

A szógyakoriság gépi vizsgálata műszaki szövegeken

Nagy Annamária Lilla

E-mail: annamarialillanagy@gmail.com

Kivonat: A tanulmány a műszaki szövegek szógyakoriságának jellemzőit és az ezekhez vezető okokat igyekszik feltárni, továbbá arra a kérdésre keresi a választ, milyen hatást gyakorolnak a műszaki szövegek szógyakoriságára a terminusok, és vajon a műszaki szövegekben is megvalósul-e az ismétléskerülési hipotézis, vagy a terminushasználat kötelező jellege megakadályozza egyrészt az autentikus szöveg íróját, majd az idegen nyelvre történő átültetéskor a fordítót is a rokon értelmű szavak és kifejezések használatában. A témában eddig viszonylag kevés kutatás született. A tanulmányban leírt vizsgálatokat a szerző gépi elemzéssel végezte párhuzamos, angol–magyar fordítástudományi korpuszon, majd összeveti Robin (2014 és 2015) eredményeivel, aki szórakoztatóirodalmi szövegeken végzett hasonló, a lexikai sűrűséget és a szógyakoriságot vizsgáló kutatást. Az így kapott eredmények a műszaki szövegek szógyakoriságának sajátosságairól adnak átfogóbb képet.

Kulcsszavak: műfaji sajátosságok, műszaki szöveg, Pannónia Korpusz, szógyakoriság, terminus

1. Bevezetés

Az egyre nagyobb népszerűségnek örvendő korpuszalapú kutatások között gyakran képezi vizsgáldás tárgyát a szógyakoriság vizsgálata, azonban az elemzésre szánt szövegek, feltehetőleg nagy terjedelműeknek köszönhetően, elsősorban a szépirodalmi művek soraiból kerülnek ki. Nem kérdés azonban, hogy az ilyen jellegű kutatások rendkívül érdekes eredményekre vezetnek más műfajok, illetve szövegtípusok esetében is.

A műszaki szövegek vizsgálatai általában azok fordítási problémáit, problematikáit helyezik középpontba (ld. Sturcz 2010), vagy a más nyelvre történő átültetésükhöz szükséges speciális fordítói kompetencia kérdéskörét vetik fel (ld. Fischer 2012). Felmerül azonban a kérdés, vajon milyen eredményt hozhat, ha nem a már jól ismert kutatási területek ismereteit bővítjük, hanem megvizsgáljuk és összevetjük, milyen arányú szógyakoriság fedezhető fel a forrásnyelvi és a célnyelvi műszaki szövegekben, illetve, milyen mértékben tér el a szógyakoriságuk az irodalmi szövegekhez képest – Robin 2014 vizsgálati eredményei

alapján. Szükséges továbbá annak feltárása is, mennyiben térnek el a fordított szövegek eredményei az autentikus változatokétól.

2. A műszaki szövegek jellegzetességei

A műszaki szövegek létrejötte és ezzel együtt a fordításuk is a folyamatosan fejlődő világ és technika termékei. A műszaki szövegek említésekor az első kérdések egyike mindig a fordító szakmai ismeretére, illetve szaknyelvi tudására vonatkozik. Az elterjedt szemlélet szerint az ilyen jellegű szövegek elfogadható minőségű fordításához nem elegendő csupán az idegen nyelv megfelelő szintű ismerete, hanem az adott műszaki területen, vagy annak tudományában való jártasság is elengedhetetlen. Hans-R. Fluck is említi egyik szaknyelvoktatással foglalkozó könyvében a szaktudás, a nyelvi és szakmai ismeretek ideális esetben meglévő egységét, amely azonban a valóságban csak nagyon ritkán fordul elő (Fluck 1992: 221 idézi Zabóné 2015: 27).

Sturcz 2010-es írásában olvashatjuk, hogy a műszaki szaknyelvhez a nyelv egésze felől indulva legjobban a Grétsy-féle „virágmotívum” alapján közelíthetünk, amely az egyes nyelvi rétegeket egymásba kapcsolódó virágszirmokként mutatja be. Ebben az elgondolásban a szaknyelv „szírom” csak főegység, amely több elemből épül fel. Ezek egyike a műszaki szaknyelv. Másként fogalmazva tehát a műszaki nyelv csupán a szaknyelvek halmazának egy kisebb csoportja további, hosszan sorolható alcsoportok mellett, mint pl. gazdasági szaknyelv, jogi szaknyelv, orvosi szaknyelv, tudományos szaknyelvek stb. Sturcz ebben a tanulmányában rámutat továbbá arra, hogy a különböző csoportosítási lehetőségek és az ebből fakadó viták következtében a szaknyelvi halmaz elemsora nehezen meghatározható és körülményesen behatárolható.

A német nyelvterületen Stolze (2005: 84) öt pontban foglalta össze a természettudományos-műszaki műfajra jellemző fordítási szabályszerűségeket, amelyeket feltétlenül érdemes figyelembe venni az ilyen szövegeken végzett kutatások és elemzések során:

1. A fordítás elsődleges célja a tartalmi ekvivalenciára való törekvés
2. Az ilyen szövegek szintaxisa korlátozott mennyiségű mintát követ, és minimális a variációs lehetőségek száma.
3. A terminusok használata miatt a lexikában a variációs lehetőségek csökkenése figyelhető meg.
4. A stilisztikában állandósult felhasználási normákat követnek, egyedi variációkra így csak korlátozottan van lehetőség.
5. A célnyelvi olvasók háttérismerete és szaktudása hasonló, vagy megegyezik a forrásnyelvi célközönségével, így alapvető probléma nem merül fel.¹

¹ *A tanulmány szerzőjének fordítása.*

A szakszövegek egy további jellemvonása, hogy a szövegek úgynevezett „puha” és „kemény” részekre oszlanak (Klaudy 2007). A „kemény” részek azok, ahol a fordítónak a szöveg célnyelvre történő átültetése során nincs választási lehetősége több különböző megoldás között, itt valójában egy adott lexikai egység behelyettesítéséről van szó. A szöveg „kemény” részei közé soroljuk többek között a terminus értékű kifejezéseket. A szöveg „puha” vagy más szóval „rugalmas” részei pedig azok, amelyeknél a fordító több lehetséges megoldás közül választhat, így a szöveg végső változata a fordítói döntés végeredményeként különböző fordítók esetében eltérő lehet, ellentétben a kötött „kemény” elemekkel.

2.1. Fordítói műveletek

A fordítási feladatok során a nyelvek eltérő jellegéből fakadóan különböző műveleteket kell elvégezni, hogy a célnyelvi szöveg ugyanazt a helyet és szerepet töltsen be a célnyelvi rendszerben, mint a forrásnyelvi szöveg az adott, másik nyelvi rendszerben. Az idők során számos kutatás és tudományos értekezés helyezte középpontba a fordítói műveleteket és igyekezett azok univerzális vagy fakultatív voltára rámutatni. Ezek közül úttörőnek számít Mona Baker (1993) tanulmánya. A szerző a fordítási viselkedés alapelveként tekint az univerzálákra, amelyeket tipikusan a fordítási szövegek sajátosságaként jelöl meg. A fordítási univerzálék az autentikus szövegekre nem jellemzőek, azonban nem a nyelvszisztemek interferenciájaként jönnek létre, hanem a fordítási folyamat során felmerülő kényszerhelyzetek hatására (Baker 1993).

Az egyik általánosan elfogadott és elsőként felfedezett fordítási univerzálék az explicitáció és az implicitáció párosa. Az explicitáció fogalmát Vinay és Darbelnet (1958) határozta meg először, a gondolatmenet kiegészítésére és a jelenség okainak feltárására később számtalan kutató végzett elemzéseket, többek között Blum-Kulka (1986), Toury (1995) Olohan és Baker (2000) és Klaudy és Károly (2005). Explicitációnak azt a jelenséget nevezzük, amikor a fordító egyértelműbben fogalmaz meg valamit a célnyelvi szövegben, amely a forrásnyelvi szövegben implicit maradt a kontextusból való meghatározhatóság miatt. Klaudy (1999) tipológiájában rávilágít, hogy az explicitáció nem csupán tudatos fordítói döntés lehet, hanem a nyelvek különbözőségéből fakadó elkerülhetetlen művelet is. Ennek megfelelően négy fő kategóriát azonosít: kötelező, fakultatív, pragmatikai és fordításspecifikus. Klaudy tipológiájára támaszkodva Robin (2015) értekezésében az alábbi felosztást javasolja: szabálykövető, normakövető, valamint szerkesztési stratégiai.

Az explicitáció és az implicitáció használatával kapcsolatban az ismétlések kerülését is szükséges megemlíteni. Az ismétléskerülés témájának Toury (1995) nagy jelentőséget tulajdonított, és kutatásában arra a következtetésre jutott, hogy a fordítások során a fordítók a lexikai ismétléseket jelentősen csökkentik. Klaudy és Károly (2005) hivatásos fordítókon és hallgatókon végzett vizsgálata pedig azt az eredményt hozta, hogy a gyakorlott fordítók nem változtatnak jelentősen az ismétléskapcsolatokon, a kezdő fordítók azonban igen.

Robin (2015) szintén arra a következtetésre jutott, hogy az explicitáció és az implicitáció használatával a fordító kerüli az ismétléseket. A szakirodalom azt mutatja, hogy bizonyos esetekben még a gyakorlott fordítók is csökkentik a lexikai ismétléseket, előfordul azonban az is, hogy nem jelentkeznek nagyobb eltérés a forrásnyelvi és a célnyelvi szövegek lexikai ismétlései között. Ez arra enged következtetni, hogy léteznek olyan befolyásoló tényezők, amelyek hatással vannak a lexikai ismétlések kerülésre, illetve meghagyására. E tényezők egyike a szöveg fajtája lehet. Bizonyos szövegtípusok, mint például az irodalmi esetén az ismétléskerülés könnyebben kivitelezhető, a végeredmény pedig koherens, jó felépítésű, választékos célnyelvi szöveg lesz. Műszaki szövegek esetén azonban, mivel ott egész más szempontok játszanak jelentős szerepet, és a terminusok is nagy fontosságot kapnak, az ismétléskerülés negatív hatást gyakorolna a célnyelvi szöveg minőségére.

2.3. A terminusok és fordításuk

Az explicitáció és az implicitáció a különböző szövegtípusoknál más-más gyakorisággal jelentkezik, így tehát ezen a területen a műszaki szövegek fordításánál is érdekes kérdések merülnek fel. Az általánosan elterjedt felfogás szerint a műszaki szaknyelv egyik tipikus jellemzője a terminusok gyakori használata. Annak meghatározása azonban, mit tekinthetünk terminusnak, korántsem egyszerű. Fischer (2012) cikkében rámutat, hogy egy adott lexikai egység terminusként történő kezelésére a fordítói szabadság adja meg a választ, hiszen mindig a fordító feladata, hogy megtalálja azt a célnyelvi megfelelőt, amely a terminus által meghatározott fogalmat jelöli, amennyiben pedig nincs ilyen megfelelő, meg kell alkotnia azt.

A legnehezebb dolga akkor van a fordítónak, amikor olyan terminusokkal kerül szembe, amelyek már a köznyelv részévé váltak, hiszen ezek felismerése jóval nehezebb, mint az egyértelműen felismerhető terminusok esetében, ugyanis a terminushasználat egyik fő feltétele, hogy a fordító ne helyettesítse be más kifejezésekkel. Ez az elgondolás további problémákat vet fel, amelyeket Fischer (2010) doktori értekezésében készített interjúk is alátámasztanak: a fordítók fejében élénken él az a szabályszerűség, miszerint a terminusok helyett nem használhatnak szinonimákat, körülírást, más megoldásokat, és nem alkalmazhatnak az egyértelmű behelyettesítéstől eltérő egyéb fordítói műveletet. Pontosan emiatt az abban a kutatásban részt vevő fordítók is bizonytalanok voltak, vajon terminusnak tekinthető-e az olyan lexikai egység, amelynek ismert szinonimája van. Fischer azonban rávilágít, hogy egyazon jelenséget megjelölő lexikai egységek gyakran a különböző szakterületeken más-más elnevezést kaptak, így szinonimái egymásnak, mégis terminusnak tekinthetők. Heltai (1988) ugyancsak felhívja a figyelmet arra, hogy az angol szaknyelv gyakran alkalmaz szinonimákat terminusok esetében, habár említi azt is, hogy ezzel szemben a magyar szaknyelvre a szinonimák esetében korlátozottabb felhasználás jellemző, sőt, a magyar szaknyelv még laikus közönség számára készült szövegek vagy beszédek esetén is a szakterminusokat alkalmazza.

Az a tény azonban, hogy néhány terminusnak van szinonimája, nem jelenti azt, hogy felhasználásuk szabadon történhet, sőt, a rossz, széttöredezett célnyelvi szövegek egy részének is ez az oka. A fordítóhallgatók gyakran alkalmaznak szinonimákat korábbi iskolai tanulmányaik következményeként, ahol a pedagógusok a szóismétlés kerülésére és a szinonimák használatára ösztönözték őket. A szakszövegekben azonban a terminushasználatnak kohézióteremtő ereje van, így a felesleges vagy esetleg az adott környezetben rossz szinonimák használata negatív hatást gyakorolhat a szöveg kohéziójára (Fischer 2012).

A jelen vizsgálódás szempontjából ezen elméletek és elgondolások nagy jelentőségűek, mivel a szinonimák használatának mértéke, illetve az explicit vagy implicit megfogalmazási módok jelentős hatást gyakorolnak a célnyelvi szövegek szógyakoriságára, sőt, akár a forrásnyelvi és autentikus szövegek esetében is kijelenthetjük ezt, hiszen az ilyen szövegeknél is meghatározó tényezőként szerepelnek az adott jelenségek.

3. A kutatás bemutatása

A jelen kutatás elsődleges célja annak feltárása volt, vajon a számítógépes vizsgálatok milyen eredményeket mutatnak a műszaki szövegek szógyakoriságot célzó vizsgálatánál. Választ kerestünk arra a kérdésre, hogy valóban magasabb szógyakoriság mutatkozik-e a műszaki szövegeknél az úgynevezett lexikai szavak esetében a terminusok fordítási „rugalmatlansága” miatt, bepillantást adva ezáltal a műszaki szövegek világának egy részletébe, és lehetőséget nyújtva további hipotézisek felvetéséhez esetleges későbbi, kvalitatív módszereket is alkalmazó kutatásokhoz.

Feltételezésem szerint a műszaki szakszövegeknél a szórakoztató irodalmi szövegeken végzett elemzéstől eltérően nem fog érvényesülni az ismétléskerülés, mivel a terminusok megakadályozzák a szinonimahasználatot. A vizsgálat-hoz alapul vettem Robin (2014) irodalmi műveken végzett statisztikai elemzését, amelynek során a fordított és lektorált szövegek, illetve autentikus magyar művek viszonyát, szövegjellemzőit vizsgálta korpusznyelvészeti módszerekkel. Ennek keretében arra a következtetésre jutott, hogy a fordításokban az autentikus szövegekhez képest alacsonyabb a lexikai sűrűség, míg a fordítás során keletkezett szövegekben magasabbnak bizonyult a gyakori és kevésbé gyakori kifejezések aránya, és a gyakori szavak sűrűbben ismétlődnek, mint az eredetileg is magyar nyelven született írásokban – alátámasztva a Laviosa (1998) által megfigyelt alapvető lexikai sajátosságokat. A Robin által vizsgált szövegekben kimutatható magas *type/token* arányok azt mutatták, hogy a szórakoztató irodalmi szövegek fordítása sokszínű és változatos, tehát a fordítók szándékosan kerülnek az ismétlést, a lektorok pedig igyekeznek még erőteljesebben hozzájárulni a szinonimák használatához, illetve még gyakrabban nyúlnak az ismétléskerülés egyéb eszközeihez – még az eredetileg magyar nyelven keletkezett szövegeknél fellelhető adatokhoz képest is.

3.1 A korpusz és módszerek

A fordítási korpuszok komoly szerepet töltenek be a nemzetközi nyelvészeti kutatások terén. A korpuszalapú nyelvészeti kutatások Mona Baker (1993) iskolateremtő cikkének következtében alakultak ki és terjedtek el. Baker ebben az írásában azt hangsúlyozta, hogy a korpuszalapú szövegelemzések új távlatokat nyitnak meg, és olyan jellegű kutatásokra adnak lehetőséget, amelyekben az autentikus és fordított szövegek közötti hasonlóságokat és különbségeket lehet felmérni egyszerűen, hatalmas szövegmennyiség vizsgálatával. Habár nemzetközi szinten már számos gépileg olvasható korpusz jött létre, Magyarországon nagyobb méretű, milliós szövegszót tartalmazó szövegtár sokáig nem létezett. A jelen cikk vizsgálatának alapjául szolgáló korpuszt az Eötvös Loránd Tudományegyetem Nyelv-tudományi Doktori Iskoláján belüli Fordítástudományi Doktori Program egyik projektjeként létrehozott Pannónia Korpusz (Robin et al. 2016) szövegei adják. A kutatáshoz a korpusz műszaki szövegeiből kiválasztott autentikus magyar műszaki témájú doktori értekezéseket, illetve angol nyelven írt műszaki témájú könyveket és használati utasításokat, valamint azok magyarra fordítását használtam fel.

Az elemzéshez 23 autentikus értekezést és 7 angol–magyar párhuzamos forrásnyelvi és célnyelvi szöveget vizsgáltam. A korpuszból vett szövegek teljes terjedelme megközelítőleg 500 000 szövegszó mindhárom esetben, így a vizsgálódás teljes terjedelme több mint 1 500 000 szövegszóból áll. Az egyes szövegek hosszúságánál a nyelvek különbözőségéből fakadó okok miatt jelentkezik némi eltérés. Habár először meglepően hangzik, nem a magyar fordítások, hanem az angol eredeti szövegek voltak hosszabbak, aminek oka valószínűsíthetően az, hogy az angol nyelv analitikus. A párhuzamos szövegek hossza nem arányosan oszlik el, tekintve, hogy az informatikai szövegek könyvformátumúak, míg a használati utasítások értelemszerűen rövidebbek, azonban a legrövidebb is eléri a 20 000 szövegszót. Ezzel szemben az autentikus magyar szövegek egyenkénti hossza nagyjából megegyező, minden értekezés megközelítőleg 20 000 és 30 000 szövegszó közötti. Az autentikus szövegtörzset elsősorban a gépészeti, járműtudományi és informatikatudományi munkák alkotják. Szeretném leszögezni, hogy bár az informatikai szövegeket gyakran külön szaknyelvi csoportként kezelik, a jelen tanulmány egy átfogóbb csoportosítás szerint a műszaki szakszövegek közé sorolja őket, és az elemzésül szolgáló korpuszban is található ilyen jellegű szövegek.

A korpusz szövegein a WordSmith Tools 6.0 nevű számítógépes szoftverrel végeztem kvantitatív vizsgálatokat. A szoftveres vizsgálat során Robin (2014) mintáját követve én is Laviosa (1998) kutatását vettem alapul, amely szerint a szógyakorisági vizsgálatoknál a leggyakoribbnak számító szavakat (*list head*) kell figyelembe venni, tehát azokat a szavakat, amelyek az adott szöveg 0,10 százalékát teszik ki. A gépi elemzéssel kigyűjtött leggyakoribb szavak listáján szereplő kifejezéseket grammatikai és funkciószavak szerint csoportosítottam kézi elemzéssel, továbbá a lexikai szavak esetében szintén kézi elemzéssel kiválogattam a terminusnak számító elemeket. Az elemzés megerősítése és az esetleges hibák elkerülése céljából felkértem kollégámat, Ábrányi Henriettát az eredmények kontrollelemzésére. A lexikai és grammatikai szavak eloszlása a szöveg információs terhére és a terminusok mint lexikai egységek gyakoriságá-

ra enged következtetni. Minél nagyobb a tartalmi szavak száma egy szövegben, annál magasabb az információs teher. Stubbs kutatásában (1996) arra jutott, hogy az információs teher mértéke műfajfüggő. Az irodalmi művekben az információs teher alacsonyabb, a szakszövegekben ez az érték jóval magasabb. A besoroláshoz Stubbs (1996: 72) útmutatását követtem, a lexikai szavak közé soroltam a főneveket, igéket, mellékneveket, számneveket, igeneveket, indulatszókat és mondatszókat, a grammatikai szavak közé pedig a névelőket, névutókat, igekötőket, határozószókat, segédigéket, névmásokat és kötőszókat. A szavak szófaját a *Magyar értelmező kéziszótár* (2014) segítségével határoztam meg, majd grammatikai és lexikai csoportokba rendeztem őket. A szótárt a megbízhatóság érdekében használtam, mivel az egyes szócikkek egyértelműen meghatározzák a keresett szó szófaját, ahogyan az alábbi szócikkekben is látható:

fényképezőgép ■■■■ fn

Fényképfelvétel készítésére szolgáló optikai berendezés.

Szin.: (rég) *fotográfmasina*

Etim.: Összetétel; *fényképező*: a → *fényképez* ige származéka + → *gép*.

a ■■■■ határozott névelő (magánhangzóval kezdődő szó előtt) **az**

Etim.: Ősi uráli kori szó.

A névelő két alakjáról és használatukról lásd az II. szócikket!

Erre való tekintettel a grammatikai és lexikai csoportokba rendezés nem szubjektív megítélés alapján történt, hanem objektív döntésnek tekinthető.

A műszaki szövegekben a lexikai szavak között természetesen előfordulnak olyan terminusok is, amelyek csak néhány alkalommal szerepelnek a szövegben. Habár természetesen terminus voltukat ezáltal nem veszítik el, a jelen tanulmány szempontjából az ilyen típusú terminusok a gyakori ismétlődés hiánya miatt nem relevánsak, mivel a szinonimahasználat kérdése csak a gyakori terminusokat érinti. Ezeknek az azonosítása tehát kívül esett a jelen kutatás célkitűzésein.

A kutatás során a szoftver segítségével kiszámítottam az egyes szövegek type/token (szótípus/szövegszó) arányát is. Az elemzés már nem csupán a leggyakoribbnak számító szavakra, hanem a teljes szövegekre irányult. A szószám (token) az összes előforduló szövegszót mutatja, a szótípus (type) pedig a korpuszban fellelhető szótári szavak számát. Ezek aránya a korpuszra jellemző lexikai változatosságot mutatja (Seidl-Pécs 2011: 106). Ha tehát alacsonyabbnak bizonyul a szóalakok száma az összes szövegszó mennyiségéhez képest, mint az irodalmi szövegekben (Robin 2014), akkor ugyancsak közvetett bizonyítékot nyerünk a terminushasználatból eredő gyakoribb szóismétlésekre.

4. Eredmények bemutatása

A leggyakoribbnak számító szavakat megvizsgálva a következő táblázatokban bemutatott értékek rajzolódnak ki a lexikai és grammatikai elemek arányával kapcsolatban. A táblázatok harmadik oszlopában a lexikai szavakban található

terminusok százalékos aránya látható. Ebben a számításban a leggyakrabban előforduló lexikai szavak számát 100%-nak vettem, majd kiszámoltam, ennek hány százalékát teszik ki a kézi elemzés és az azt követő kontrollelemzés után a terminusokra vonatkozó értékek. A táblázatok negyedik oszlopában a vizsgált szövegekben található szövegszó és szótípus (type/token) aránya, illetve azok átlaga tekinthető meg. Ezzel a vizsgálattal a szóismétlésekre, a szókincs gazdagságára következtethetünk, illetve a szavak előfordulásának gyakoriságára. Ezek az értékek minden esetben standardizált értékek. A számok azt mutatják, hogy a magyar nyelvű szövegekben, legyen szó akár fordított, akár autentikus szövegről, magasabb a type/token arány, mint az angol szövegekben.

Az 1. táblázat a párhuzamos szövegek angol eredeti változatánál mutatja a leggyakoribb szavak eloszlását.

1. táblázat

Forrásnyelvi angol szövegek elemzési adatai

	Lexikai (tartalmi) szavak %	Grammatikai (funkció) szavak %	Terminus értékű lexikai szavak %-os aránya	Type/token arány
1. szöveg	61	39	49	36,75
2. szöveg	72	28	58	28,11
3. szöveg	75	25	60	28,05
4. szöveg	62	38	37	37,52
5. szöveg	70	30	67	28,11
6. szöveg	68	32	72	33,13
7. szöveg	69	31	62	26,83
Átlag	68	32	59	31,21

A számokból úgy látszik, hogy a szövegekben a lexikai, más néven tartalmi elemek szerepelnek többségben a leggyakoribb szavak között. Ez arra enged következtetni, hogy a gyakori szavak között több olyan kifejezés is található, amely terminus értékű szó, állandó ismétlésük szükséges a szöveg teljes terjedelmében. Érdekes jelenség, hogy az egyes szövegek között, habár témájukat tekintve nem egyformák, mégsem figyelhető meg nagy különbség, a lexikai és a grammatikai elemek aránya szövegenként nagyjából megegyező: a lexikai elemek eloszlásánál az adatok terjedelme a legkisebb elemszámtól egészen a legnagyobbig 14 (61–75). A terminusok százalékos arányánál megfigyelhetjük, hogy a 4. szöveg kivételével kifejezetten magas százalékban szerepelnek terminusok a szövegek leggyakoribb lexikai szavai között, de a legalacsonyabb százalékérték is 37, kijelenthetjük tehát, hogy az angol eredeti szövegek leggyakoribb szavaiban a terminusok előfordulási aránya több, mint 1/3. A type/token arány a forrásnyelvi

szövegek esetében 31,21. Ugyanez az érték Robin (2014) irodalmi szövegekre vonatkozó vizsgálatában 42,05 volt, amely jóval magasabb szám. Ennek értelmében arra következtethetünk, hogy a terminusok miatt jóval szerényebb a szövegek szókinccse.

2. táblázat

Célnyelvi magyar szövegek elemzési adatai

	Lexikai (tartalmi) szavak %	Grammatikai (funkció) szavak %	Terminus értékű lexikai szavak %-os aránya	Type/token arány
1. szöveg	71	29	55	50,41
2. szöveg	84	16	45	43,92
3. szöveg	78	22	62	44,40
4. szöveg	75	25	57	50,56
5. szöveg	79	21	50	39,46
6. szöveg	75	25	49	44,01
7. szöveg	75	25	58	38,29
Átlag	77	23	53	44,43

A 2. táblázat az első táblázatban bemutatott angol szövegek magyar nyelvű fordításainál mutatja a leggyakoribb szavak eloszlását. Habár a szövegpárok között nem figyelhető meg kiugrásszerű különbség a szógyakoriság eloszlásában, mégis jól látszik, hogy a fordított magyar szövegekben a lexikai szavak aránya valamivel magasabb számokat mutat, mint azok angol forrásnyelvi változatainál. Ez első sorban a magyar nyelv szintetikus jellegével magyarázható, míg az analitikus angol nyelvben akkor is szükség van különálló funkciószavakra, amikor a magyar nyelvben nem. Az eltérő rendszer mellett ugyanakkor az is elképzelhető, hogy a fordítók grammatikai elemek helyett lexikai elemeket alkalmaztak, pl. terminusok betoldásával gazdagították a szöveget. Ezt azonban csak olyan kvalitatív kutatás mutatná ki, amely a fordítási műveleteket vizsgálja. A magyarban ez a szógyakorisági vizsgálatot tekintve alapvetően inkább hátrányt kellene, hogy jelentsen, hiszen a WordSmith Tools és a hozzá hasonló programok nem képesek arra, hogy a szavakat a szótó alapján rendezzék gyakorisági listákba, így a ragozott alakok külön egységként jelennek meg. A lexikai elemek eloszlásánál az adatok terjedelme 15 (69–84), ami magasabb, mint az angol autentikus szövegeknél, tehát a célnyelvi alkorpusz szövegei nyelvezetükben jobban különböznek egymástól, mint a forrásnyelvi alkorpuszt felépítő szövegek.

A type/token arány a fordított szövegek esetében 44,43, ez magasabb a forrásnyelvi szövegek type/token arányánál, ami értelemszerű, hiszen a két nyelv rendszere jelentősen eltér, és a magyar nyelvben alapvetően is kevesebb ismét-

lódó funkciószó van, mint az angolban. Robin (2014) irodalmi szövegeinél azonban ez az érték 58,93 volt, így az látszik, hogy az eredeti szövegek esetében megmutatózó tendencia a fordításoknál is megjelenik: a műszaki szövegek szókinccse szerényebb az irodalmi szövegekénél az információs teher ellenére is, hiszen kevesebb a szinonima, többször ismétlődnek a lexikai elemek. Érdeemes megfigyelni azonban a 2. táblázat 1. és 2. szövegét: a másodiknál sokkal alacsonyabb a type/token arány, mégis sokkal kevesebb terminus értékű lexikát tartalmaz, pedig kimagasló a lexikai szavak aránya. Ha azonban az 1. táblázat 5. és 6. szövegét emeljük ki, az 5-nél jóval alacsonyabb a type/token arány, mégis sokkal kevesebb terminus értékű lexikát tartalmaz, pedig a lexikai szavak aránya jelentős. Hasonló jelenség figyelhető meg továbbá az 1. táblázat 2., 3. és 5. szövegénél is, ahol a type/token aránya közel azonos, de a terminus értékű szavak aránya eltér egymástól. A vizsgálat eredményei tehát ugyan nem konkluzívak, mégis megállapíthatjuk, hogy a többségük egy irányba mutat.

Bár a szövegek többségénél az látható, hogy a type/token arány alacsonyabb, mint Robin (2014) vizsgálatában, felmerül azonban a lehetőség, hogy a type/token arány nem kizárólag a terminológia ismétlődésének köszönhető, hanem más tényező is szerepet játszhat, akár a formulaszerű nyelvhasználat is.

Megfigyelhető továbbá, hogy a két műfaj szövegei között a különbség még nagyobb, mint a forrásnyelvi szövegek esetében. Az eltérésekből arra is következtethetünk, hogy a fordítók talán jobban ügyelnek a terminusok használatára a műszaki szövegek esetén, és jobban követik a forrásnyelvi szöveget is.

3. táblázat

Autentikus magyar szövegek elemzési adatai

	Lexikai (tartalmi) szavak %	Grammatikai (funkció) szavak %	Terminus értékű lexikai szavak %-os aránya	Type/token arány
1. szöveg	75	25	34	56,08
2. szöveg	64	36	50	47,15
3. szöveg	67	33	37	52,28
4. szöveg	62	38	48	52,37
5. szöveg	62	38	40	62,90
6. szöveg	67	33	32	53,97
7. szöveg	74	26	47	51,62
8. szöveg	73	27	41	51,62
9. szöveg	67	33	31	57,40
10. szöveg	72	28	42	50,06
11. szöveg	80	20	40	50,23

	Lexikai (tartalmi) szavak %	Grammatikai (funkció) szavak %	Terminus értékű lexikai szavak %-os aránya	Type/token arány
12. szöveg	77	23	43	49,25
13. szöveg	69	31	33	55,76
14. szöveg	77	23	36	53,97
15. szöveg	71	29	34	51,19
16. szöveg	71	29	34	53,58
17. szöveg	60	40	45	48,12
18. szöveg	72	28	33	50,83
19. szöveg	63	37	41	53,41
20. szöveg	79	21	35	54,81
21. szöveg	72	28	36	53,04
22. szöveg	68	32	32	56,51
23. szöveg	76	24	42	53,04
Átlag	70	30	39	53,00

A 3. táblázat az autentikus magyar szövegeknél mutatja be, milyen a lexikai és a grammatikai szavak eloszlása a szógyakorisági listák élén. Az autentikus magyar szövegeket adó értekezések esetében a táblázat azért áll ennyi elemből, mert úgy döntöttem, realisabb képet ad, ha az egyes szövegeket külön egységként kezeltem. Témájukat tekintve ugyan több értekezést is össze lehetne vonni, hiszen az egyes tanszékek doktori disszertációi egy-egy nagy téma köré csoportosulnak, az egyes munkák azonban mégis jelentősen eltérnek altémájuk szerint, így ebből a megfontolásból nem vontam össze őket egymással. A táblázatból jól kivehető, hogy az egyes szövegek lexikai és grammatikai szavainak eloszlása nagyjából azonos síkon mozog, és minden értekezés esetében elmondható, hogy a tartalmi szavak a leggyakoribb szavak listáját figyelembe véve többségben vannak a funkciószavakhoz képest. Érdekes jelenségnek vélem azonban, hogy a magyar nyelvű fordított szövegeknél alapvetően magasabb értékeket kaptam a lexikai kifejezések között. Arra a kérdésre, hogy ezt – az egyébként csekély – különbséget valóban az adja, hogy az egyik halmaz szövegei autentikus magyar szövegek, a másik halmazban pedig fordított szövegek vannak, és ezáltal a különbséget a fordítási folyamat, a fordítási műveletek adják-e, a jelen tanulmány még nem tud választ adni. Mindenesetre az elemzés arra enged következtetni, hogy az autentikus és a fordított műszaki szövegek szógyakorisági mutatói között különbség fedezhető fel, a különbség pedig abban áll, hogy az autentikus szövegek leggyakoribb szavai között több a tartalmi szó. A lexikai elemek eloszlásánál az adatok terjedelme 20

(60–80). A három szövegcsoporthoz itt a legnagyobb a különbség az elemadatok terjedelmében, tehát itt a legnagyobb az eltérés a szövegek között abban a tekintetben, hogy mennyi azon lexikai elemek száma, amelyeket a szöveg teljes terjedelmében folyamatosan ismételt a szerző.

A terminus értékű szavak százalékos arányánál itt mutatkoznak a legalacsonyabb számok. A legalacsonyabb százalékos érték 31%, a legmagasabb pedig 50%, így tehát a leggyakoribb lexikai szavak maximum fele terminus értékű szó. Habár ezek az értékek alacsonyabbak, mint a forrásnyelvi és a fordított célnyelvi szövegek listáiban találhatóak, mégsem tekinthetjük csekély számnak, hiszen a 31%-os legalacsonyabb arány is azt mutatja, hogy a szavak közel egyharmada terminusértékű.

Az autentikus magyar műszaki szövegeknél kimutatott type/token arány 53,00, ami jóval magasabb a fordított szövegekéénél. Ez arra utalhat, hogy a fordítók sokkal jobban ragaszkodnak a forrásnyelvi szöveghez, nem kerülnek az ismétléseket, ehelyett inkább kerülnek a kockázatot – vagyis a szinonimahasználat kockázatát. Robin (2014) kutatásában az autentikus irodalmi szövegek type/token aránya 61,8, amely ugyan nem olyan sokkal magasabb, mint a forrás- és célnyelvi szövegek esetében, mégis utal a terminushasználatból eredő szegényebb szókincsre, amely eredhet a formulaszerű nyelvhasználatból. A számok mutatják továbbá, hogy míg Robin vizsgálatában a magyarra fordított és az autentikus magyar szövegek type/token aránya között nincs nagy különbség, a jelen vizsgálat elemzése során nagy különbség mutatkozott. A két kutatás összevetése során jelentkező tendencia azt mutatja, hogy az irodalmi szövegek fordítói a szókincs gazdagítására törekcszenek, míg műszaki szövegeknél pedig ismételtetik a terminusokat, a bevett szókapcsolatokat, frázisértékű kifejezéseket.

Robin (2014) vizsgálatában arra jutott, hogy az irodalmi szövegek esetében a fordításokban kisebb lexikai sűrűség jelentkezett, mint az autentikus szövegekben, a fordítók pedig jobban kerültek az ismétléseket, mint a magyar írók. Ennek következtében az autentikus magyar szövegek szógyakorisága nagyobb volt, mint a fordított magyar szövegeké. A műszaki szövegeket vizsgálva az mutatkozott, hogy a fordított szövegekben a leggyakoribb szavak között nagyobb a lexikai elemek száma a grammatikai szavakéhoz képest. Továbbá az elemzés eredményeiből kiolvasható, hogy a fordított szövegek esetében nem csupán az rajzolódik ki, hogy a szógyakoriság tekintetében egyértelmű többségben vannak a tartalmi szavak, de érdekes módon megközelítőleg hasonló értékek születtek a forrásnyelvi angol és az autentikus magyar nyelvű írásokban, holott a két vizsgált nyelv felépítését és rendszerét tekintve eltérő.

5. Összefoglalás

A tanulmányban különböző műszaki témájú szövegek főbb lexikai és grammatikai szavainak eloszlását számoltam ki és vizsgáltam meg WordSmith Toolsszal elkészített szógyakorisági listák alapján. Az elemzéshez 23 autentikus értekezést és 7 angol–magyar párhuzamos forrásnyelvi és célnyelvi szöveget vizsgáltam.

Mindhárom szövegcsoport esetében ugyanaz az összesített eredmény mutatkozott: a műszaki szövegeket vizsgálva minden esetben a lexikai, más néven tartalmi elemek kerültek többségbe a szógyakorisági listák leggyakoribb szavainak számba vétele során. Kijelenthetjük továbbá, hogy a szógyakoriság tekintetében az angol, a magyar fordított és a magyar autentikus szövegek elemzésekor egyaránt nagyjából hasonló arányú eloszlás figyelhető meg, függetlenül attól, hogy az adott szöveg a műszaki szakszövegek nagyobb halmazán belül mely kisebb alkategóriába sorolható.

A szövegek type/token arányának vizsgálatakor megmutatkozott továbbá, hogy összevetve Robin szórakoztatóirodalmi szövegeken végzett vizsgálatával, a műszaki szövegek esetében rendre alacsonyabb értékeket kaptunk. A type/token arányok továbbá rámutatnak arra is, hogy a műszaki szövegek esetében a magyar autentikus szövegek írói jobban ügyelnek a változatos szóhasználatra, és lehetőségeikhez mérten igyekeznek kerülni az ismétléseket. Ezzel szemben a fordítóknál azt mutatják az arányszámok, hogy feltehetőleg ők jobban ragaszkodtak ahhoz, hogy betartsák a műszaki szövegek terminusokat megtartó jellegét. A lexikai szavak magas szintje és a type/token arányok közvetett bizonyítékként szolgálhatnak arra, hogy ugyanazok a kifejezések ismétlődnek a szövegben, amelyek a szógyakorisági listák leggyakoribb lexikai szavait vizsgálva nagy arányban terminusok. Ez arra enged következtetni, hogy a műszaki szövegek szókincse szegényebb az irodalmi szövegekénél, ennek magyarázata pedig a terminushasználat erőteljes jelenléte, ami a fordítók számára is követendő, ugyanis a megfelelő minőségű célnyelvi szöveg létrehozása érdekében kerülniük kell a szinonimahasználatot, és sokkal jobban ragaszkodni a forrásnyelvi szöveghez is.

A jelen kutatás mindössze kvantitatív eredményekkel szolgált a felvetett vizsgálatot illetően, azonban a téma teljes körűjárása érdekében egy további elemzés során kvalitatív eredményekkel is ki lehetne egészíteni, amihez a Pannónia Korpusz kiváló alapként szolgál.

A tanulmány a műszaki szövegek szógyakoriságának jellemzőit igyekezett feltárni, illetve választ találni arra a kérdésre, milyen hatást gyakorolnak a terminusok az ilyen műfajú szövegek szógyakoriságára. A vizsgálatokat követően arra a következtetésre jutottam, hogy az eredmények nem konkluzívak, azonban a vizsgált szövegek többsége egy irányba mutat. Olyan eredményeket is kaptam, amelyek arra engednek következtetni, hogy a kutatást érdemes volna más irányba is bővíteni, például a formulaszerű nyelvhasználat hatásának figyelembe vétele, lemmatizált listák elemzése, illetve a két vagy több tagból álló kifejezések vizsgálata N-gram vizsgálatokkal. Ez utóbbi azért volna szükségszerű, mert a szétbontott lexikai elemek nem feltétlenül jelentkeznek terminusként, egyben azonban már igen.

Kitekintésként megemlíteném a természettudományos szövegek csoportját is, mivel az oda sorolt írások jelentős hasonlóságokat mutatnak a műszaki szakszövegekkel, gyakran említik őket rokon szövegtípusként is. A hasonlóságot főként a cikkben többször is említett gyakori terminushasználatban láthatjuk, így egy következő kutatás keretében érdekes volna elvégezni egy szógyakorisági elemzést természettudományi szövegekből álló korpuszon is, majd összevetni a műszaki szövegeken végzett vizsgálatok eredményeivel.

Irodalom

- Baker, M. 1993. Corpus linguistics and translation studies: implications and applications. In: Baker, M., Francis, G., Tognini-Bonelli, E. (eds) 1993. *Text and Technology: in Honour of John Sinclair*. Amsterdam/Philadelphia: John Benjamins. 233–50.
- Blum-Kulka, S. 1986. Shifts of Cohesion and Coherence in Translation. In: House, J. Blum-Kulka, S. (eds) *Interlingual and Intercultural Communication. Discourse and Cognition in Translation and Second Language Acquisition*. Tübingen: Narr. 17–35.
- Fischer M. 2010. *A fordító mint terminológus, különös tekintettel az európai uniós kontextusra*. Doktori disszertáció. Budapest: ELTE BTK.
- Fischer M. 2012. Elméleti és módszertani adalék a terminológia oktatásához I. Terminológiaelméleti alapkérdések a fordításban. *Fordítástudomány* XIV. évf. 2. szám. 5–30.
- Heltai P. 1988. Contrastive Analysis of Terminological Systems and Bilingual Technical Dictionaries. *International Journal of Lexicography* Vol. 1. No. 1. 32–40.
- Klaudy K. 1999. Az explicitációs hipotézisről. *Fordítástudomány* II. évf. 1. szám. 5–22.
- Klaudy K. 2007. *Nyelv és fordítás*. Budapest: Tinta Könyvkiadó.
- Klaudy K., Károly K. 2005. Implication in Translation: Empirical Evidence for Operational Asymmetry in Translation. *Across Languages and Cultures* Vol. 6. No. 1. 13–29.
- Laviosa, S. 1998. The English Comparable Corpus: A Resource and a Methodology. In: Bowker, L., Cronin, M., Kenny, D., Pearson, J. (eds) *Unity in Diversity: Current Trends in Translation Studies*. Manchester: St. Jerome. 101–112.
- Olohan, M., Baker, M. 2000. Reporting that in Translated English. Evidence for Subconscious Processes of Explication. *Across Languages and Cultures* Vol. 1. No. 2. 141–159.
- Pusztai F. (szerk.) 2014. *Magyar értelmező kéziszótár*. Budapest: Akadémiai Kiadó.
- Robin E. 2014. Lektorált fordítások és eredeti magyar szövegek gépi összehasonlítása. Elhangzott: *Nyelv, kultúra és társadalom*. ANYK Konferencia. Kodolányi János Főiskola, Budapest 2014. november 3–4.
- Robin E. 2015. *Fordítási univerzálék a lektorált szövegekben*. Doktori disszertáció. Budapest: ELTE BTK.
- Robin E., Dankó Sz., Götz A., Nagy A. L., Pataky É., Szegh H., Zolczer P. 2016. Fordítástudomány és korpuszkutatás: bemutatkozik a Pannónia Korpusz. *Fordítástudomány* XVIII. évf. 2. szám. 5–26.
- Seidl-Péché O. 2011. *Fordított szövegek számítógépes összevetése*. Doktori disszertáció. Budapest: ELTE BTK.
- Stolze, R. 1999. 2005. *Die Fachübersetzung: eine Einführung*. Tübingen: Narr.
- Stubbs, M. 1996. *Text and Corpus Analysis. Computer-assisted Studies of Language and Culture*. London: Blackwell.
- Sturcz Z. 2010. Megállapítások és felvetések a műszaki tudományok nyelvéről. In: Dobos Cs. (szerk.) *Szaknyelvi kommunikáció*. Budapest: Tinta Könyvkiadó.
- Toury, G. 1995. *Descriptive Translation Studies and Beyond*. Amsterdam: John Benjamins
- Vinay, J. P., Darbelnet, J. [1958] 1995. *Comparative Stylistics of French and English. A methodology for Translation*. Amsterdam: John Benjamins.
- Zabóné Varga I. 2015. *Műszaki szövegek fordításának terminológiai problémái német-magyar nyelvpárban járműipari szövegek alapján*. Doktori disszertáció. Budapest: ELTE BTK

Források

PANNÓNIA KORPUSZ PÁRHUZAMOS KORPUSZ

Eredeti szakirodalom:

2011. *NIKON P550 User manual*. Nikon Corporation.
 2012. *NIKON D3200 User manual*. Nikon Corporation.
 2014. *NIKON D810 User manual*. Nikon Corporation.
 Chappell, D. 2008. *Introducing a .NET Framework 3.5*. Boston (USA): Addison-Wesley Professional.
 Mackey, A. 2010. *Introducing .NET 4.0 With Visual Studio 2010*. New York: Apress
 Microsoft Windows Server Team. 2004. *Migrating from Microsoft Windows NT Server 4.0 to Windows Server 2003*. Washington: Microsoft
 Stanek, W. 2008. *Microsoft® Exchange Server 2007 Administrator's Pocket Consultant, Second Edition*. Washington: Microsoft.

Fordított szakirodalom:

2011. *NIKON P550 Használati utasítás*. Nikon Corporation.
 2012. *NIKON D3200 használati utasítás*. Nikon Corporation.
 2014. *NIKON D810 használati utasítás*. Nikon Corporation.
 Chappell, D. 2008. *Bemutatózik a .NET Framework 3.5*. Microsoft. Bicske: Szak Kiadó.
 Mackey, A. 2010. *A .NET 4.0 és a Visual Studio*. Bicske: Szak Kiadó.
 Microsoft Windows Server Team. 2004. *Áttérés a Windows Server 2003-ra. Segédlet kis és közepes vállalatoknak*. Bicske: Szak Kiadó.
 Stanek, W. 2008. *Microsoft Exchange Server 2007 – A rendszergazda zsebkönyve*. Bicske: Szak Kiadó.

PANNÓNIA KORPUSZ ÖSSZEHASONLÍTHATÓ KORPUSZ

Eredeti szakirodalom:

- Bánlaki P. 2014. *Gépjármű hajtáslánc fődarabok rezgés- és zajdiagnosztikai végellenőrző rendszerének továbbfejlesztése*. Doktori disszertáció. Budapest: BME.
 Bauernhuber A. 2015. *Lézersugaras fém-polimer kötés kialakításának és tulajdonságainak vizsgálata*. Doktori disszertáció. Budapest: BME.
 Bede Zs. 2013. *Változtatható irányú forgalmi sávok analízise nagyméretű közúti közlekedési hálózatokon*. Doktori disszertáció. Budapest: BME.
 Boronkai L. 2011. *Villamos vontatójárművek hajtásdinamikai folyamatainak szimulációja a jármű keresztirányú mozgásának figyelembevételével*. Doktori disszertáció. Budapest: BME.
 Dömötör F. 2013. *Összetett szerkezetű járműanyagok (hibridek, kompozitok) határfelületei megmunkálási folyamatainak korszerű technológiai diagnosztikája*. Doktori disszertáció. Budapest: BME.
 Ficzer P. 2014. *Gyors prototípus numerikus és kísérleti szilárdsági analízise*. Doktori disszertáció. Budapest: BME
 Hargitai L. 2016. *Folyami áruszállító hajók manőverképességét előrejelző mozgás-szimulációs módszer kidolgozása*. Doktori disszertáció. Budapest: BME.
 Hokstok Cs. 2013. *Vasúti infrastruktúragazdálkodás kontrolling bázisú döntéselőkészítő rendszerek alkalmazásával*. Doktori disszertáció. Budapest: BME.
 Korody E. 2007. *Multifunkcionális, új generációs repülésszimulátor kifejlesztése és különböző kormányzási megoldások vizsgálata*. Doktori disszertáció. Budapest: BME.

- Kovács G. 2011. *Elektronikus fuvar- és raktérbörze rendszermodellje*. Doktori disszertáció. Budapest: BME.
- Kovács K. 2011. *Járműalkatrészek és szerkezetek élettartamának és kifáradási folyamatának statisztikai vizsgálata*. Doktori disszertáció. Budapest: BME.
- Meyer D. 2015. *Repülésbiztonsági szint alapú eljárás-befolyásolás a polgári célú légitörekedésben: Az airside, pre-take-off objektum- és folyamatcsoport biztonságintegritása*. Doktori disszertáció. Budapest: BME.
- Nagy A. 2014. *Kisméretű légi eszközök mozgásfolyamatát meghatározó mérési és szimulációs környezet fejlesztése: siklóernyők fordulási tulajdonságainak elemzése*. Doktori disszertáció. Budapest: BME.
- Ozsváth P. 2009. *Magnézium alapú hibrid járműanyagok környezetbarát forgácsolásának optimalálása*. Doktori disszertáció. Budapest: BME.
- Selymes P. 2011. *A légi személyszállítás értékképzési folyamatának modellezése*. Doktori disszertáció. Budapest: BME.
- Simongáti Gy. 2009. *STPI (a Fenntartható Közlekedés Mutatója) kidolgozása a belvárosi hajózás fenntarthatóság elve szerinti értékeléséhez*. Doktori disszertáció. Budapest: BME.
- Szabó A. 2014. *Összetétel, fázisviszonyok és feszültségállapot vizsgálata járműipari ötvözetekben termofeszültségméréssel*. Doktori disszertáció. Budapest: BME.
- Szabó B. 2014. *Gumiabroncsos járművek kissebességű pályamozgásának és a gumiabroncs deformációjának kapcsolata*. Doktori disszertáció. Budapest: BME.
- Szabó G. 2008. *Nagy megbízhatóságú elektronikus közlekedési alrendszerek RAMS paramétereinek kezelése*. Doktori disszertáció. Budapest: BME.
- Török Á. 2010. *A fenntartható városi közlekedés feltételei és a megvalósítás eszközrendszere*. Doktori disszertáció. Budapest: BME.
- Vámossy Z. 2009. *Mobil robotok navigációja PAL-optikára alapozott gépi látással*. Doktori disszertáció. Budapest: BME.
- Vincze-Pap S. 2008. *Autóbuszok ütközésállósági vizsgálatai és vizsgálati módszerei, különös tekintettel a borulásbiztonságra, a vázszerkezetek képlékeny csuklóira és zónáira*. Doktori disszertáció. Budapest: BME.

Függelék

Szógyakorisági lista minta

kurzivalva – lexikai elemek, kurzivalva és vastagítva – terminus értékű lexikai elemek

N	Word	Freq.	N	Word	Freq.
1	#	1676	28	<i>emberi</i>	40
2	A	1401	29	<i>integrált</i>	40
3	Az	490	30	<i>paraméterek</i>	39
4	És	465	31	<i>radar</i>	39
5	<i>rendszer</i>	158	32	<i>során</i>	38
6	hogy	99	33	amely	37
7	<i>biztonsági</i>	85	34	Q	37
8	is	85	35	<i>hibafa</i>	36
9	<i>fuzzy</i>	83	36	illetve	36
10	<i>IKLR</i>	72	37	<i>pre</i>	36
11	valamint	68	38	<i>légiközlekedési</i>	35
12	<i>légijármű</i>	67	39	<i>eljárás</i>	33
13	<i>futópálya</i>	62	40	<i>légiközlekedés</i>	33
14	egy	59	41	<i>polgári</i>	33
15	vagy	58	42	<i>kockázatelemzés</i>	32
16	<i>veszély</i>	57	43	<i>meteorológiai</i>	32
17	nem	54	44	<i>megfelelő</i>	31
18	<i>ábra</i>	52	45	<i>eljárások</i>	30
19	<i>alapján</i>	52	46	<i>gurulóút</i>	29
20	<i>elemi</i>	52	47	<i>repülőtér</i>	29
21	<i>szint</i>	51	48	<i>vonatkozó</i>	29
22	<i>repülésbiztonsági</i>	50	49	<i>esemény</i>	28
23	meg	47	50	ha	28
24	<i>modell</i>	46	51	<i>hiba</i>	28
25	<i>rendszerek</i>	46	52	<i>palarp</i>	28
26	<i>légiforgalmi</i>	43	53	<i>tagsági</i>	28
27	<i>egyés</i>	40	54	tehát	28

N	Word	Freq.	N	Word	Freq.
55	forgalmi	27	78	való	23
56	szerint	27	79	azok	22
57	alapú	26	80	figyelemben	22
58	fények	26	81	komplex	22
59	következtető	26	82	off	22
60	alkalmazott	25	83	adott	21
61	bekövetkezési	25	84	látható	21
62	hibája	25	85	működő	21
63	aktuális	24	86	szolgáló	21
64	csúcsesemény	24	87	annak	20
65	ILS	24	88	elemzés	20
66	kockázati	24	89	értékek	20
67	minden	24	90	események	20
68	mint	24	91	földi	20
69	műszaki	24	92	humán	20
70	airside	23	93	kapcsolatos	20
71	amelyek	23	94	ki	20
72	biztonságigazolási	23	95	vonatkozóan	20
73	esetében	23	96	által	19
74	irányítás	23	97	el	19
75	kell	23	98	fizikai	19
76	megfelelően	23	99	szükséges	19
77	olyan	23	100	take	19