

Lektorált fordítások és eredeti magyar szövegek gépi összehasonlítása

Robin Edina

E-mail : robin.edina@gmail.com

Kivonat: A leíró fordítástudomány lényeges feladatának tekinti a fordítási szöveg általános törvényszerűségeinek, sajátosságainak feltárását, amelyek megkülönböztetik a nem fordítás útján keletkezett szövegektől. Ámde a párhuzamos és összehasonlítható korpuszokon végzett elemzések nem számolnak a lektorálás során végrehajtott módosításokkal, amelyek szintén befolyásolhatják a fordítások szövegszintű tulajdonságait. Felmerül tehát a kérdés, hogy a lektorálás eredményeképpen mi történik a fordítási univerzálékkal a fordított szövegekben, vajon a lektorok módosításai inkább a célnyelvi normához közelítik-e a fordítások jellemzőit. Jelen kutatás során általános statisztikai gépi elemzéseket végeztünk tíz szórakoztató irodalmi regény fordítói és lektori kéziratán, majd a kapott eredményeket tíz, eredetileg magyar nyelven született regény statisztikai adataival is egybevetettük. Az elemzések között szerepelt a mondatok átlagos hosszának megállapítása, a standardizált szóalak és szótípus (type/token) arányok közötti különbségek vizsgálata, a lexikai gyakoriság elemzése, a lexikai sűrűség kimutatása és az adatok szórásának megállapítása. Az eredmények arra engednek következtetni, hogy a lektorálás során a fordítási szöveg tulajdonságai többnyire az eredetileg is célnyelven született, autentikus szövegek jellemzőihez, vagyis a célnyelvi normához közelítenek, a lektorálásban ugyanakkor megmutatkoznak bizonyos univerzális szerkesztési jellegzetességek is.

Kulcsszavak: univerzálék, fordítás, lektorálás, optimalizáció, korpuszkutatás

1. Bevezetés

A leíró fordítástudományban a normák és az általános szabályszerűségek tudományos vizsgálatai nyomán fogalmazódtak meg az úgynevezett fordítási univerzálék, amelyek leírják, milyen pontokon és minőségben különböznek a fordítások a forrásnyelvi szövegektől, illetve a nem fordítás nyomán, hanem eredetileg célnyelven született írásoktól (Baker 1993, 1995). A fordítási szöveg általánosan megfigyelhető univerzális nyelvi jellemzői közt tartjuk számon többek között az egyszerűsítést, a konvencionális nyelvhasználatot, a forrásnyelvi szövegtulajdonságok átvételét, a lexikai egységek sajátos megoszlását, a szóismétlések kerülését, az explicitésre való törekvést, a redundanciát és a kiegyenlítődést. A fordításokat vizsgáló empirikus kutatások többnyire

alátámasztják ezeknek az univerzális nyelvi sajátosságoknak az előfordulását a fordítással keletkezett szövegekben (Laviosa 2009). Nem veszik azonban figyelembe azt a tényt, miszerint az elemzéseik alapjául szolgáló korpuszok általában publikált, lektorált fordításokból állnak, amelyek már nemcsak a fordítók stratégiáinak és műveleteinek, hanem a lektorok beavatkozásainak nyomait is magukon viselik.

Korábbi kutatásaim során megállapítottam (Robin 2014a), hogy a lektorok módosításainak nyomán megfigyelhetők bizonyos statisztikai eszközökkel kimutatható különbségek a fordítói és lektori kéziratok között, általános tendenciákra utalva. Az eredmények tanúsága szerint a lektorok az ekvivalencia ellenőrzése és a helyesírási, nyelvtani hibák javítása mellett töreksenek a szöveg redundanciájának enyhítésére, a lexika gazdagítására, a szöveg feldolgozhatóságának könnyítésére, és igyekeznek kerülni a szóismétléseket. Munkájuk nyomán tehát bizonyos fordítási univerzálénak tartott jelenségek enyhülnek, míg mások felerősödnek, ami felveti azt a gondolatot, hogy némelyik fordítási sajátosságnak kikiáltott jelenség valójában inkább szövegszerkesztési univerzálét takar, hiszen ez a tevékenység a fordításnak és a lektorálásnak egyaránt szerves része (Robin 2014b). Ám egyelőre még megválaszolatlan a kérdés, vajon a lektor módosításai valóban a célnyelvi szövegekre jellemző normákhoz közelítik-e a fordítások jellemzőit, vagy egyfajta konvencionális, szerkesztési norma irányába viszik el inkább a fordított szöveget. A felvetés megvizsgálásához egy lektorálási korpusz és egy magyar nyelvű összehasonlítható korpusz egybevetése vált szükségessé.

2. Fordítási univerzálék és korpuszkutatások

A fordítási szöveget vizsgáló empirikus kutatásoknak Mona Baker (1995) programadó tanulmánya nyomán új lendületet adott a kétnyelvű párhuzamos korpuszok és az egynyelvű összehasonlítható korpuszok elemzése, amelyek lehetővé teszik a fordítások statisztikai összehasonlítását a célnyelven keletkezett szövegekkel. Utóbbi vizsgálatokat Baker különösen fontosnak tartja, hiszen amennyiben a fordítások valóban rendszerszerűen különböznek az eredeti célnyelvi szövegektől, szükségszerűen létezniük kell olyan nyelvi sajátosságoknak is, amelyek a fordított szövegeket megkülönböztetik az eredetileg is a célnyelven született írásművektől, mivel csak fordítás nyomán született szövegekben fordulnak elő, illetve nagyobb vagy kisebb gyakorisággal szerepelnek fordításokban, mint eredeti szövegekben (Baker 1995: 235), a fordításban részt vevő nyelvpároktól függetlenül. A különböző empirikus módszerekre (Translog, hangos gondolkodtatás, interjú, képernyőrögzítés, szemkövetés, korpuszvizsgálat) épülő kutatások statisztikailag is jelentős adatokkal támasztották alá a fordítási univerzálékról alkotott hipotéziseket.

Laviosa-Braithwaite (1998) összefoglaló szócikke az *Encyclopedia of Translation Studies* fordítási univerzálékról szóló fejezetében első helyre teszi az egyszerűsítést az univerzálék sorában, megkülönböztetve egymástól a lexikai, szintaktikai és stilisztikai egyszerűsítést. A későbbiekben saját kutatásokat

is végzett az egyszerűsítés vizsgálatára (Laviosa 1998, 2000), amelyek igazolták a korábbi eredményeket. Egynyelvű összehasonlítható korpuszában (ECC – English Comparable Corpus) fordítás és nem fordítás során létrejött angol nyelvű narratív és sajtószövegeket vizsgált korpusznyelvészeti eszközökkel, és elemzései nyomán a szóhasználat négy alapvető sajátosságát azonosította a fordítási szövegben (Laviosa-Braithwaite 1998: 290):

- (1) a fordított szövegekben alacsonyabb a lexikai sűrűség;
- (2) a gyakori és kevésbé gyakori kifejezések aránya a fordítási szövegben magasabb;
- (3) egy adott fordítási korpuszban a leggyakoribb szavak gyakrabban ismétlődnek;
- (4) a fordítási szöveg leggyakoribb szavainak listája kevesebb lemmát tartalmaz.

Laviosa szógyakoriságra vonatkozó eredményeit később Xiao (2010) elemzései is alátámasztották, aki a kínai fordítások sajátosságait vizsgálva elkészítette a kutatás alapjául szolgáló alkorpuszok lexikai gyakorisági profilját. A profilok tartalmazzák a szövegekben szereplő gyakoribb szótípusok számát, a gyakori szavak számának az összes szóhoz viszonyított arányát, gyakoriságát, a gyakori és kevésbé gyakori szavak számának eloszlását.

A magyar fordítástudományban Pápai (2001, 2004) folytatott elsőként korpusznyelvészeti elemzést az explicitáció kutatására saját összeállítású, magyar–angol párhuzamos és magyar nyelvű összehasonlítható korpuszán (Arrabona). Irodalmi és társadalomtudományi írásművek fordításait vetette egybe a forrásnyelvi eredetikkel, valamint nem fordítás során keletkezett, összehasonlítható szövegekkel. Pápai megvizsgálta a szótípusok és szövegszók arányát, amely a lexikai változatosság jelzésére, összehasonlítására szolgál; minél magasabb a szótípusok aránya az összes szó számához képest, annál változatosabb a szöveg szókészlete. A statisztikai számítógépes elemzés eredményei alátámasztották Laviosa (1998, 2000) kutatási eredményeit. Pápai alacsonyabb szótípus/szövegszó arányokat talált a fordítás során létrejött szövegekben, mint az eredetileg célnyelven keletkezett művekben; a fordítások tehát kisebb lexikai változatosságot mutatnak.

Scarpa (2006) szintén korpusz alapú kutatásokat végzett az egyszerűsítés és az explicitáció fordítási szövegminőségre gyakorolt hatásait vizsgálva, angolról olaszra fordított írásművekben. A fordításokat összevetette a forrásnyelvi eredetivel a szövegek terjedelmének, a mondatok számának és átlagos hosszúságának, a standardizált szótípus/szövegszó arányoknak a megállapításával. A kutatás eredményei rámutattak az összefüggésre a fordított szövegek jobb minősége, a magasabb explicitási szint és az alacsonyabb egyszerűsítési tendencia között. Hasonlóképpen Rabadan és munkatársai (2009) is korpuszalapú kontrasztív elemzéseket végeztek a fordított szövegek minőségi értékelésére a grammatikai eszközök használatának tekintetében, a fordítási univerzálékát hívva segítségül a vizsgálatokhoz. Megállapították, hogy az alacsonyabb szövegminőségre utaló adatok egyértelműen kapcsolatba hozhatók az egyszerűsítéssel, az interferenci-

ával és a grammatikai nyelvi eszközök sajátos eloszlásával. Ezek az empirikus korpuszkutatások tehát alátámasztani látszanak Chesterman (1993) megállapítását, hogy a potenciálisan univerzális fordítási sajátosságok valójában nem mindig kívánatos jelenségek. Márpedig ha a lektor munkájának célja a fordítás minőségének javítása, elvárható, hogy visszaszorítsa a kedvezőtlen szövegjellemzőket, és a célnyelvi normához közelítve azoknak adjon nagyobb hangsúlyt, amelyek pozitívan befolyásolják a szövegminőséget.

3. A kutatás bemutatása

Jelen kutatás célja az volt, hogy megvizsgálja, mi történik szövegszinten a fordítási univerzálékkal a lektorálás eredményeként, vajon a lektor módosításai az eredetileg is a célnyelven született írásokra jellemző normához közelítik-e a fordított szövegek jellemzőit. Mivel a lektor arra törekszik, hogy a célnyelv normáinak megfelelő szöveget hozzon létre (Horváth 2009), az ekvivalencia ellenőrzésén és a hibák javításán kívül nagy valószínűséggel olyan módosításokat is végez az általa gondozott fordított szövegekben, amelyek csökkenthetik a fordítási univerzáléknak tartott nyelvi sajátosságok dominanciáját (Robin 2014b). Feltételeztem tehát, hogy a lektorálás során megfigyelhetők olyan statisztikai eszközökkel is mérhető különbségek a fordítói és a lektori kéziratok között, amelyek a fordítási univerzálék háttérbe szorulására, ugyanakkor a célnyelvre jellemző normák felerősödésére utalnak a lektorált fordításokban.

Hipotézisem bizonyítására számítógépes statisztikai elemzéseket végeztem egy lektorálási és egy magyar nyelvű összehasonlítható korpusz alapján. A lektorálási korpusz tíz teljes regény angol eredeti (961 598 szövegszó), illetve magyar nyelvű fordítói (862 420 szövegszó) és lektori (870 765 szövegszó) kéziratát foglalja magába, amelyek párhuzamos viszonyban állnak egymással. A magyar nyelvű összehasonlítható korpusz szintén tíz regény teljes szövegét öleli fel (904 009 szövegszó). A vizsgált regények a szórakoztató irodalom műfajába tartoznak, az elmúlt tíz évben jelentek meg, így a mai modern nyelvhasználatot tükrözik. A kiválasztott műfaj a művelt köznyelv normáit követi, írásbeli és szóbeli nyelvhasználatot is tartalmaz, és mivel sem a szépirodalom, sem a szaknyelv kötöttségei nem jellemzők rá, nagyobb mozgásteret nyújt a fordítóknak és a lektoroknak munkájuk végzéséhez. Ilyen módon kiváló alapul szolgál a fordító műveleteit, a fordítás általános sajátosságait vizsgáló empirikus kutatásokhoz.

A szövegeket digitális formátumban bocsátották rendelkezésemre a könyvek kiadói, amivel lehetővé vált a kéziratok számítógépes elemzése. A korpuszokat annotálás nélkül dolgoztam fel, ami lexikogrammatikai elemzéseket ugyan nem engedett meg, de bizonyos statisztikai vizsgálatokat el tudtam végezni. A teljes szövegállományokra irányuló gépi elemzések között szerepelt a szókincs változatosságára utaló szótípusok és szövegszók standardizált arányának kiszámítása, a mondatok átlagos hosszának megállapítása, az információs teher és a szókincs gazdagságát kimutató lexikai sűrűség vizsgálata, a szövegek lexikai gyakorisági profiljának elkészítése és végül a számadatok szórásának kiszámítá-

sa annak megállapítására, mennyire különböznek egymástól az egyes szövegek az eredeti, fordítói és lektori alkorpuszon, illetve összehasonlítható korpuszon belül. A statisztikai elemzéseket a Wordsmith Tools 6.0 program segítségével hajtottam végre. A kapott eredményeket rögzítettem, majd összevettem az egyes korpuszok és alkorpuszok összesített adatait az esetleges különbségek, illetve tendenciák megállapítására. A szám adatok szignifikancia- és szórásvizsgálatát az SPSS 14.0 statisztikai elemzőprogrammal végeztem el.

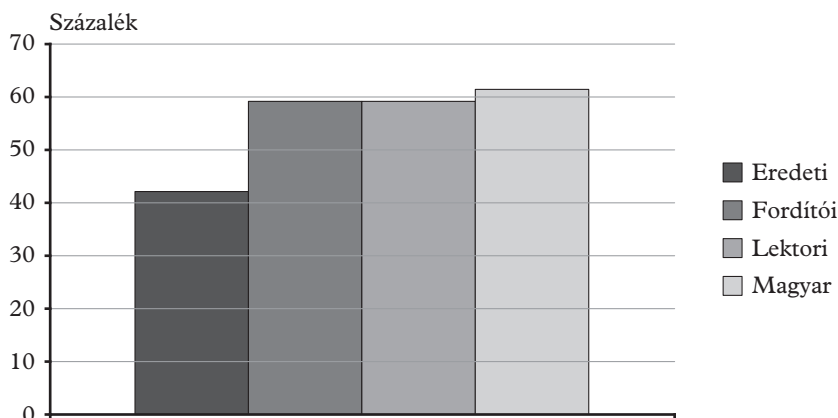
4. Eredmények

A gépi elemzés első lépésében megvizsgáltam a két korpuszt felépítő kéziratokban a szótípusok és szövegszók (*type/token*) arányát, amely adat a szókincs változatosságának mértékére utal (Laviosa 1998). Az elemzőprogram számítása szerint az angol forrásnyelvi korpusz összesített standardizált szótípus/szövegszó aránya 42,05; a magyar fordítói kéziratoké 58,93; a lektori kéziratoké 59,41; a magyar nyelvű összehasonlítható korpuszé pedig 61,8. A standardizált, vagyis ezer szóra számított értékeket az 1. ábra oszlopai illusztrálják.

Azonnal szembetűnik, hogy a szótípusok és a szóalakok aránya angolról magyarra történő fordítás és a lektorálás eredményeképpen egyaránt növekszik, jóllehet eltérő mértékben. A fordítás során mutatkozó, nagyobb mértékű növekedés nyilvánvalóan az angol és a magyar nyelv közötti eltérésekre, morfológiai-szintaktikai okokra vezethető vissza. A szótípusok és szövegszók aránya a lektorálás nyomán tovább növekszik, és bár az eltérés statisztikailag nem szignifikáns, egyértelmű tendencia látszik a szókincs változatosságának növekedésére a lektorált szövegekben. A növekedés mértéke minden bizonnyal az autentikus célnyelvi szövegekre jellemző norma felé közelít, jóllehet még nem éri el, hiszen a számítógépes elemzés a magyar nyelvű összehasonlítható korpuszban még magasabb szótípus/szóalak értéket mutatott ki.

1. ábra

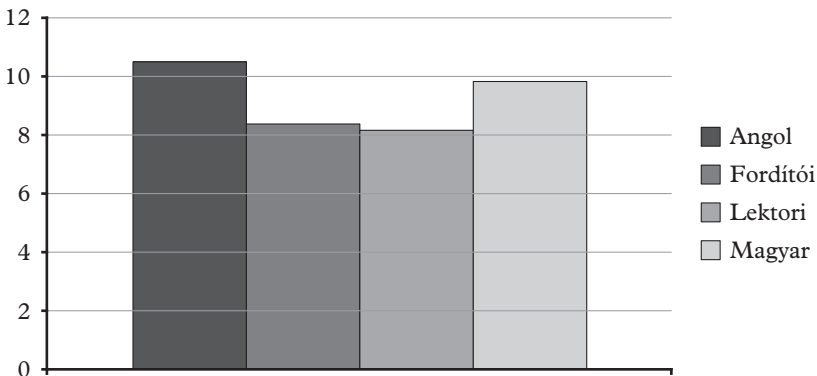
Szótípusok és szövegszók aránya



A fordítók és a fordításokat gondozó lektorok tehát együttes erőfeszítéssel dolgoznak azon, hogy a célnyelvi szöveg szókincsének változatossága megfeleljen a célnyelvi normának. Ha a fordító nem felel meg az igényeknek, a lektor közbeléphet, optimalizálva a szöveg lexikai összetételét. Mivel a szótípusok és szövegszók aránya alacsonyabbnak bizonyult a fordított, mint az eredetileg magyar nyelven íródott szövegekben, a lektorálás során viszont nőtt az arány, elmondható, hogy a lektori beavatkozások nyomán a fordítási univerzálénak tartott egyszerűsítés dominanciája csökken.

Az elemzés következő lépésében megvizsgáltam a két korpusz szövegeiben lévő mondatok számát és az átlagosan előforduló szavak adatait. Megállapítottam, hogy angolról magyar nyelvre történő fordítás során csökken a mondatok átlagos hosszúsága, ami elsősorban a két nyelv közötti rendszerbeli különbségekből fakad – ezért jelentős az eltérés a két alkorpusz adatai között, ahogyan az a 2. ábra oszlopain jól látható. Ugyanakkor a mondatok száma szintén növekszik, ami már nem feltétlenül vezethető vissza nyelvrendszerbeli, strukturális szabályokra, vagyis a fordítók a hosszú mondategységek grammatikai felbontásával, explicitáció segítségével igyekeznek megkönnyíteni a feldolgozást a befogadók számára. A lektori alkorpusz vizsgálatának eredményei szerint a lektorált szövegekben a mondatok száma – jóllehet kisebb mértékben – jellemzően tovább növekszik, miáltal a bennük szereplő szavak száma csökken.

2. ábra
Átlagos mondathosszúság



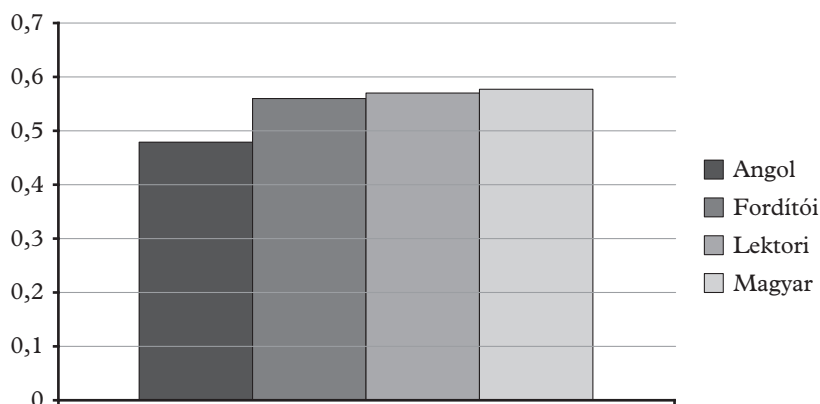
A fenti 2. ábra oszlopai jól mutatják, hogy bár az eredeti és a fordítói kéziratok között jelentős eltérés tapasztalható, a fordítói és lektori alkorpuszba tartozó szövegek között nem látszik jelentős különbség, ezt igazolta az eredmények szignifikanciavizsgálata is. A tendencia azonban világosan megmutatkozik: a lektorálás eredményeként a fordításokban tovább növekszik a mondatok száma, az átlagos mondathosszúság csökken, így igyekeznek mérsékelni a lektorok a Laviosa (1998) által megfigyelt jelenséget, miszerint a fordított szövegekben hosszabbak a mondatok, mint az eredeti célnyelvi művekben.

A lektorok tehát ebben esetben nem fordítják vissza a fordításokra jellemző explicitációs tendenciát, hanem kisebb mértékben, egyensúlyra törekedve, de szintén mondatokat bontanak fel a feldolgozhatóság érdekében. Az összehasonlítható korpuszsal való egybevetés során arra is fény derült, hogy erre nem az eredetileg is célnyelven íródott szövegekben megfigyelhető tényleges norma vezérelti őket, hiszen az autentikus magyar szövegek jóval hosszabb mondatokat tartalmaznak. Az adatok arra utalnak, hogy a fordítók és a lektorok feldolgozhatóságra való törekvése mintha egyfajta szövegszerkesztési normát követne, így próbálva megfelelni a célnyelvi befogadók *feltételezett* igényeinek, egy *vélt* célnyelvi normának.

Az elemzések során megvizsgáltam a két korpuszban szereplő kéziratok lexikai sűrűségét is, amely a szótípusok és szövegszók arányához hasonlóan a szókincs gazdagságának vizsgálatára és az információs teher megállapítására szolgál. A lexikai sűrűség az úgynevezett lexikai (tartalmas) szavak mennyiségét hasonlítja össze a grammatikai (funkció-) szavak számával. A kiszámításához a tartalmas szavak mennyiségét szükséges elosztani az összes szövegszó számával. Minél nagyobb százalékban szerepelnek tartalmas szavak egy szövegben, annál gazdagabb a szókincs, ugyanakkor magasabb az üzenet információs terhe is.

Az elemzés eredményeiből kiderült, hogy angolról magyar nyelvre történő fordítás során – a szótípusok és szövegszók arányához hasonlóan – növekszik a fordított szövegek lexikai sűrűsége, ami az angol és a magyar nyelv morfológiai-szintaktikai jellemzőiből, rendszerbeli különbségeiből fakad. Az egyes alkorpuszok számadatait a 3. ábra szemlélteti. A fordítói és a lektori kéziratokat összehasonlítva már kisebb eltérést találunk. Ahogyan első pillantásra is látszik, nincs szignifikáns különbség a fordítói és lektori alkorpusz adatai között, ugyanakkor további növekedés tapasztalható a lektorált szövegekben a lexikai sűrűség terén, ez azonban már nem vezethető vissza a fordításban részt vevő nyelvpár rendszerbeli különbségeire.

3. ábra

Lexikai sűrűség

Az adatokból megállapítható, hogy a fordításokra jellemző lexikai elszegényedés és grammatikai redundancia a lektorálás eredményeként, bár csekély mértékben, de mérséklődik, így az információs teher is megnövekszik – a lektorálás eredményeképpen tehát csökken az egyszerűsítési univerzálé jelenléte a szövegekben. Az összehasonlítható korpusz eredetileg is célnyelven keletkezett írásaiban a lexikai sűrűség magasabb értéket mutat még a másik két magyar nyelvű alkorpusz értékeinél is. Az eredmények alapján tehát megállapítható, hogy a fordított szövegek lexikai sűrűsége valóban alacsonyabb az autentikus írásművekenél, a lektorálás azonban az eredetileg is célnyelven született szövegek változatos szókincséhez, információs terhéhez közelíti a fordított szövegeket – miközben a tényleges célnyelvi normát tartja szem előtt, bár azt még nem éri el.

Laviosa (1998, 2000) egynyelvű, összehasonlítható korpuszán végzett kutatási eredményei arra mutatnak, hogy a gyakori és kevésbé gyakori kifejezések aránya a fordításokban magasabb, a gyakori szavak pedig sűrűbben ismétlődnek. A gyakorisági szólisták alapján ezért elkészítettem az eredeti, a fordítói és a lektori alkorpuszok, illetve a magyar nyelvű autentikus szövegeket tartalmazó összehasonlítható korpusz lexikai gyakorisági profilját (Xiao 2010: 207). A gyakorisági profilok megmutatják a két korpuszban található összes szövegszó számát, a leggyakoribb szótípusok és szövegszók mennyiségét, a leggyakoribb szavak számának az összes szóhoz viszonyított arányát, valamint gyakoriságát, a gyakori és a kevésbé gyakori szavak eloszlását.

1. táblázat

Lexikai gyakorisági profilok

<i>Teljes korpusz</i>	Eredeti	Fordítói	Lektori	Magyar
Összes szövegszó	1 094 083	862 420	870 765	904 009
Gyakori szavak (típusok)	122	85	85	91
Gyakori szavak (szövegszók)	601 636	317 539	316 652	255 056
Összesített arány	54,99	36,81	36,36	36,59
Gyakoriság	4931,44	3735,75	3725,31	2802,81
Gyakori/kevésbé gyakori szavak	1,2217	0,5827	0,5714	0,5771

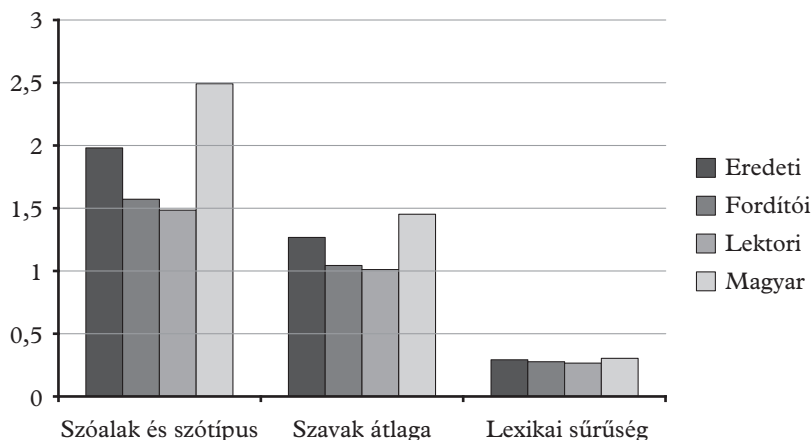
Az 1. táblázat számadatai alapján megállapíthatjuk, hogy angolról magyar nyelvre történő fordítás során csökken a gyakori szavaknak az összes szövegszóhoz viszonyított aránya, gyakorisága, illetve a gyakori és a kevésbé gyakori szavak eloszlása is; mindez részben ismét a két nyelv strukturális sajátosságainak különbségeiből, a magyar nyelvnek az angollal ellentétben szintetikus szerkesztési módjából ered. Ugyancsak csökkenő tendencia mutatkozik a lektori kéziratokban, ami azonban már nem tudható be nyelvrendszerbeli különbségeknek, inkább a lektorok azon igyekezetére utal, hogy gazdagítsák az általuk gondozott fordítások szókincsét, kerüljék a gyakran ismétlődő kifejezéseket, megfelelően a célnyelvi befogadók elvárásainak. Ismét az egyszerűsítési univerzálénak

tulajdonított lexikai elszegényedés visszaszorításának lehetünk tehát tanúi, akárcsak a szótípus/szövegszó arány és a lexikai sűrűség esetében, noha az eltérés ezúttal sem szignifikáns.

Az összehasonlítható korpusz elemzésének adatai igazolni látszanak Laviosa (1998, 2000) kutatási eredményeit: a fordítás során keletkezett szövegekben ténylegesen magasabb a gyakori és kevésbé gyakori kifejezések aránya, és a gyakori szavak sűrűbben ismétlődnek, mint az autentikus magyar nyelvű írásokban. A lektorálás mérsékli a tendenciát, sőt érdekes módon az elemzések még az eredetileg célnyelven született szövegekenél is alacsonyabb gyakorisági értékeket mutattak ki a leggyakoribb szavak összesített és kevésbé gyakori szavakhoz viszonyított arányában. A lektorok tehát a fordítások ellenőrzése, szerkesztése során az eredeti szövegek szerzőinél is jobban ügyelnek a leggyakoribb szavak arányára, bár ezek gyakorisága még mindig jóval magasabb értéket mutat.

A gépi elemzések után a lexikai sűrűség, szótípus/szövegszó arány és átlagos mondatösszeg vizsgálataival, standardizált értékek alapján nyert eredményeket szórásvizsgálatnak is alávettem. A szórás ugyanis megmutatja, hogy bizonyos változók esetében milyen mértékben térnek el az adott szövegek számadatai a középértéktől, vagyis megállapítható, mennyire különböznek egymástól az egyes szövegek az eredeti, fordítói és lektori alkorpuszon, illetve az összehasonlítható korpuszon belül. Baker (1995) a fordítási univerzálék közé sorolja az úgynevezett kiegyenlítődést is, amely a fordítási szöveg azon tulajdonságára utal, hogy egy fordítási korpuszon belül jobban hasonlítanak egymásra a szövegek a lexikai sűrűség, a szótípus/szövegszó arány és az átlagos mondatösszeg tekintetében, mint az eredetileg is célnyelven született írások.

4. ábra
Szórásvizsgálat



A gépi elemzés adatai alátámasztják a kiegyenlítődési univerzáléről szóló hipotézist, ahogyan azt a 4. ábra oszlopai is mutatják. A szórás ugyanis minden változó esetében csökkent, nemcsak angolról magyarra történő fordítás során,

hanem a lektorálás következtében is. Az angol eredeti szövegek esetében mutatkoznak a legnagyobb különbségek a lektorálási korpuszon belül. A magyar nyelvű fordítói és lektori kéziratok jobban hasonlítanak egymásra statisztikai tulajdonságaikat illetően, és ez a jelenség a lektorált változatokban a legerőteljesebb, habár a különbség most sem szignifikáns.

Ha az eredményeket az összehasonlítható korpuszsal is egybevetjük, rögtön kiderül, hogy az eredetileg is magyar nyelven született, autentikus szövegek jóval nagyobb változatosságot mutatnak statisztikai tulajdonságaikat illetően, mint akár az angol nyelvű eredeti, akár a fordított és a lektorált szövegek. A fordítás és a lektorálás tehát ebben az értelemben nem a célnyelvi normát követi; a fordítói és a lektori kéziratok egyértelműen egy bizonyos szövegépítési cél, norma felé közelítenek, ez a törekvés pedig kiegyenlítődést, optimalizációt eredményez a szövegben. Baker (1996) fordítási szövegre vonatkoztatott, kiegyenlítődési univerzáléja így a lektorálásra is érvényes; makroszinten jelentkező, szövegszerkesztési stratégia részeként működik.

5. Összefoglalás

A fentiekben bemutatott kutatás kiindulópontja az a megfigyelés volt, miszerint a fordítási szöveg általános jellemzőit, univerzáléit vizsgáló kutatások olyan szövegeket is bevonnak az elemzésekbe, amelyek a lektor módosításainak nyomait is magukon viselik, így nem világos, hogy az úgynevezett univerzálék, vagyis a fordítási szöveg sajátos nyelvi mintázatai ténylegesen a fordítási folyamatnak tulajdoníthatók-e. Jelen korpuszkutatás arra a kérdésre kereste a választ, mi történik szövegszinten a fordítási univerzálékkal a lektorálás eredményeként, vajon a lektori módosítások az eredetileg is a célnyelven született írásokra jellemző normához közelítik-e a fordított szövegek jellemzőit.

A számítógépes elemzések számadatai elsősorban igazolni látszanak Laviosa (1998, 2000), Pápai (2001, 2004), Scarpa (2006) és Xiao (2010) korábbi kutatási eredményeit az univerzálék közé sorolt egyszerűsítés és az explicitációs tendencia fordított szövegekben való előfordulását illetően. A lektorálási korpusz fordítói kézírataiban az autentikus magyar szövegekhez képest alacsonyabb szótípus/szövegszó arányt és lexikai sűrűséget találtam. A fordítás során keletkezett szövegekben továbbá ténylegesen 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, 2000) által megfigyelt alapvető lexikai sajátosságokat.

Jelen kutatás eredményei azonban azt is bizonyítják, hogy kimutathatók bizonyos statisztikai eltérések a fordítói és lektori kéziratok között, habár a különbségek többnyire nem szignifikánsak. Érdeemes megjegyezni, hogy ha a különbségek szignifikánsak volnának, már nem is harmadlagos szövegalkotásról, lektorálásról volna szó, hanem újrafordításról. Megfigyeltem, hogy a szótípusok és szövegszók arányát és a lexikai sűrűséget, tehát a szókincs gazdagságát, változatosságát illetően a szöveg javítását célzó lektorálás a célnyelvi normához közelíti a fordítást, csökkenti a gyengébb szövegminőséget eredményező lexikai

elszegényedést, mérsékli a fordítási sajátosság dominanciáját – jöllehet a lektorált kéziratok adatai még így sem érik el az autentikus magyar szövegek statisztikai számértékeit, ezért lehetséges, hogy a lektorált szövegekben is kimutatható az egyszerűsítés.

A gyakori szavak ismétlése, az átlagos mondathosszúság és a kiegyenlítő-dés terén azonban a fordításban is jelentkező, de még a célnyelvi normán is túlmutató lektorálási tendenciának lehetünk tanúi. A lektori kéziratok statisztikai adatainak szórása kevesebb a fordítói kéziratokéhoz és magyar nyelvű autentikus szövegekhez képest is, így erőteljesebb kiegyenlítő-dés tapasztalható; a mondatok feldolgozást segítő, explicitációval járó felbontásával tovább mérséklődik a mondatok hosszúsága; a leggyakoribb szavak összesített és kevésbé gyakori szavakhoz viszonyított aránya még az eredeti magyar szövegeknél is alacsonyabb értéket mutat – ezek a sajátosságok azonban nem befolyásolják kedvezőtlenül a szövegek minőségét, a lektor optimalizációval ügyel az egyensúlyra.

A jelenség magyarázatát abban kereshetjük, hogy a lektorált fordítások valójában kétszeres „szerkesztésen” esnek át, először a fordítás, majd a lektorálás folytán (Robin 2014b), miközben a nyelvi közvetítők a kommunikációs kokárdákat kerülve a befogadók *feltételezett* elvárásainak megfelelően igyekeznek megalkotni a szöveget, néha már babonásan. Egyfajta *szerkesztési* normát követnek, amely nem feltétlenül esik egybe az eredetileg is célnyelven keletkezett írásművekben fellelhető tényleges célnyelvi szövegalkotási normákkal. Szöveg-szerkesztési sajátosságról van szó, amely bár a fordításban is érezteti hatását, a lektorálásban még erőteljesebben jelentkezik.

Irodalom

- Baker, M. 1993. Corpus Linguistics and Translation Studies: Implications and Applications. In: Baker, M., Francis, G., Tognini-Bonelli, E. (eds) *Text and Technology: In Honour of John Sinclair*. Amsterdam/Philadelphia: Benjamins. 233–250.
- Baker, M. 1995. Corpora in Translation Studies. An Overview and Suggestion for Future Research. *Target* Vol. 7. No. 2. 223–245.
- Chesterman, A. 1993. From ‘Is’ to ‘Ought’: Laws, Norms and Strategies in Translation Studies. *Target* Vol. 5. No. 1. 1–20.
- Horváth P. I. 2009. *A lektori kompetencia*. Doktori értekezés. Kézirat. Budapest.
- 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.
- Laviosa, S. 2000. TEC: a Resource for studying what is „in” and „of” translational English. *Across Languages and Cultures* Vol. 1. No. 2. 159–177.
- Laviosa, S. 2009. Universals. In: Baker, M. (ed.) 2009. *Encyclopedia of Translation Studies*. London: Routledge. 306–311.
- Laviosa-Braithwaite, S. 1998. Universals of Translation. In: Baker, M. (ed.) *Routledge Encyclopedia of Translation*. London: Routledge. 288–291.
- Pápai V. 2001. *Az explicitációs hipotézis vizsgálata angol–magyar és magyar–magyar párhuzamos korpuszok egybevetésével*. Doktori értekezés. Kézirat. Pécs–Győr: PTE–SZIE.
- Pápai, V. 2004. Explicitation: A universal of translated text? In: Mauranen, A., Kujamaki, P. (eds) *Translation Universals. Do they exist?* Amsterdam: Benjamins. 143–164.

- Rabdan, R., Labrador, B., Ramon, N. 2009. Corpus-based contrastive analysis and translation universals. A tool for translation quality assessment. *Babel* Vol. 55. No. 4. 303–328.
- Robin E. 2014a. Lektorált fordítások és eredeti magyar szövegek gépi összehasonlítása. Elhangzott: ANYK Konferencia. *Nyelv, kultúra és társadalom*. Budapest: Kodolányi János Főiskola. (2014. november 3–4.)
- Robin E. 2014b. Fordítási univerzálék a lektorált fordításokban. Budapest. Doktori értekezés.
- Scarpa, F. 2006. Corpus-based Quality-Assessment of Specialist Translation: A Study Using Parallel and Comparable Corpora in English and Italian. In: Gotti, M., Sarcevic, S. (eds) *Insights into specialized translation*. Bern/Frankfurt: Peter Lang. 155–172.
- Xiao, R. 2010. How different is translated Chinese from native Chinese? A corpus-based study of translation universals. *International Journal of Corpus Linguistics* Vol. 15. No. 1. 5–35.

Források

Lektorálási korpusz

- Bray, L. 2010. *Lázadó angyalok*. Szeged: Könyvmolyképző.
- Cashore, K. 2010. *A garabonc*. Szeged: Könyvmolyképző.
- Clare, C. 2009. *Csontváros*. Szeged: Könyvmolyképző.
- Diamand, E. 2009. *Kalózsok nyomában*. Budapest: Agave.
- Gray, C. 2009. *Evernight – Örökéj*. Szeged: Könyvmolyképző.
- Harris, J. 2009. *Rúnajelek*. Budapest: Ulpius.
- Penney, S. 2010. *Gyengéd, mint a farkasok*. Szeged: Könyvmolyképző.
- Pullman, Ph. 2008. *Árny északon*. Pécs: Alexandra.
- Stiefvater, M. 2009. *Shiver – Borzongás*. Szeged: Könyvmolyképző.
- Westerfield, S. 2007. *Csúfok*. Szeged: Könyvmolyképző.

Összehasonlítható korpusz

- Bene V. 2006. *Eszterlánc*. Budapest: Európa.
- Benina 2010. *A Boszorka fénye*. Szeged: Könyvmolyképző.
- Benina 2013. *A Boszorka démona*. Szeged: Könyvmolyképző.
- Bosnyák V. 2011. *Lovessence*. Szeged: Könyvmolyképző.
- Böszörményi Gy. 2002. *Gergő és az álomfogók*. Szeged: Könyvmolyképző.
- Böszörményi Gy. 2009. *3... 2... 1...* Szeged: Könyvmolyképző.
- On Sai 2012. *Calderon*. Szeged: Könyvmolyképző.
- Spirit B. 2010. *Árnyékvilág*. Szeged: Könyvmolyképző.
- Szurovecz K. 2011. *Gyémántfiú*. Szeged: Könyvmolyképző.
- Szurovecz K. 2011. *Fényemberek*. Szeged: Könyvmolyképző.