZK)



¹ A melléknevek részletes adatait a tanulmány terjedelmi korlátai miatt nem közöljük.

Megbeszélés

A szaknyelv szembevető lexikai sajátosságai a szakszavak és a terminusok (Szépe, 1982). A szaknyelvi korpusz reprezentativitásának köszönhetően információt szolgáltathat egy-egy szakszó, terminus műfajra és regiszterre jellemző használatáról (Seidl-Péché, 2018). A morfológiai vizsgálat során azonosított főbb különbségek a szaknyelvi és a köznyelvi korpuszban a következők voltak: a SZEMEK arányaiban lényegesen több nominalizációt (főnevet és melléknevet, 44,9%) tartalmaz, mint az SZK (24,5%), de kevesebb igét (SZEMEK: 6,3%, SZK: 9,9%).

A vizsgálatból nyert adatok (2. táblázat) az orvosi szaknyelvi szövegek nominális jellegét mutatják a magyar nyelvben (vö. angol orvosi nyelv, Halliday – Martin, 1993; Gao, 2012). Az igei stílus névszói stílussal való helyettesítése, a grammatikai metafora alkalmazása (vö. Halliday – Martin, 1993), a nominalizáció, valamint a személytelen szerkezetek általában jellemzőek a szaknyelvekre (Kurtán, 2003; Keresztes, 2014). A nominalizáció fontos szerepet játszik az orvosi cikkek logikai szerkezetének felépítésében, növeli a formalitást (Gao, 2012). A főnévi szócsoportok, tagmondatok teszik lehetővé a szöveg argumentációjának kifejezését (Keresztes, 2014). A tudományos gondolkodás és ezzel összefüggésben a tudományos nyelv fejlődésében megfigyelhető, hogy a nominalizáció jelentős szereppel bír a tudományos terminológia kialakításában (Kuna, 2011).

A nominális stílus legfőbb jellemzője „a névszók egyeduralma” (Szikszai 2006: 480). A szaknyelv az igéket főnevekké alakítva, az absztrakt entitások megnevezésére, szakszavakat hoz létre szóképzéssel. Igen sok igéből *-ás*, *-és*, *-at*, *-et* képzővel képzett főnevet (pl. *hat*–*hatás*, *kezel*–*kezelés*, *vizsgál*–*vizsgálat*) azonosítottunk a SZEMEK-ben.

A SZEMEK alkorpuszaiban leggyakrabban előforduló főnevek többsége az egészségügyet, gyógykezelést, kutatást (eredeti közlemények!) leíró főnév (pl. *beteg*, *eredmény*, *eset*, *kezelés*, *kérdőív*, *minta*, *vizsgálat*). Gyakoriak az idő szemantikai mezőbe tartozó szavak is (pl. *nap*, *év*, *idő*). Emellett jellemzőek az alkorpusz-specifikus görög-latin eredetű orvosi szakszavak (pl. *hipertónia*, *koronária*, *melanóma*, *proszata*, *stádium*, *tumor*). Az alkorpusz-specifikus nem görög-latin eredetű főnevek többsége összetett szó (pl. *bőrtünet*, *sejtvonal*, *sugárkezelés*, *vérnyomás*). Összességében a főnevekről elmondható, hogy a SZEMEK-ben jellemző a köznyelvinél hosszabb szavak használata.

A köznyelvi korpuszban is gyakoriak az időre utaló főnevek (pl. *nap*, *óra*, *év*, *idő*, illetve a hét napjai), azonban ezen kívül nem találtunk azonosságot a SZEMEK és az SZK főnévi készlete között. Az SZK

leggyakoribb főnevei általános emberi kapcsolatokat, tevékenységeket jelölnek (pl. *ember, élet, iskola*), az üzleti étellel (pl. *cég, forint, részvény, százalék, tőzsde*) vagy az informatikával kapcsolatosak (pl. *állomány, felhasználó, hálózat, lemez, számítógép*).

A szaknyelvekben a nominalizáció térhódítása mellett az igék szerepének csökkenése jellemző (Kurtán, 2006). Ennek a megállapításnak megfelel, hogy a SZEMEK-ben kevesebb az ige (6,3%), mint az SZK-ban (9,9%). A szaknyelvi korpuszban azonosított leggyakoribb igék a létezést kifejező igék (*lesz, van*). A többi ige is inkább tényeket és állapotokat ír le, történést fejez ki, semmint eseményeket (pl. *csökken, emelkedik, kerül, történik*). A további igék a kutatáshoz vagy az orvosi vizsgálathoz kapcsolódnak (*alkalmaz, észlel, igazol, végez, vizsgál*). Általánosságban megállapítható, hogy a tudományos orvosi szaknyelv bizonyos szintaktikai szerkezetek használatát részesíti előnyben: az ige gyakran kopula funkciót tölt be (pl. *diagnózist igazol, emelkedést mutat, vizsgálatra kerül*) (vö. Ablonczyné, 2002, 2006; Kurtán 2006; példák – SZEMEK). Hasonlóságot mutat, hogy a köznyelvi korpuszban is a létezést jelentő igék (*lesz, van*) fordulnak elő leggyakrabban, emellett azonban gyakori a kötelezettség kifejeződése (*kell*), valamint a képességet, szándékot kifejező ige (pl. *tud, szeret, akar*).

A SZEMEK-ben több a melléknév (16,9% vs. SZK: 8,5%) és a számnév is (4,6% vs. SZK: 2,1%)². A melléknevek közül az összesített szaknyelvi korpuszban a *klinikai* és a *nagy* szavak fordultak elő leggyakrabban, a köznyelvi korpuszban a *nagy* melléknév gyakorisága volt a legmagasabb, emellett az *első* és *jó* melléknevek fordultak elő leggyakrabban. Eddig nem végeztük még el a *nagy* melléknév szöveggörnyezetének vizsgálatát a SZEMEK-ben, de valószínűsítjük, hogy míg a köznyelvben ez a melléknév gyakran a jelölt főnév méretére utal, a szaknyelvi korpuszban valószínűleg a fontosság/jelentőség tartományban profilál (Papp, 2008). A számneveknek fontos szerepük van a tények alátámasztásában, a pontos adatok közlése fontos jellemzője a tudományos cikkeknek, elősegítik az objektivitást. A SZEMEK-ben arányaiban lényegesen több számnevet azonosítottunk (4,6% vs. SZK: 2,1%).

A SZEMEK-ben kevesebb a névmás (2,2%) mint a köznyelvi korpuszban (3,8%): a tudományos orvosi szövegek a pontosabb megfogalmazás érdekében kerülnek a névmási visszautalásokat, azok esetleges többértelműségét, helyette más módszereket alkalmaznak a korreferencia kifejezésére (ismétlés, helyettesítés).

² A SZEMEK-ben azonosított melléknevek és számnevek részletes elemzésére ebben a tanulmányban nem kerül sor.

A SZEMEK-ben szintén kevesebb a határozószó (hely és idő) (SZEMEK: 4% vs. SZK: 7,7%) és a főnévi igenév (SZEMEK: 0,2% vs. 1%) is. Ez magyarázható azzal, hogy a határozószavak igét módosítanak, a főnévi igenév az ige személytelen alakja (képzős igei származék), igéből viszont kevesebb szerepel a szaknyelvi korpuszban.

Jelen tanulmányban nem térünk ki a szaknyelvi korpusz szövegeinek részletes szintaktikai elemzésére, azonban összességében állíthatjuk, hogy a SZEMEK több alá- és mellérendelést, valamint közbevetést tartalmaz, mint az SZK. Az orvosi tanulmányok mondatszerkesztése bonyolultabb, mint a köznyelvi szövegeké.

Konklúzió

A SZEMEK annotált szaknyelvi korpusz nyelvtechnológiai feladatokban elméleti és alkalmazott nyelvészeti kutatásokban egyaránt alkalmazható tesztelési és fejlesztési célokra is. Elősegíti a kortárs, írott magyar orvosi nyelv mélyebb elemzését, rendszeres leírását (Kuna et al., 2017). Mivel elektronikus korpusz, bővíthető vagy szükség esetén javítható. Az adott szövegtípusok további részletes vizsgálatára alkalmas, így beilleszthető a hazai orvosi szaknyelvi szövegek kutatásába.

A szakfordítói munkához ma már elengedhetetlenek a fordítási környezetet biztosító szoftverek, a fordítástámogató eszközök. A korpusz továbbfejleszhető két vagy többnyelvű glosszáriummá, illetve megfelelő szakértők bevonásával akár elektronikus szótárrá (vö. Seidl-Péché – Pálinkás, 2015).

A SZEMEK lehetővé teszi a csatlakozást más nemzeti orvosi nyelvek vizsgálatához: nemzetközi összehasonlító korpuszvizsgálatokba való bekapcsolódást az orvosi szaknyelv területén.

Hivatkozások

- Ablonczyné, M. L. (2002): A gazdasági szaknyelv sajátosságai. In: Földes, Cs. (szerk.) (2002): *MMI Annum tempus Linguarum Europae. Scripta Philologica Pannoniensis*. Veszprém
- Ablonczyné, M. L. (2006): *Gazdaság és nyelv*. Lexikográfia Kiadó: Pécs
- Bősze, P. (2004). Gondolatok az orvosi nyelv magyarításáról. In: Balázs, G. (szerk.): *A magyar nyelvi kultúra jelene és jövője*. Magyar Tudományos Akadémia: Budapest
- Csendes, D. – Csirik, J. – Gyimóthy, T. – Kocsor, A. (2005): The Szeged Treebank. In: Matoušek, V. et al. (eds.): *Proceedings of the 8th International Conference on Text, Speech and Dialogue (TSD 2005)*. Springer: Karlovy Vary
- Gao, W. (2012): Nominalization in Medical Papers: A Comparative Study. *Studies in Literature and Language*. 4 (1). 86–93

- Halliday, M.A.K. – Martin, J.R. (1993): *Writing science: Literacy and discursive power*. The Falmer Press: London/Washington
- Keresztes, Cs. (2003): Az angol nyelv hatása a magyar szakmai nyelvre. In: Tóth, Sz. (szerk.). *Nyelvek és kultúrák találkozása*. Officina Press Kft.: Szeged
- Keresztes, Cs. (2014): A magyar orvosi nyelv napjainkban. In: Demeter, É. – Nagy, T. (szerk.) *Ünnepi kötet Rébék-Nagy Gábor 60. születésnapjára*. PTE ÁOK: Pécs
- Kontra, M. (1981): *A nyelvek közötti kölcsönzés néhány kérdéséről, különös tekintettel „elangolosodó” orvosi nyelvünkre*. Akadémiai Kiadó: Budapest
- Kráncz, R. (2014): Kórházi és hagyományos tanórai interakciók összehasonlító elemzése. *Anyanyelv-pedagógia*. 7 (4). , <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=>
- Kuna, Á. (2016): Person deixis and self-representation in medical discourse. *Jezyk Komunikacja Informacja / Language Communication Information*. 11. 99–121
- Kuna, Á. – Kocsis, Zs. – Ludányi, Zs. (2017): A Magyar orvosi nyelvi korpusz 16-17. századi alkorpusza: Tervezet, átírás, annotálás. In: Forgács, T. – Németh, M. – Sinkovics, B. (szerk.): *A nyelvtörténeti kutatások újabb eredményei IX*. Szegedi Tudományegyetem: Szeged
- Kuna, Á. (2011): *A 16–17. századi magyar nyelvű orvosi recept szövegtipológiai és pragmatikai vizsgálata funkcionális kognitív keretben*. Doktori disszertáció. ELTE: Budapest
- Kurtán, Zs. (2003): *Szakmai nyelvhasználat*. Nemzeti Tankönyvkiadó: Budapest
- Kurtán, Zs. (2006): Szaknyelv. In: Kiefer, F. (szerk.) *Magyar nyelv*. Akadémiai Kiadó: Budapest
- Lüdeling, A. – Kytö, M. (eds.) (2008): *Corpus Linguistics An International Handbook*. Volume 1. Walter de Gruyter: Berlin–New York
- Papp, K. (2008): Tartomány-specifikáció és hatókörszűkítés a nem tipikus melléknévi módosításokban. In: Gecső, T. – Sárdi, Cs. (szerk.): *Jel és jelentés*. Tinta Könyvkiadó: Budapest
- Seidl-Péché, O. (2018): Melyek a (szak)fordító és a fordításkutató munkáját segítő legfontosabb nyelvi korpuszok? In: Robin, E. – Zachar, V. (szerk.): *Fordítástudomány ma és holnap*. L'Harmattan Kiadó: Budapest
- Seidl-Péché O. – Pálincás, M. (2015): Lépést tud-e tartani a műszaki szaklexikográfia a technikai fejlődéssel? In: Bocz, Zs. (szerk.): *Porta Lingua – 2015: A XXI. századi szakmai, szaknyelvi kommunikáció kihívásai: tanári és tanulói kompetenciák*. Szaknyelvtanítók és -Kutatók Országos Egyesülete (SZOKOE): Budapest
- Szépe, Gy. (1982): A szaknyelv és a mindennapi nyelv kapcsolata. *A technika tanítása*. 5. 129–139
- Szikszaíné, N. I. (2006): *Leíró magyar szövegtan*. Osiris Kiadó: Budapest
- Vincze, V. – Simkó, K. I. – Szántó, Zs. – Farkas, R. (2017): Universal Dependencies and Morphology for Hungarian – and on the Price of Universality. In: *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2017) – Volume 1*. Valencia
- Vincze, V. et al. (2010): Hungarian Dependency Treebank. In: *Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC'10)*. Valletta, Malta
- Vincze, V. (2013): Domain Differences in the Distribution of Parts of Speech and Dependency Relations in Hungarian. *Journal of Quantitative Linguistics*. 20 (4). 314–338
- Zsibrita, J. – Vincze, V. – Farkas, R. (2013): ‘magyarlanc’: A toolkit for morphological and dependency parsing of Hungarian. In: *Proceedings of RANLP*