Fókuszban a fordított és a tolmácsolt szöveg

KORPUSZALAPÚ FORDÍTÁSKUTATÁS MAGYARORSZÁGON

Szerkesztette:

ROBIN EDINA ÉS SEIDL-PÉCH OLIVIA

Fókuszban a fordított és a tolmácsolt szöveg

KORPUSZALAPÚ FORDÍTÁSKUTATÁS MAGYARORSZÁGON

Szerkesztette:

ROBIN EDINA ÉS SEIDL-PÉCH OLIVIA

Segédkönyvek a nyelvi közvetítésről

Fókuszban a fordított és a tolmácsolt szöveg

Korpuszalapú fordításkutatás Magyarországon

Szerkesztők:

Robin Edina és Seidl-Péch Olívia

Szakmai lektor: Zachar Viktor Kristóf

Korrektúra: Juhász Dóra, Sulyok Kamilla, Zeman Gabriella

Borítóterv: Berki Éva

Tördelés, nyomdai előkészítés: Communication And Design Stúdió Kft.

ISBN 978-615-00-7338-5

DOI: https://doi.org/10.36252/Nyelvkozvsegedkonyv1

Közreadó:

Eötvös Loránd Tudományegyetem Bölcsészettudományi Kar, Nyelvtudományi Doktori Iskola, Fordítástudományi Doktori Program

Magyar Alkalmazott Nyelvészek és Nyelvtanárok Egyesülete, Fordítástudományi Szakosztály

A kiadást az Eötvös Loránd Tudományegyetem Hallgatói Önkormányzata támogatta.

Budapest, 2020

Tartalomjegyzék

Recenziók

Korpuszalapú fordításkutatás Magyarországon

Bevezetés

Robin Edina és Seidl-Péch Olívia

robin.edina@btk.elte.hu

Eötvös Loránd Tudományegyetem Bölcsészettudományi Kar

Fordító- és Tolmácsképző Tanszék

olivia@inyk.bme.hu

Budapesti Műszaki és Gazdaságtudományi Egyetem

Idegen Nyelvi Központ

Kivonat: A korpuszalapú fordításkutatás az utóbbi évtizedek egyik legjelentősebb és legtermékenyebb kutatási paradigmájává vált. A fordítási folyamat és a fordítási szöveg sajátosságainak leírását célzó vizsgálódásoknak mostanra már nélkülözhetetlen segítséget nyújtanak a számítógépes korpusznyelvészeti eszközök, ezeken belül is a nyelvi közvetítés eredményeképpen létrejött szövegeket tartalmazó, nagyméretű digitális korpuszok, amelyek gépi elemzése statisztikailag jelentős adatokkal támaszthatja alá a kutatási eredményeket. A technika fejlődése és az egyre modernebb eszközök újabb lendületet adtak a korpuszalapú vizsgálatoknak nem csupán a nemzetközi színtéren, hanem a magyarországi fordításkutatók körében is.

Kulcsszavak: fordítási szöveg, korpuszalapú fordítástudomány, megakorpusz, multimodális korpusz, Pannónia Korpusz

Robin Edina és Seidl-Péch Olívia: Korpuszalapú fordításkutatás Magyarországon. In: Robin E., Seidl-Péch O. (szerk.) 2020. *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordításkutatás Magyarországon.* Segédkönyvek a nyelvi közvetítésről I. Budapest: ELTE BTK Fordítástudományi Doktori Program, MANYE Fordítástudományi Szakosztály. DOI: https://doi.org/10.36252/Nyelvikozvsegedkonyv1.1

1. A hazai kutatási körkép

A magyar korpuszalapú fordítástudományi kutatások gyakorlatilag kis időbeli elmaradással tartottak lépést a nemzetközi trendekkel, és a múlt század 90-es éveitől kezdve jelentek meg folyamatosan. Klaudy Kinga 1987-es, autentikus és oroszról fordított magyar célnyelvi szövegek összehasonlító korpuszán végzett kutatása (Klaudy 1987) még manuális elemzésre támaszkodott, ugyanakkor Kohn János egy évtizeddel később már szakirodalmi leírások, valamint német és angol minták alapján építette fel számítógéppel olvasható és kutatható, párhuzamos szövegeket tartalmazó fordítási korpuszát. Kutatásának fókuszában a fordítás nyelvészeti és irodalmi megközelítésének összekapcsolása állt (Kohn 1997). A Magyar Tudományos Akadémia Nyelvtudományi Intézetében is megindultak a korpusznyelvészeti kutatások, és az első iniciatívák között szerepelt az a projekt, amely más európai nyelvészeti intézményekhez kapcsolódva 1999-ben létrehozta Orwell *1984* című műve alapján az angol–magyar párhuzamos korpuszt. A Nyelvtudományi Intézet honlapján a mai napig elérhető és gépileg kereshető az *Orwell: 1984 – angol–magyar párhuzamos korpusz*[1], amely lehetővé teszi egy-egy angol kifejezés összes előfordulásának és azok magyar fordításának listázását.

Az ezredfordulót jellemző technológiai fejlődés hazánkban is lehetővé tette a korpuszalapú fordítástudományi kutatások elterjedését és az egyes kutatók személyi számítógépén létrehozott korpuszok megjelenését. Az első fordítástudományi korpuszalapú doktori értekezést Pápai Vilma írta 2001-ben (Pápai 2001). Az általa vizsgált, összesen 45 ezer szövegszóból álló Arrabona korpusz angol–magyar párhuzamos és összehasonlítható szövegeket tartalmazott. Kutatásának középpontjában az explicitációs hipotézis vizsgálata állt, miszerint a fordítás eredményeként keletkezett szövegek általában explicitebbek a forrásnyelvi szövegeknél, azaz „a fordítók felerősítik a forrásnyelvi szöveg logikai kapcsolatait és utalásrendszerét, kiegészítik az elliptikus szerkezeteket, magyarázó betoldásokat alkalmaznak a forrásnyelvi kultúrába ágyazott fogalmak, eszmék, események, az

1 Orwell: 1984 - angol-magyar párhuzamos korpusz.
Elérhető: http://corpus.nytud.hu/demo/infotrend/orwell/

ún. reáliák stb. megvilágítására" (Pápai 2001: 4). Pápai projektjével szinte párhuzamosan indult meg a Szent István Egyetem Nyelvi Intézetében az első angol–magyar párhuzamos szövegeket tartalmazó szaknyelvi korpusz építése (Heltai 2007). A korpusz mérete átlépte az egymillió szövegszavas határt (1,1 millió szó), és a kutatások elsősorban a szakfordítások jellegzetességeinek vizsgálatát tették lehetővé, illetve kiszolgálták a szakfordítóképzés céljait és a fordítói kompetenciák feltárását, mivel a korpuszba gyűjtött szövegek részben hivatásos fordítóktól, részben szakfordító hallgatóktól származtak. A Szegedi Tudományegyetem keretei között zajlott az orosz–magyar kétirányú párhuzamos és összehasonlító korpusz építése (Szabó et al. 2011), amely mindkét nyelven tartalmaz autentikus és fordítás útján keletkezett szövegeket. A 130 ezer szövegszóból álló, szépirodalmi, tudományos, és hivatalos szövegeket tartalmazó HunOr korpusz annotálását (mondatokra bontás, tulajdonnévi annotáció) és mondatszintű párhuzamosítását a kutatási célok pontosabb végrehajtása érdekében valósították meg.

A második magyar nyelvű, korpuszalapú fordítástudományi disszertáció 2004-ben született meg. Balaskó Mária a fordításnyelv vizsgálata során arra helyezte a hangsúlyt, hogy a korpuszelemzés segítségével a konkrét nyelvi adatok alapján írja le a fordítási jelenségeket, és a kimutatott szabályszerűségek mentén vonja le az általánosítható következtetéseket (Balaskó 2004). Ezek után a 2010-es évektől kezdve egymás után jelentek meg a korpuszalapú fordítástudományi kutatások az ELTE Fordítástudományi Programjának keretében. A kutatások olyan témakörökkel foglalkoztak, mint a filmfordítás (Polcz 2012, Somodi 2014), a fordítási kompetencia (Makkos 2015), a géppel támogatott fordítás (Lengyel 2013, Ábrányi 2016), az információs szerkezet (Nagy 2015), az interkulturális kommunikáció (Kovács 2015), a lektorálás (Mohácsi-Gorove 2014, Robin 2014), a lexikai kohézió (Bozsik 2015, Seidl-Péch 2011a) és a tolmácsolás (Sato 2014). A felsorolt kutatók mindegyike létrehozta a saját kutatási céljainak megfelelően összeállított korpuszokat, ugyanakkor egyre sürgetőbb igénnyé vált egy gépileg elemezhető, akár közösen is kutatható fordítástudományi referenciakorpusz építése.

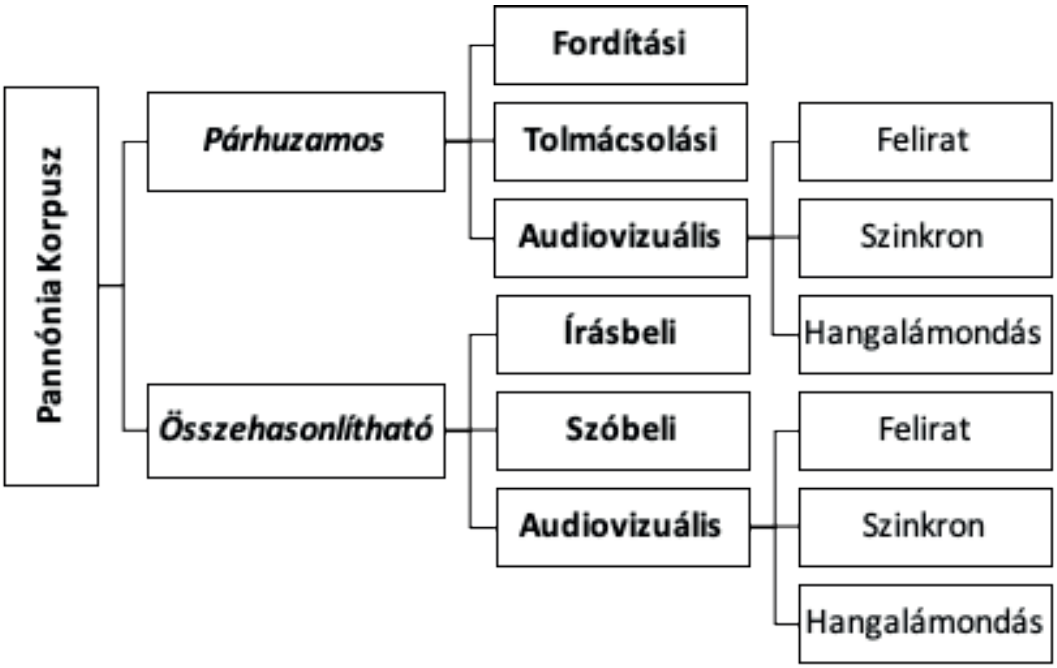
2. A magyar fordítási korpuszprojekt

Néhány évvel ezelőtt egy, az Eötvös Loránd Tudományegyetem Fordítástudományi Doktori Programján belül elindított projekt tűzte ki célul egy több millió szövegszóból álló magyar fordítási megakorpusz felállítását, illetve szorgalmazta a korpusz szövegállományán végzett kutatásokat a fordítási nyelvhasználat átfogó vizsgálatára és a korpuszalapú fordítástudomány magyarországi támogatására. A munkában részt vevő kutatócsoport egy olyan digitális szövegtár felállítását irányozta elő, amely megkönnyíti a fordításkutatók munkáját, nagy mennyiségű szöveget kínálva kvantitatív és kvalitatív kutatásokhoz egyaránt. Egy efféle szövegkorpusz lehetőséget nyújthat hipotézisek tesztelésére és újabb kutatási hipotézisek felállítására, új eredményekkel gazdagítva a tudományt, és hozzájárul a diszciplína fejlődéséhez. Az ebből a megfontolásból alapított Pannónia Korpusz olyan számítógépen tárolható szövegtár, amely teljes mértékben megfelel a mai modern korpuszokkal szemben támasztott kritériumoknak. Multimodális korpusz, tehát fordított, tolmácsolt és audiovizuális szövegeket egyaránt magába foglal, párhuzamos és összehasonlítható komponenst is tartalmaz, lehetőséget nyújtva a magyar fordítási piacon folyó, sokoldalú fordítási tevékenység kutatására, beleértve a lektorálást is (Robin et al. 2016).

A párhuzamos korpusz idegen nyelvről, egyelőre főként angolról magyar nyelvre és magyarról idegen nyelvre fordított természetes szövegeket tartalmaz. A Pannónia Korpusz bizonyos mértékig tehát kétirányú korpusznak tekinthető, mivel a magyarról idegen nyelvre fordított szövegek nem csupán a párhuzamos, hanem az összehasonlítható főkorpuszba is beletartoznak. Az összehasonlítható korpusz eredetileg is magyar nyelven született, írásbeli és szóbeli autentikus magyar szövegeket tartalmaz, felépítését tekintve a fordítási, tolmácsolási és audiovizuális alkorpuszra tagolt párhuzamos főkorpuszt követi. A korpusz tervezésekor a projekt résztvevői egy olyan magyar nyelvű fordításokat és autentikus szövegeket tartalmazó digitális gyűjtemény összeállítását tűzték ki célul, amely a modern nyelvhasználatot tükrözi – 2000 után keletkezett természetes szövegeket válogatva a gyűjteménybe –, ilyen módon több millió szavas méretével, számos szövegfajtát felvonultató sokszínűségével és reprezentatív összetételével gondoskodik a jövőbeli fordítástudományi kutatások eredményeinek általánosíthatóságáról.

1. ábra

A Pannónia Korpusz összetétele



A Pannónia Korpusz jelenlegi állapotában több mint 15 millió szövegszót számlál, minden alkorpuszában tartalmaz fordítás eredményeként született, illetve autentikus szövegeket is. Szövegfajták tekintetében a korpusz az eredeti célkitűzésekhez képest tovább bővült, immár tíz témában kínál kutatásra alkalmas szövegeket a vizsgálatok számára, ahogyan azt a 2. ábra felsorolása is tükrözi. A korpusz legújabb komponense az egészségtudományi alkorpusz – nem meglepő módon, hiszen az orvosi szövegek vizsgálata friss kutatási területként jelentkezik a fordítástudományban.

A korpuszalapú fordítástudományi kutatások egyik kritikus pontja a korpuszépítés nehézsége, az adatgyűjtés. Baker (1995) kifejti, hogy a korpuszok felépítéséhez több okból is nagyon nehéz megfelelő forrásokat találni. Először is a fordítók többnyire gyanakodva fogadják a felkérést, hogy fordított szövegeket bocsássanak a kutatók rendelkezésére, tartva az esetleges értékeléstől.

2. ábra

A Pannónia Korpusz alkorpuszai

Gyakran nem is nyilvánvaló számukra, milyen előnnyel jár munkájukra nézve a kutatás támogatása, ráadásul számos szövegfajta esetében titoktartási szerződés korlátozza a szövegek elérhetőségét. További nehézséget jelent, hogy a szöveggyűjtés rengeteg időt és energiabefektetést igénylő vállalkozás. Mindez korlátozza a korpuszok fejlődését. Ez alól a Pannónia Korpusz sem kivétel; mindazonáltal a folyamatos bővítés következtében továbbra is fokozatosan növekszik a mérete.

A 2018/2019-es tanévben az ELTE Fordító- és Tolmácsképző Tanszék mesterszakos hallgatóinak közreműködésével sikeresen befejeződött az intermodális tolmácsolási korpusz fejlesztése az Európai Parlament interneten elérhető adatbázisából származó eredeti, tolmácsolt és fordított szövegek gyűjtésével. A tolmácsolási alkorpusz ilyen módon számos kutatási lehetőséget kínál a magyar, francia, angol és német nyelv viszonylatában (Szegh 2017, Götz 2017), párhuzamos és összehasonlítható elemzésekhez egyaránt.

Az audiovizuális alkorpusz (Robin és Szegh 2018) ugyancsak jelentős fejlődésen ment keresztül; angol–magyar nyelvi irányban befejeződött a szövegek, mozifilmek, sorozatok és dokumentumfilmek feliratának, valamint szinkronjának gyűjtése és transzkrip-

ciója, illetve folyamatban van az audiovizuális szövegek feldolgozása német–magyar és francia–magyar nyelvi irányban is. Az audiovizuális szövegek manuális átírásához szintén segítséget kapott a Pannónia Korpusz az ELTE Fordító- és Tolmácsképző Tanszékétől. Először a 2016/2017 őszi félévében meghirdetett közismereti kurzus keretén belül végezték mesterszakos hallgatók az átírást a korpuszprojekt tagjainak irányításával. A magyar nyelvű autentikus audiovizuális tartalmak esetében komoly nehézségek merültek fel: magyar nyelvről csak ritkán készül angol, német vagy francia szinkron, így csupán a feliratok szövegéből építhető kétirányú párhuzamos korpusz. Jelenleg a fordítási alkorpusz építése folyik egészségtudományi témájú szakszövegek gyűjtésével; és bár a szövegtár továbbra is folyamatos fejlesztés alatt áll, máris lehetőséget kínál a kutatásokra.

3. A tanulmánykötet felépítése

A jelen kötet a magyar fordítástudomány berkeiben jelenleg folyó, elsősorban a fentiekben bemutatott Pannónia Korpusz szövegeit felhasználó, korpuszalapú kutatásokba igyekszik bepillantást nyújtani elméleti kérdések tárgyalásával és empirikus kutatások bemutatásával. A kutatások a korpusz mindkét főkorpuszából – párhuzamos és összehasonlítható –, valamint összes alkorpuszából – fordítási, tolmácsolási és audiovizuális – merítenek a vizsgálatokhoz. A kötet tanulmányai áttekintik a korpusznyelvészeti vizsgálatok fő elméleti és módszertani dilemmáit, rávilágítanak a feltárt normák és univerzális fordítási sajátosságok különbségeire, valamint érdekes ismeretekkel és eredményekkel gazdagítják a korpuszalapú fordításkutatás iránt érdeklődők tudását a fordítás és tolmácsolás eredményeként keletkezett szövegek általános nyelvi jellemzőiről és a nyelvi közvetítők stratégiáiról.

A kötet első tanulmányának szerzője a magyar fordítástudomány megteremtője és az ELTE Fordítástudományi Doktori Programjának vezetője, **Klaudy Kinga** professzor asszony, aki a hazai fordításkutatók körében úttörőként alkalmazta a párhuzamos és az összehasonlítható korpuszokat vizsgálati eszközként (Klaudy 1987, 2005). *Visszatekintés egy korai korpuszelemzésre* című tanulmányában egy olyan kutatást elevenít fel, amelyet

1980-ban nyújtott be kandidátusi értekezésként. A szövegek elemzését akkor még manuálisan végezte, de az eredmények ma is érvényesek. Oroszból fordított magyar szakmai és tudományos szövegeket hasonlított össze hasonló időben keletkezett és hasonló témájú autentikus magyar szövegekkel. A vizsgálat középpontjában bizonyos grammatikai mutatók (a mondatok és a mondategységek hosszúsága, szerkesztettsége, mélysége, bővítettsége), valamint bizonyos aktuális tagolási típusok álltak (visszautaló téma, informatív téma, retorikus téma, emelkedő réma, ereszkedő réma, egyenletes réma). A kutatás eredményei azt mutatják, hogy a vizsgált grammatikai mutatók és aktuális tagolási típusok gyakoriságában eltérések mutatkoznak a fordított magyar szövegek és az autentikus magyar szövegek között, és ezen eltérő disztribúciós jegyek alapján a fordított szövegek kvázi-helyes magyar szövegnek nevezhetők.

A tanulmánykötetben szereplő második cikk szerzője **Balaskó Mária**, akit ugyancsak a magyarországi korpuszalapú fordítástudomány úttörőjeként tartunk számon. Jelenleg az ELTE Savaria Egyetemi Központ oktatója, jelentősebb kutatási területei között a fordítástudomány és a korpusznyelvészet kapcsolata mellett megjelenik a lexikális grammatika és az interkulturális kommunikáció is. *Korpuszalapú fordítástudomány, és ami mögötte van* című tanulmányában részletesen ismerteti a brit nyelvészeti tradíció meghatározó elveit, illetve bemutatja, miként járult hozzá a fordítástudomány leíró paradigmáján belül a korpuszalapú fordításkutatás kialakulásához. A cikk amellett érvel, hogy a korpusznyelvészetben kidolgozott analitikus eszközöket alkalmazni lehet a fordítás eredményeként keletkezett szöveg általános nyelvi mintázatainak vizsgálatára.

A kötetben szereplő következő tanulmány **Seidl-Péch Olívia** írása, aki a Budapesti Műszaki és Gazdaságtudományi Egyetem Idegen Nyelvi Központjának oktatója. A korpusznyelvészeti kutatások sorában először végezte el szövegek lexikai-szemantikai tulajdonságának, nevezetesen a lexikai kohézió vizsgálatának feltárását nyelvtechnológiai eszközök segítségével (Seidl-Péch 2011b). További kutatási területei közé tartozik elsősorban fordított szakszövegek esetében e szövegek lexikai sajátosságainak, különös tekintettel a terminológia menedzseléssel kapcsolatos preferenciáinak vizsgálata. Jelen tanulmánya *Korpuszalapú fordításkutatás: lehetőségek és nehézségek. Fókuszban a kor-*

puszépítés és a korpuszalapú elemzés címmel azokat a legfontosabb szempontokat kíséreli meg számba venni, amelyek figyelembevétele elengedhetetlen a korpuszalapú kutatások tervezésénél és a korpuszok összeállításánál. Az összegyűjtött szempontrendszer mintegy iránymutatásként is szolgálhat a jövőbeni kutatások, korpuszépítők munkájához.

Az empirikus kutatásokról beszámoló tanulmányok sorában az első **Pápai Vilma**, a Széchenyi István Egyetem egykori oktatójának mostanáig csupán angol nyelven elérhető 2004-es cikke, amely eredetileg Anna Mauranen és Pekka Kujamäki *Translation Universals. Do they exist?* című tanulmánykötetében jelent meg a John Benjamins könyvkiadó gondozásában. A tanulmányt, amelyben a szerző disszertációja legfontosabb eredményeit foglalja össze, immár magyarul olvashatjuk az ELTE Fordító- és Tolmácsképző Tanszék mesterszakos hallgatóinak, Heltai Edit Fruzsinának, Juhász Dórának, Kis Elizabetnek, Nagy Laurának és Szász Barnának fordításában, *Explicitáció – a fordítás univerzális jellemzője?* címmel. A szerző korpuszalapú vizsgálat alapján járja körül az explicitáció jelenségét, amelyre a szakirodalom a fordítási szöveg univerzális jellemzőjeként hivatkozik. A kutatás egy fordított szövegeket tartalmazó angol–magyar nyelvű párhuzamos korpusz, valamint egy autentikus magyar szövegekből álló összehasonlítható korpusz alapján végzett kettős elemzésre épül, rámutatva a különbségre az explicitáció mint fordítási művelet és az explicitség mint szövegsajátosság között. Fényt derít továbbá arra a kérdésre, vajon van-e szorosabb összefüggés az explicitáció és az ugyancsak fordítási univerzálénak tartott egyszerűsítés között.

A kötet második tanulmánya **Mány Dániel** írása, aki az ELTE BTK fordító–tolmács mesterszakos hallgatójaként szerzett diplomát, jelenleg az ELTE FTT és az SE ETK megbízott oktatójaként francia fordítást és fordítói projektmunkát tanít. Az ELTE Fordítástudományi Doktori Programjának hallgatójaként kutatási területe az orvosi szakfordítás, különös tekintettel a betegtájékoztatókra, továbbá a Pannónia Korpusz egészségtudományi alkorpuszának fejlesztőjeként is tevékenykedik. A kötetben olvasható, *Idegen szavak félautomatikus elemzése autentikus és fordított orvosi szövegekben: fordítási stratégiák angolról franciára és magyarra fordított betegtájékoztatók tükrében* című feltáró tanulmányának középpontjában az idegen szavak jelensége áll, azon belül is az orvostu-

dományi terminusok előfordulása és fordítása autentikus angol, illetve angolról magyarra és franciára fordított betegtájékoztatók párhuzamos korpuszában. Az idegen szavak fordításának elemzése során kísérletet tesz a fordítók stratégiáinak feltérképezésére, amelyek segítségével a nyelvi közvetítők a forrásnyelvi szöveg célnyelvre való átültetésekor próbálnak megfelelni a laikus olvasók, illetve az egészségügyi személyzet elvárásainak.

A kötet soron következő cikkének szerzője **Dankó Szilvia**, az ELTE Fordítástudományi Doktori Programjának hallgatója és a Pannónia Korpuszprojekt egyik alapító tagja. Korábbi kutatásai során a célnyelv-specifikus egyedi elemek alulreprezentáltságáról szóló hipotézist (Tirkkonen-Condit 2004) vizsgálta az angol és a magyar nyelv viszonylatában (Dankó 2017). Jelen tanulmányában Chesterman (2007) javaslatát követve nem *a priori*, hanem a fordított és a nem fordított szövegek korpuszalapú elemzésének eredményeképpen létrejövő gyakorisági minta alapján kísérli meg azonosítani a magyar nyelvre jellemző egyedi elemeket – *a posteriori*. Ugyanis ha bizonyos szavak előfordulási gyakorisága eltér a fordított és az autentikus szövegeket tartalmazó korpuszban, akkor feltételezhetően előfordulnak alulreprezentált, célnyelv-specifikus „egyedi" nyelvi elemek is a gyakorisági különbségeket felmutató szólistában.

Nagy Annamária Lilla az ELTE BTK germanisztika alapszakosaként, majd fordító–tolmács mesterszakosaként szerzett diplomát. Évek óta dolgozik fordítóként és tolmácsként magyar, német és angol munkanyelven, továbbá nemzetközi szinten tevékenykedik szabadúszó újságíróként és fotóriporterként a könnyűrepülés világában. Az ELTE Nyelvtudományi Doktori Iskolájának Fordítástudományi Doktori Programjában kutatási területe a műszaki szakfordítás sajátosságainak feltárása új módszerekkel (Nagy 2017), elsősorban a Pannónia Korpusz fordított és autentikus szövegeit vizsgálva, amelynek alapításában és fejlesztésében is szerepelt vállalt. Jelen tanulmányában – *A szógyakoriság gépi vizsgálata műszaki szövegeken: terminusok és műszaki formulaszerű elemek* – a műszaki szövegek szógyakorisági jellemzőit igyekszik feltárni, választ keresve a kérdésre, vajon milyen hatást gyakorolnak a műszaki szövegek szógyakoriságára a terminusok, a műszaki szövegkörnyezet formulaszerű kifejezései.

Bakaja Zoltán *Az indiai vallási irodalomban használatos beszélőjelölő igék magyar fordítása* című írása az utolsó a jelen kötet cikkeinek sorában, amely a nyelvi közvetítés írásbeli módjának szövegsajátosságait vizsgálja. A tanulmány szerzője 1994-ben szerzett mélyhegedűművész, tanár képesítést a Liszt Ferenc Zeneművészeti Főiskola Egyetemi Tagozatán. 2011-ben a Bhaktivedanta Hittudományi Főiskola vaisnava teológia, 2014-ben vaisnava jógamester alapszakán szerzett diplomát. 1997 óta tanít, 2012-től a Bhaktivedanta Hittudományi Főiskola oktatójaként. Jelenleg az ELTE Irodalomtudományi Doktori Iskolájának hallgatója, készülő disszertációja témájául a *Bhagavad-gítá* magyar nyelvű fordításainak bemutatását választotta. A cikk a *Bhagavad-gītā* beszélőjelölőjében található ige magyar fordítási lehetőségeit vizsgálja. Számba veszi a mostanáig megjelent kiadások megoldásait, és a Magyar Nemzeti Szövegtár, valamint a Pannónia Korpusz fordított szövegeinek segítségével keres választ arra a kérdésre, hogy egy modern magyar kiadásban vajon melyik magyar ige lehetne a legjobb választás a beszélőjelölő idéző igéjének fordítására.

A szóbeli és az írásbeli nyelvi közvetítés határmezsgyéjén elhelyezkedő audiovizuális fordítás sajátosságainak kutatásával foglalkozik **Malaczkov Szilvia**, a Budapesti Gazdasági Egyetem oktatója, aki egyrészt angol üzleti, diplomáciai és média szaknyelvet, másrészt angol–magyar, magyar–angol szakfordítást tanít. Az ELTE BTK Fordítástudományi doktori programjának doktori hallgatójaként az önkéntes feliratozás jellemzőit vizsgálja nyelvészeti megközelítésű elemzésekkel. Ebben a kutatási témában született *A nem hivatásos feliratozás a fordítástudomány szolgálatában: a vonatkozó mellékmondat használata az audiovizuális fordításban* című tanulmánya is. A cikk alapjául szolgáló empirikus kutatás célja az önkéntes TED-fordítók online feliratiban megjelenő nyelvhasználat összehasonlítása az eredeti magyar nyelvű TED előadások nyelvhasználatával, valamint a sztenderd nyelvhasználattal.

Harmadik kód a tolmácsolásban: létezik tolmácsolási szöveg? Intermodális korpuszok és esettanulmány című írásában **Szegh Henriett** a tolmácsolt szövegek sajátosságait vizsgálja. A szerző 2002-ben diplomázott az ELTE BTK francia nyelv- és irodalom szakán. 2011-ben végzett az ELTE Európai Uniós konferenciatolmács képzésén (magyar–

francia–angol). 2012-ben szakfordító vizsgát tett (társadalomtudomány, magyar–francia, francia–magyar). Jelenleg az ELTE BTK Nyelvtudományi Doktori Iskolájának Fordítástudományi Doktori Programjában folytatja tanulmányait. Szakterülete a tolmácsolás vizsgálata, azon belül a tolmácsolási nyelv jellemzőinek feltérképezése, cikkeit is ebben, illetve az anticipáció témájában publikálta (Szegh 2016, 2017). Jelen tanulmányában kutatásainak második szakaszáról számol be: intermodális elemzéssel megvizsgálja ugyanazon forrásnyelvi szövegek tolmácsolt és fordított változatát, először statisztikai vizsgálatoknak vetve alá a szövegeket, majd kvalitatív elemzést alkalmazva igyekszik feltárni, vajon mi állhat a tolmácsok döntéseinek hátterében.

Az utolsó, kutatási eredményeket közlő tanulmány – *Diskurzusjelölők* és kötőelemek *a magyar szinkrontolmácsolt és eredeti diskurzusban* – ugyancsak a tolmácsolás témakörével foglalkozik. Szerzője, **Götz Andrea** az ELTE BTK Fordítástudományi Doktori Programján szerzett PhD fokozatot, a Károli Gáspár Református Egyetemen tudományos segédmunkatársként tevékenykedik. A Pannónia Korpuszprojekt egyik alapítótagjaként többek között a tolmácsolási alkorpusz fejlesztésével foglalkozott (Götz 2017), kutatásainak középpontjában a mediált szövegek diskurzusjelölő-használata áll (Götz 2016). Jelen tanulmányában ezúttal arra a kérdésre keresi a választ, mennyiben motiválják a forrásnyelvi beszédek a magyar tolmácsolt szövegek pragmatikai és diskurzusjelölőit, és vajon gyakoribbak-e ezek a nyelvi elemek a tolmácsolt szövegekben, mint az autentikus szóbeli nyelvhasználatban.

A korpuszalapú fordításkutatásokkal foglalkozó elméleti értekezések és empirikus korpuszkutatások bemutatása után két recenzió következik, mindkét kiadvány a korpuszalapú vizsgálatok kihívásaival foglalkozik. Az első ismertető, **Benedek Enikő** doktorandusz hallgató írása, Mariachiara Russo, Claudio Bendazzoli és Bart Defrancq (2018) *Making Way in Corpus-based Interpreting Studies* címmel megjelent kiadványát mutatja be, amely a korpuszalapú tolmácsolástudomány elméleti keretrendszerének ismertetése után a szakképzett tolmácsok produkciójának korpuszalapú elemzéséből levont következtetéseiről tartalmaz tanulmányokat. A második recenzióban **Klenk Márk,** szintén a Fordítástudományi Doktori Program hallgatója, ismerteti Claudio Fantinuoli és Federico

Zanettin (2015) *New directions in corpus-based translation studies* című tanulmánykötetét. A szerzők gyakorlati példákon keresztül mutatják be a korpuszok fordítástudományi kutatásokban való alkalmazásának lehetőségeit.

A tanulmánykötet fentebb ismertetett cikkeinek mindegyike hozzájárul a magyarországi korpuszalapú fordítástudomány fejlődéséhez, ötvözve a jelentős méretű szövegállományra épülő kvantitatív vizsgálatok általánosíthatóságának előnyeit a kvalitatív elemzések mélyreható és felszíni szövegsajátosságok okait feltáró jellegével – hipotéziseket tesztelve, ugyanakkor újabb hipotéziseket vetve fel későbbi elemzések céljára. A tanulmányokban szereplő, empirikus alapokon nyugvó új eredmények és új elméleti fejtegetések így remélhetőleg alapot nyújtanak számos további vizsgálódás számára, gondoskodva az új kutatási paradigma hazai tudományos körökben való meggyökerezéséről.

Irodalom

Ábrányi H. 2017. *A fordítási környezetek hatása a fordítás minőségére.* Doktori értekezés. Budapest: ELTE. Elérhető: https://edit.elte.hu/xmlui/static/pdfjs/web/viewer.html?file=https://edit.elte.hu/xmlui/bitstream/handle/10831/32524/dissz_abranyi_henrietta_nyelvtud.pdf?sequence=1&isAllowed=y

Bozsik Gy. 2015. *A lexikai kohézió és az énekelhetőség vizsgálata operaszövegek fordításában*. Doktori értekezés. Budapest: ELTE. Elérhető: http://doktori.btk.elte.hu/lingv/bozsikgyongyver/diss.pdf

Chesterman, A. 1993. From ‘Is’ to ‘Ought’: Laws, Norms and Strategies in Translation Studies. *Target* Vol. 5. No. 1. 1–20. https://doi.org/10.1075/target.5.1.02che

Chesterman, A. 2007. What is a Unique Item? In: Gambier, Y., Shlesinger, M., Stolze, R. (eds) 2007. *Doubts and Directions in Translation Studies*. Amsterdam: Benjamins. 3–14. https://doi.org/10.1075/btl.72.04che

Dankó Sz. 2017. Alulreprezentált célnyelvi elemek a fordításban, avagy a „szokott” esete

Fordítástudomány 19. évf. 1. szám 75–84.

Götz A. 2016. A diskurzusjelölők fordításának kérdései. *Fordítástudomány* 18. évf. 1. szám. 5–18.

Götz A. 2017. Az első magyar intermodális korpusz bemutatása. Kutatási lehetőségek. *Argumentum* Vol. 13. 126–139.

Heltai P. 2007. Párhuzamos szaknyelvi korpusz munkálatai a Szent István Egyetemen. In: Feketéné Silye Magdolna (szerk.) *Porta Lingua*. Debrecen: DE ATC. 285–293.

Klaudy K. 1987. *Fordítás és aktuális tagolás.* Budapest: Akadémiai Kiadó.

Klaudy K. 2005. Párhuzamos korpuszok felhasználása a fordításkutatásban. In: Lanstyák I., Vančoné Kremmer I. (szerk.). *Nyelvészetről – változatosan. Segédkönyvek egyetemisták és a nyelvészet iránt érdeklődők számára*. Dunaszerdahely: Gramma Nyelvi Iroda. 153–185.

Kohn, J. 1997. Mit den beiden Flügeln atmen.... Computergestützte Analyse literarischer Texte. In: Henrici, G. és Kohn, J. (eds) *DaF-Unterricht im Spannungsfeld zwischen Forschung und Praxis, Sombathelyer Didaktik-Symposium 1995*. Szombathely: BDTF: 41–49.

Kovács M. 2015. *A frazeológiai univerzálék fordítási aspektusai és üzenetközvetítő szerepe európai uniós kontextusban.* Doktori értekezés. Budapest: ELTE. Elérhető: http://doktori.btk.elte.hu/lingv/kovacsmarietta/diss.pdf

Lengyel I. 2013. *A fordítási hiba fogalma funkcionális megközelítésben*. Doktori értekezés. Budapest: ELTE. Elérhető: http://doktori.btk.elte.hu/lingv/lengyelistvan/diss.pdf

Makkos A. 2015. Összehasonlítható kompetenciák anyanyelvi és fordított szövegekben. Az anyanyelvi fogalmazási kompetencia és a fordítási kompetencia összefüggései egyetemi hallgatók magyar nyelvű, összehasonlítható szövegei alapján. Doktori értekezés. Budapest: ELTE. Elérhető: http://doktori.btk.elte.hu/lingv/makkosani-

ko/diss.pdf

Mauranen, A., Kujamäki, P. (eds) 2004. *Translation Universals. Do they exist?* Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/btl.48

Mohácsi-Gorove A. 2014. *A minőség fogalma a fordítástudományban és a lektorálás mint minőségbiztosítási garancia.* Doktori értekezés. Budapest: ELTE. Elérhető: http://doktori.btk.elte.hu/lingv/mohacsigoroveanna/diss.pdf

Nagy A. L. 2017. A szógyakoriság gépi vizsgálata műszaki szövegeken. *Fordítástudomány* 19. évf. 2. szám. 40–57.

Nagy J. 2015. *A kommunikatív dinamizmus (relatív) egyensúlya a fordításban.* Doktori értekezés. Budapest: ELTE. Elérhető: https://edit.elte.hu/xmlui/bitstream/handle/10831/32440/dissz_Nagy_J%E1nos_nyelvtud.pdf;jsessionid=6A4C5AC4617F2FA25A75A5D97F1D15C5?sequence=1

Polcz K. 2012. *Konvencionálisan indirekt beszédaktusok az angol–magyar filmfordításban.* Doktori értekezés. Budapest: ELTE. Elérhető: http://doktori.btk.elte.hu/lingv/polczkaroly/diss.pdf

Robin E. 2014. *Fordítási univerzálék a lektorált szövegekben*. Doktori értekezés. Budapest: ELTE. Elérhető: http://doktori.btk.elte.hu/lingv/robinedina/diss.pdf

Robin E., Szegh H. 2018. A Pannónia Korpusz audiovizuális alkorpusza. In: Dróth J. (szerk.) *Gépiesség és kreativitás a fordítási piacon és az oktatás különböző szintjein.* Budapest: KGRE, L'Harmattan Kiadó. 93–110.

Sato N. 2014. *A vállalati és üzleti tolmács kettős lojalitása a magyar–japán és a japán-magyar interperszonális kommunikációban.* Doktori értekezés. Budapest: ELTE. Elérhető: http://doktori.btk.elte.hu/lingv/satonoriko/diss.pdf

Seidl-Péch O. 2011a. *Fordított szövegek számítógépes összevetése.* Doktori értekezés. Budapest: ELTE. Elérhető: http://doktori.btk.elte.hu/lingv/seidlpecholivia/diss.pdf

Seidl-Péch O. 2011b. Fordított szövegek számítógépes összevetése. In: Bocz Zs., Sárvári

J. (szerk.) (2013) Válogatott cikkek, tanulmányok, 2010–2013. Budapest: BME GTK Idegennyelvi Központ. 369–386.

Somodi J. 2014. *Megszólítások pragmatikája japán-magyar összevetésben. A japán apellatív megszólítások fordításának vizsgálata magyar filmszövegekben.* Doktori értekezés. Budapest: ELTE. Elérhető: http://doktori.btk.elte.hu/lingv/somodijulia/diss.pdf

Szegh H. 2017. Harmadik kód a tolmácsolásban: létezik tolmácsolási szöveg *Fordítástudomány* 19. évf. 1. szám. 60–74.

Tirkkonen-Condit, S. 2004. Unique items – over- or under-represented in translated language? In: Mauranen, A., Kujamaki, P. (eds) *Translation Universals: Do they exist?* Amsterdam: Benjamins. 177–186. https://doi.org/10.1075/btl.48.14tir

Visszatekintés egy korai korpuszelemzésre

Klaudy Kinga

klaudy.kinga@btk.elte.hu

Eötvös Loránd Tudományegyetem Bölcsészettudományi Kar

Fordító- és Tolmácsképző Tanszék

Kivonat: A szerző 1980-ban végzett elemzést összehasonlítható (*comparable*) korpuszokon. Az elemzést manuálisan végezte, de az eredmények ma is vállalhatók. Oroszból fordított magyar szakmai és tudományos szövegeket hasonlított össze hasonló időben keletkezett és hasonló témájú autentikus magyar szövegekkel. Vizsgált bizonyos grammatikai mutatókat (a mondatok és a mondategységek hosszúsága, szerkesztettsége, mélysége, bővítettsége), valamint bizonyos aktuális tagolási típusokat (visszautaló téma, informatív téma, retorikus téma, emelkedő réma, ereszkedő réma, egyenletes réma stb.). A vizsgálat kimutatta, hogy a fenti grammatikai mutatók és aktuális tagolási típusok gyakorisága tekintetében eltérések vannak a fordított magyar szövegek és az eredetileg is magyarul fogalmazott szövegek között. A grammatikai tagolásban és az aktuális tagolásban kimutatott eltérő disztribúciós jegyek alapján a fordított szövegeket kvázi-helyes magyar szövegnek nevezte.

Kulcsszavak: fordított és eredeti magyar szöveg, grammatikai tagolás, aktuális tagolás, disztribúció, kvázi-helyesség

1. Bevezetés

Egy olyan kutatásra szeretnék visszatekinteni, amelyet 1980-ban nyújtottam be kandidátusi értekezésként, 1981-ben védtem meg, és 1987-ben jelent meg az Akadémiai Kiadónál *Fordítás és aktuális tagolás* címmel. Papp Ferencnek, a számítógépes nyelvészet magyarországi megteremtőjének tanítványaként természetes volt, hogy a fordításkutatásban is

Klaudy Kinga: Visszatekintés egy korai korpuszelemzésre. In: Robin E., Seidl-Péch O. (szerk.) 2020. *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordításkutatás Magyarországon.* Segédkönyvek a nyelvi közvetítésről I. Budapest: ELTE BTK Fordítástudományi Doktori Program, MANYE Fordítástudományi Szakosztály. DOI: https://doi.org/10.36252/Nyelvikozvsegedkonyv1.2

vonzottak az egzakt módszerek, és bár azok a szövegcsoportok, amelyeket vizsgáltam, nem voltak gépileg olvashatók és automatikusan lekérdezhetők, mégis korpusznak neveztem őket, mint ahogyan minden nyelvész annak nevezte akkoriban a nyelvi adatok forrásaként használt szövegeit. Részben azért érzem most (40 évvel később) érdemesnek feleleveníteni ezt a kutatást, mert több szempontból előrevetítette a Mona Baker által 1993-ban megindított korpuszalapú fordításkutatás elveit, részben pedig azért, mert néhány megállapítását a későbbi nagyméretű korpuszokon végzett elemzések is igazolták, vagy még igazolhatják.

Megjegyzem, hogy a visszatekintés műfajából fakadóan nem fogom sem a használt terminusokat, sem a felhasznált szakirodalmát részletezni, mindez megtalálható és olvasható az eredeti műben. Ez különösen az aktuális tagolással foglalkozó fejezetben lesz szembetűnő, ahol a téma, réma, tematikus szakasz, rematikus szakasz, rematikus csúcs, kommunikatív dinamizmus fogalmát úgy fogom használni, mint amelyek a nyelvtudományban közismertek és nem szorulnak magyarázatra. Ugyanebből az okból minimális mennyiségű lesz a nyelvi példák száma is, továbbá csak az oroszból fordított magyar (OM) és az autentikus magyar (M) szövegcsoportból hozok példákat.

2. A kvázi-helyesség fogalma és a korpusz összeállításának elvei

A kutatási témát oktatói munkám tapasztalatai hozták. Az 1973-as tanévtől az ELTE BTK Fordító- és Tolmácsképző Csoportjában folyó posztgraduális fordítóképzésben oktattam orosz–magyar fordítástechnikát, és arra próbáltam nyelvészeti magyarázatot találni, miért olyan nehezen érthetőek az oroszul amúgy jól tudó, orosz szakot végzett hallgatók magyarra fordított szövegei, amikor mondataik egyenként grammatikailag helyes magyar megfogalmazásnak tekinthetők. Véletlenül bukkantam rá Papp Ferenc 1972-ben megjelent cikkére, aki ugyanezt a jelenséget a magyar anyanyelvű orosz szakos egyetemi hallgatók orosz nyelvű fogalmazásaiban mutatta ki, és kvázi-helyességnek nevezte. Idézem az ő meghatározását orosz nyelvű cikkének 2005-ben megjelent magyar fordításából:

a) a szöveg minden mondata az adott nyelv grammatikai szabályai szerint épül fel (az anyanyelvi beszélők minden mondatot nyelvtanilag helyes mondatként fogadnak el);

b) a szövegben közvetlenül egymás mellett álló minden egyes mondatpár a *topic–comment* szabályai szerint épül fel, a megfelelő nyelvi eszközök kötik őket össze (az anyanyelvi beszélők minden szomszédos mondatpárt helyesen szerkesztett és egymáshoz helyesen kapcsolt mondatként fogadnak el);

c) az egész szöveget az anyanyelvi beszélők elutasítják, mint olyan szöveget, amely nem felel meg az adott nyelven helyesen szerkesztett szövegekről alkotott intuitív elképzeléseiknek. (Papp 2005: 122, Répási Györgyné fordítása)

A kvázi-helyesség első okaként Papp a következőt említette: „Az idegen nyelvet magas szinten elsajátító személy hibát követhet el bizonyos szövegjelenségek használatának statisztikai eloszlásában" (Papp 2005:123). Példának azt a jelenséget hozta fel, hogy az oroszul fogalmazó magyarok általában kevesebb szenvedő szerkezetet, igeneves szerkezetet és több mellékmondatot használnak, mint ahogyan az az autentikus orosz szövegekben szokás.

Nekem nagyon fontosak voltak ezek a gondolatok, mert megerősítettek abban, hogy a magyarra fordított szövegek idegenszerűségét is csak szövegszinten lehet megragadni. Méghozzá úgy, és ez már a saját leleményem volt, hogy autentikus magyar szövegekkel vetjük őket egybe. Ez volt az első „korát megelőző" ötletem, hiszen nem tudok róla, hogy akkoriban bárki végzett volna egybevetést a fordítás eredményképpen keletkezett célnyelvi szövegek és az eredetileg is célnyelven írt szövegek között. Mona Baker 1995-ös programadó cikkében írja le a fordításkutatásban használható szövegkorpuszok fajtáit: (1) párhuzamos korpusz, (2) multilingvális korpusz, (3) összehasonlítható korpusz. A nyolcvanas években az utóbbi terminust ugyan még nem ismerhettem, de a Baker definíciójának megfelelő összehasonlítható (*comparable*) korpuszt építettem, amikor az

oroszból fordított magyar szakmai és tudományos szövegcsoport mellett azonos műfajú, azonos időben keletkezett autentikus magyar szakmai és tudományos szövegcsoportot hoztam létre.

A másik ilyen „előrevetített" gondolatom a fordított szövegek értékítélettől mentes leírásának gondolata volt. Ezt akkor így fogalmaztam meg:

> (...) nekünk nem az a célunk, hogy a „germanizmusok" ellen harcolók nyomdokain haladva most „russzicizmusokra" vadásszunk a fordított szövegekben, netán „ki akarjuk irtani őket", hanem az, hogy a fordított szövegeket nyelvi ténynek fogva fel megpróbáljuk különböző szempontokból megvizsgálni és leírni őket. (Klaudy 1987:7)

A két magyar szövegcsoport szövegszintű különbségeinek kimutatásán kívül az orosz szöveg interferenciájára is kíváncsi voltam, ezért korpuszom egy orosz nyelvű szövegcsoportot is tartalmazott. Mivel manuálisan elemeztem, úgy gondoltam, hogy 200-200 mondatot elemzek mindhárom szövegcsoportból. Hogy minél változatosabb legyen a merítés, háromszor 20 könyvből választottam ki egyenként 10 mondatot a könyvek különböző részeiből véletlenszerű mintaválasztással (ahol táblázat vagy ábra volt, ott módosítottam). Volt tehát három szövegcsoportom: autentikus orosz szakmai és tudományos szövegek (O), oroszból fordított magyar szakmai és tudományos szövegek (OM), és autentikus magyar szakmai és tudományos szövegek (M).

Két megjegyzés a fordított szövegekről: (1) az OM szövegcsoportot függetlenül állítottam össze, tehát nem az O szövegcsoport fordításait tartalmazta; (2) az OM szövegcsoportban nem használtam hallgatói fordításokat, minden fordítás nyomtatásban megjelent szöveg volt (adataikat ld. Klaudy 1987: 122–124). A korpusz szövegeinek kiválasztása is arra utal, hogy nem hibás szövegekre vadásztam, hanem az oroszból magyarra fordított szövegek általános jellegzetességeit akartam megragadni.

Természetesen lehet vitatkozni arról, hogy a kvázi-helyesség terminus mennyire mentes az értékítélettől, de akkor úgy döntöttem, hogy a dolgozat céljának megfelel. Későbbi munkáimban már nem használtam, mert akkor már inkább a fordítás folyamata, azaz az átváltási műveletek rendszerének leírása foglalkoztatott, és nem a fordított szövegek jellegzetességeinek leírása.

A célnyelvi elemek eltérő disztribúcióját egyébként Mona Baker 1993-as cikkében felveszi a feltételezett és a jövőben korpusznyelvészeti eszközökkel vizsgálandó fordítási univerzálék közé, amelyek nyelvpártól és fordítási iránytól függetlenül jellemzik a fordított szövegeket (Baker 1993: 245). Baker az univerzálék felsorolásakor abból indul ki, hogy a fordítások szövege nem jobb, nem rosszabb, csak más, és ennek a másságnak a megnyilvánulásait tekinti univerzálénak (Baker 1993: 234). A másság megállapításával azonban nem zárhatjuk ki azt a tényt, hogy az olvasók számára az eltérő disztribúció megértési problémákhoz vezethet, és ezzel veszélyeztetheti a kommunikatív ekvivalenciát.

3. Kvázi-helyeség az OM mondatok grammatikai tagolásában

Az oroszból magyarra fordított szövegek kvázi-helyességét két szempontból vizsgáltam meg: a három szövegcsoport mondatainak grammatikai tagolásában és aktuális tagolásában. A grammatikai tagolás tekintetében az volt a hipotézisem, hogy a fordított magyar szövegek és az autentikus magyar szövegek közötti szintaktikai különbségeket számszerűen is ki lehet mutatni, és bizonyos szintaktikai mutatók tekintetében az oroszból fordított (OM) magyar szövegek valahol félúton lesznek az orosz szövegek (O) és az autentikus magyar (M) szövegek között.

A kutatás módszere a három szövegcsoport 600 mondategészének mondategységekre való bontása volt. A *mondategész* és a *mondategység* terminust Deme Lászlótól (1971) kölcsönöztem, mert el akartam kerülni a *főmondat* és a *mellékmondat* terminus használatát. A *mondategész* terminus magától értetődő, a *mondategység* pedig egy predikatív egységet jelent. Ezután az így kapott 1342 mondategység szintaktikai képletének leírása következett oly módon, hogy az alany (A), az állítmány (Áll), a tárgy (T), a határo-

zó (H) és a kötőszó (k) mondatszintű elemnek számított, a jelzői bővítmények pedig mondatszint alatti elemként (mélységüktől függetlenül) alsó indexbe kerültek, az alapszótól jobbra vagy balra. A mondatok képletének leírását és az adatok felvételének módját egy OM szövegcsoportból származó mondat képletének leírásával szemléltetem:

(1) Mindezek a különböző időkbe tartozó és a tudományos folyamat szerves részét alkotó elemek bonyolult, dialektikus kölcsönhatásban vannak egymással, és ebből fakad a tudomány, mint bonyolult rendszer történetiségének törvényszerű jellege. (OM 1.2)

A mondat képlete: $_{10}$A–$_{2}$H–Áll–H/k–H–Áll–$_{6}$A

1. táblázat

A három szövegcsoport összevetésének adatai

Az összevetés egységei	**O**	**OM**	**M**
mondatok száma	200	200	200
mondategységek száma	354	453	535
összes szövegszó	3 724	3 755	3 954
mondatszintű elemek száma	1 462	1 833	2 277
mondatszint alatti elemek száma	2 262	1 922	1 677
balra álló mondatszint alatti elemek száma	531	1 830	1 644
jobbra álló mondatszint alatti elemek száma	1 731	92	33

Látjuk, hogy az OM szövegcsoport első szövegének második mondatában a szövegszavak száma 26, az önálló mondategységek száma 2, a mondatszintű elemek száma 8, a balra álló mondatszint alatti elemek száma 18. Az ilyen módon manuálás számlálás alapján kapott összesített adatokat az *1. táblázat* tartalmazza.

Az arányokat a 2., 3., 4. és 5. táblázat fogja szemléltetni. A *2. táblázatból* láthatjuk, hogy az egy mondatra jutó szószám tekintetében nincs nagy különbség a három szövegcsoport között, de a mondategészek tagoltsága eltér: a legkevésbé tagoltak, azaz a legkevesebb mondategységet az orosz mondatok tartalmazzák (1,77 mondategység jut egy mondategészre), a legtagoltabbak az autentikus magyar szövegek mondatai (2,67), a fordított mondatok tagoltsága pedig a kettő között van (2,26).

2. táblázat

A mondategészek átlagos hosszúsága és tagoltsága

Mondategészek/mondategységek	**O**	**OM**	**M**
szó/mondategész	18,62	18,77	19,77
mondategység/ mondategész	1,77	2,26	2,67

A mondategészek tagoltsága, azaz hogy a mondategészek hány önálló mondategységet tartalmaznak, tipikusan olyan sajátosság, amely szabad szemmel nem látható, nem vehető észre. A *2. táblázat* mutatja, hogy a fordított szövegekben nő a mondategységek száma, tehát a fordítók próbálnak igésíteni, új predikatív egységeket alkotni, de a tagoltság nem éri el az autentikus magyar szövegekét.

Ezzel függ össze az, amit a *3. táblázat* mutat, hogy az összes szövegszónak csak 39,25% van mondatszinten az orosz szövegekben, míg az autentikus magyar szövegekben 57,58%-a. A fordított magyar szövegek újra a kettő között helyezkednek el (51,18%).

3. táblázat

Az összes szövegszó megoszlása

Összes szövegszó	**O**	**OM**	**M**
mondatszinten állók	39,25%	48,81%	57,58%
mondatszint alatt állók	60,74%	51,18%	42,41%

Az olvasó számára a mondatszintű elemek azonnal felfoghatók, míg a mondatszint alatt elhelyezkedő, jelzői szerkezetekbe zsúfolt elemek nehezebben értelmezhetők. A fordítók

igyekeznek ugyan felemelni a mondatszint alatti bővítményeket a mondat szintjére, de nem érik el az autentikus magyar szövegek szintjét.

A következő, *4. táblázat* a mondatszint alatti bővítmények elhelyezkedését mutatja. Látjuk, hogy az orosz mondatszint alatti bővítmények az alapszótól balra is, jobbra is állhatnak. Az alapszótól balra helyezkedik el 23,47%, jobbra pedig 76,52%. Az autentikus magyar szövegekben viszont az alapszótól balra áll a bővítmények 98,3%-a, a jobbra bővítés lehetősége megvan ugyan (*pl. ugrás a sötétbe, harc a végsőkig* stb.), de a magyar nyelvhasználat csak korlátozottan él vele. Az alapszótól jobbra álló határozói jelzőket leginkább csak címekben vagy mondatok végén találunk.

4. táblázat

A mondatszint alatti bővítmények elhelyezkedése

A mondatszint alatti bővítmények elhelyezkedése	**O**	**OM**	**M**
az alapszótól balra áll	23,47%	95,21%	98,03%
az alapszótól jobbra áll	76,52%	4,78%	1,96%

Köztudott volt eddig is, hogy az orosz főnévi csoportra a jobbra bővítés, a magyar főnévi csoportra a balra bővítés jellemző, elgondolkodtató azonban, hogy míg az oroszban az uralkodó jobbra bővítés csak 76,52%-os, a balra bővítésre is jut 23,47%. A mennyiségi és minőségi jelzők az alapszó előtt állnak, és csak a birtokos jelző és a határozói jelző kerül az alapszó mögé. Tehát az oroszban két lehetőség van a mondatszint alatti bővítmények elhelyezésére. A magyarban viszont az uralkodó balra bővítés 98,03%-os, tehát differenciálatlanul minden mondatszint alatti bővítmény az alapszó elé kerül.

A fordításban az orosz főnévi csoportok jobbra álló határozói jelzői is a magyar alapszó elé, azaz balra kerülnek. Ahhoz, hogy a főnév elé kerülhessek, a fordítóknak jelzősíteni kell őket, amit általában üres, deszemantizálódott melléknévi igenevekkel oldanak meg (*alapuló, célzó, fakadó, járó, történő, gyakorló, folyó*). Ez volt egyébként az egyetlen adat, amelyet gépileg kaptam. Hell György készített nekem szógyakorisági statisztikát a

BME Számítóközpontjában egy 10 ezer szavas OM és egy 10 ezer szavas M szövegről, és a fordított szövegben 16 találatot kaptunk a *folyó* szóra, míg az autentikus magyarban egyet sem (Klaudy 1987: 21). A téma jellegéből adódóan vízi útról nem lehetett szó.

A bővítési lehetőségek közötti különbségek tükröződnek az *5. táblázatban*, ahol a mondategységekkel kapcsolatos arányokat láthatjuk. A három szövegcsoport mondategységei majdnem azonos mennyiségű mondatszintű elemet tartalmaznak, de a mondatszint alatti elemek száma jelentősen eltér. Az orosz szövegcsoportban a mondategységek hosszabbak, és több mondatszint alatti elemet tartalmaznak.

5. táblázat

A mondategységek átlagos hossza és bővítettsége

Az összevetésben szereplő mutatók	**O**	**OM**	**M**
szó/mondategység	10,50	8,28	7,39
mondatszintű elem/ mondategység	4,12	4,04	4,25
mondatszint alatti elem/ mondategység	6,38	4,24	3,13

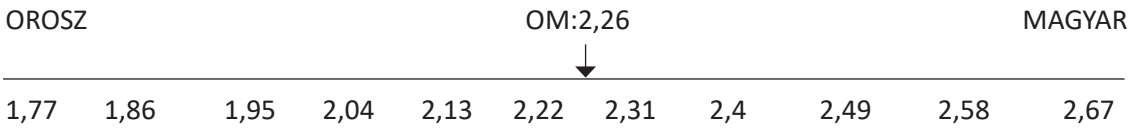
Emlékezzünk vissza arra, hogy a szövegszavak száma szerint a mondategészek átlagos hosszúságában nem volt nagy különbség a három szövegcsoport között. A különbség a mondategészek tagoltságában mutatkozik meg. A magyar nyelvben az egy mondatra jutó információmennyiséget több mondategységben helyezzük el, mint az orosz nyelvben, és ezzel természetesen együtt jár, hogy az egy mondategységre jutó mondatszint alatti bővítmények száma is kevesebb lesz. A mondategységek bővítettségi mutatója a magyar szövegcsoportban (3,13), ez kevesebb mint a fele az orosz mondategységek bővítettségi mutatójának (6,38). A fordítások adataiból látszik, hogy a fordítók igyekeztek növelni a mondategységek számát, és csökkenteni a mondatszint alatti bővítettséget (4,24), úgy is mondhatnánk, a mondat mélységét (Yngve 1973), de nem érték el az autentikus magyar szövegek szintjét.

A vizsgált 1342 mondategység képletének leírása és a számszerű adatok manuális összesítése alapján megállapítható, hogy a magyar fordítások bizonyos szintaktikai mutatók tekintetében az orosz átlag és a magyar átlag között helyezkednek el. A monda-

tegészek kevésbé tagoltak (*1. skála*), a mondategységek hosszabbak (*2. skála*), jobban bővítettek (*3. skála*), és egy mondatszintű elemre több mondatszint alatti bővítmény jut (*4. skála*), mint az autentikus magyar szövegekben.

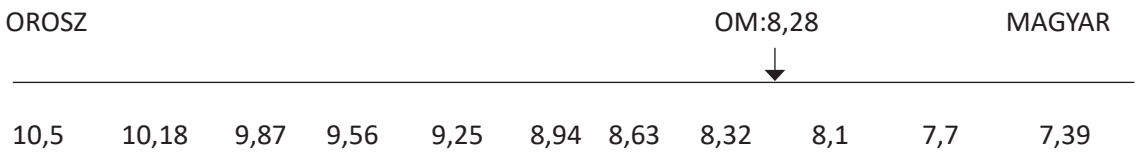
1. skála

A mondategészek tagoltsága a fordításokban (mondategység/mondategész)



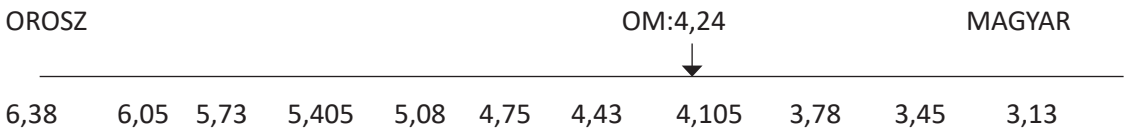
2. skála

A mondategységek hossza a fordításban (szövegszó/mondategység)



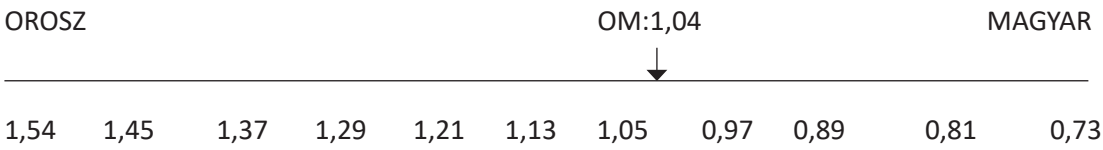
3. skála

A mondategységek mondatszint alatti bővítettsége
(mondatszint alatti elem/mondategység)



4. skála

A mondatszintű elemek mondatszint alatti bővítettsége
(mondatszint alatti elem/mondatszintű elem)



Számításaink alátámasztják azt a tényt, amit fordítók, lektorok és szerkesztők intuitíve nyilván mindig is érzékeltek, hogy az oroszból fordított magyar szövegekben a mondategészek kevésbé tagoltak, azaz kevesebb mondategységből állnak, a mondategységek több mondatszint alatti bővítményt tartalmaznak, és hosszabbak bennük az alapszó előtt álló mondaszint alatti bővítményláncok. Ezek az eltérések csak szövegszinten érzékelhetők, hiszen a mondatok egyenként jól formált magyar mondatoknak tekinthetők.

4. Kvázi-helyesség az OM mondatok aktuális tagolásában

Szintén gyakorló fordításoktatói és lektorálási tapasztalataim mutatták, hogy a fordított szövegek jellegzetességeinek leírásában a téma–réma (a továbbiakban TR) viszonyokat is érdemes vizsgálni, hiszen a hallgatói fordításokban és a kiadói lektorálási megbízásaimban nagyon sokszor kellett szórendet javítanom. Az oroszból fordított magyar szövegekben későn világosodott meg a mondat szerkezete, eltolódtak a hangsúlyok, egyszóval fárasztó volt olvasni, megérteni őket. Mivel a magyar mondatok szórendje viszonylag szabad, a szófajok és mondatrészek nem pozícionálisan, hanem morfológiailag vannak meghatározva, ezeket a szórendi javításokat valóban csak szövegszinten lehetett indokolni. Ilyen indoklásokat adtam a javításaimban, hogy „Rövidítsük meg a mondat elejét, így szorosabbra fűzzük a kapcsolatát az előző mondattal…”, „Hozzuk előre a támpontot, mert ha az igei állítmány a mondat végére csúszik, későn világosodik meg a mondat szerkezete…” stb. Bár tudományos szakszövegeket javítottam, ezeknek a tudományos szakszö-

vegeknek is volt dallamuk, és gyakran éreztem, hogy a hangsúlyok a fordításban nem jó helyen vannak. Papp Ferenc azt írta, hogy a kvázi-helyes szövegek mondatai topic–comment szempontból is rendben vannak (lásd a cikk elején említett meghatározásának b) pontját), ebben én nem voltam olyan biztos.

A TR viszonyok feltárásához a három szövegcsoportot megpróbáltam kommunikatív egységekre (KE) bontani, azaz mindegyik mondategységében elkülönítettem a tematikus szakaszt a rematikus szakasztól, és a rematikus szakaszon belül elkülönítettem a rematikus csúcsot, a rematikus szakasz leghangsúlyosabb részét. Ezt úgy végeztem, hogy a mondategységek már meglévő szintaktikai képletében szögletes zárójellel jelöltem ki a kommunikatív tagolást, azaz a tematikus és rematikus szakaszt, és gömbölyű zárójellel a rematikus csúcsot. Azért használtam saját terminusokat, mert az általam használt terminusok nem egészen estek egybe a szakirodalomban akkor már széles körben használt topik, komment és fókusz terminussal (É. Kiss 1978).

A képlet leírását az M szövegcsoport egyik mondatával szemléltetem, először a mondatot mutatom be kommunikatív egységekre (KE) bontva, majd annak képletét. Az alsó indexekben a mondatszint alatt álló jelzői bővítményeket tömörítettem. A mondategész egyetlen mondategységből áll, a mondategységet bevezető tárgyi bővítmény a mondat tematikus szakasza, utána kezdődik a rematikus szakasz, ezen belül a határozó a rematikus csúcs, és az állítmány a rematikus szakasz leszálló ágában van, mint ezt majd a rématípusoknál látni fogjuk.

(2) $_{T}$[A 48-as olasz szabadságharc bukásának végső okát] # $_{R}$ [($_{RCS}$ a polgárság viszonylagos fejletlenségében) kell keresnünk.]

Képlete: $_{T}$ [$_{5}$T] #$_{R}$ [($_{RCS\ 2}$H)–$_{1}$Áll]. (M 8.1.)

A szakmai és tudományos szövegekben a mondatok nagy része természetesen objektív szórendű, az ismerttől (téma) halad az ismeretlen (réma) felé, tehát a mondategészek nagy része téma–réma (TR) tagolású. Ha meseszövegeket vizsgáltunk volna, ott

találtunk volna sok szubjektív szórendű, azaz a rémától a téma felé haladó mondatot (pl. *Ment mendegélt a farkas*…). Kommunikatív egységekre lebontva a korpusz kis számban tartalmazott réma–téma (RT) tagolású mondategységeket, valamint csak rémát tartalmazó mondategységeket (komplex réma=KR) és csak témát tartalmazó mondategységeket (komplex téma=KT). A kommunikatív egységek száma minimálisan tért csak el a mondategységek számától: 600 mondat 1342 mondategységet és 1332 kommunikatív egységet tartalmazott, ezért a dolgozat további részeiben is a mondategység terminussal dolgoztam.

6. táblázat

A négy fő aktuális tagolási típus abszolút és százalékos megoszlása a három szövegcsoportban

	O		**OM**		**M**	
	KE száma	**%**	**KE száma**	**%**	**KE száma**	**%**
TR	270	77,42	320	70,17	368	71,04
RT	8	3,14	11	2,41	20	3,86
KR	63	15,42	115	25,21	122	23,55
KT	14	4,00	10	2,19	8	1,84
Összesen	358	100,00	456	100,00	518	100,00

Amint a *6. táblázatból* látható, a három szövegcsoport nem különbözik jelentősen egymástól a fő aktuális tagolási típusok számszerű megoszlása tekintetében. A mondategységek határai megegyeznek a kommunikatív egységek határaival, és nagy részük objektív szórendű, azaz téma–réma tagolású. A további vizsgálat megmutatta, hogy az eltérések a tematikus és a rematikus szakaszok jellegzetes típusainak megoszlásában keresendők.

4.1. Kvázi-helyesség a fordított szövegek tematikus szakaszában

Röviden ki kell térnünk itt a szintaktikai mondattagolás és az aktuális (más terminussal értelmi vagy kommunikatív) mondattagolás viszonyára is, hiszen köztudott, hogy a két tagolás nem esik egybe: az alany nem mindig esik egybe a témával, az állítmány nem mindig esik egybe a rémával. Nyelvenként és műfajonként eltérő, hogy mi töltheti be a

mondat témájának szerepét, mi állhat témaként. Az angolban például a szegényes morfológiai jellegzetesség miatt a mondat témája legtöbbször megegyezik a mondat alanyával, hiszen az alany pozícionálisan van meghatározva a mondat elején. A magyarban viszont gyakran állnak hely és időhatározók a mondat elején, és még az igei állítmánnyal való mondatkezdés sincs kizárva. Ilyen szempontból az orosz is gazdag morfológiájú nyelv, és bármi állhatna a mondat elején, számításaink viszont azt mutatták, hogy mégis az alanyi téma a leggyakoribb.

7. táblázat

A tematikus alanyok gyakorisága, hosszúsága és mondatszint alatti bővítettsége az O és az M szövegcsoportban

Az összevetésben szereplő mutatók	**O**	**M**
alanyi TSZ/összes TSZ	69,3 %	52,2 %
alanyi TSZ szószáma/alanyok száma	3,81 szó	2,53 szó
alanyi TSZ mondatszint alatti bővítettsége / alanyok száma	2,25 szó	1,41 szó

Sokéves kontrollszerkesztési, fordatásoktatási gyakorlatom alapján feltételeztem, hogy a fordított szövegekben a kvázi-helyeség egyik megnyilvánulása a tematikus alanyok nagyobb gyakorisága, hosszúsága és bővítettsége lesz. Ilyen kezdetű mondatokra gondoltam az OM szövegcsoportból:

(3) A szerzők általánosításai és javaslatai a művészi-tervezői megformálásra és a különféle termékfajták legjobb felhasználásra vonatkozólag… (OM 19.10)

(4) Mindezek a különböző időkbe tartozó és a folyamat szerves részét alkotó elemek… (OM 1.5.)

A számszerű adatok ezt a feltételezésemet nem igazolták. A *6. táblázatban* látható különbségek az orosz és a magyar szövegcsoport alanyi témáinak gyakorisági, hosszúsági és bővítettségi mutatói között nem tükröződtek a magyar fordításban, ahol a következő

mutatókat kaptam: az OM szövegcsoportban a tematikus alanyok gyakorisága: 53,7%, hosszúság 2,66 szó, mondatszint alatti bővítettség 1,36 szó. Ezek a mutatók alig térnek el az autentikus magyar szövegek mutatóitól (ld. *7. táblázat*). Nem azt mondom, hogy az orosz tematikus szakaszok szintaktikai megformálásnak nincs hatása a fordított magyar szövegekre, de ez a hatás másképp jelentkezik: olyan mondatkezdéstípusok jelennek meg a fordított magyar szövegekben, amelyek az autentikus magyar szövegekben szokatlanok, ritkán vagy sohasem fordulnak elő.

Ezeknek a mondatkezdéstípusoknak a feltárásához háromféle tématípust különítettem el:

- Összefoglaló-visszautaló téma: az előző mondategység, mondategész vagy bekezdés tartalmát foglalja össze, vagy arra utal vissza.
- Informatív téma: új információt közöl, de ez az információ a szerző szándéka szerint csak bevezetés valami még fontosabbhoz, ami a mondat rematikus szakaszában található. Az informatív mondatkezdés azért téma, mert rajta kívül a mondatban van még fontosabb információt tartalmazó réma is.
- Retorikus téma: Olyan, rendszerint halmozott és bővített névszói szerkezettel megformált tematikus szakasz, amely tartalmát tekintve lehet akár visszautaló, akár informatív, az előző két típustól eltérően azonban olyan rematikus szakasz követi, amely nem hordoz fontosabb információt, csak leszögezi a tematikus szakaszban mondottak fontosságát (pl. mikor a szónoki beszédben az előadó a célok hosszas felsorolását úgy fejezi be, hogy …*ez a mi feladatunk*).

4.1.1. Az összefoglaló-visszautaló téma

Az összefoglaló-visszautaló témában (a továbbiakban visszautaló téma) a visszautalás lehet explicit (pl. mutatónévmással), és lehet implicit, amikor a szerző a tematikus szakaszban mintegy összefoglalva megismétli az előző mondatban vagy mondatokban közölt információkat. Az orosz szerzők gyakran használják a szövegkoherencia megteremtésé-

nek azt a módszerét, hogy névszói szerkezetté alakítva építik be az előző mondat vagy mondatok rematikus szakaszát a következő mondat tematikus szakaszába. Ez a lehetőség a magyarban is megvan, de az autentikus magyar szövegek szerzői a magyar főnevek korlátozottabb (szinte csak balra irányuló) bővítési lehetőségei miatt ritkábban élnek vele. Ezt mutatják korpuszom számszerű adatai is. Magyar szövegcsoportom első mondategységeinek visszautaló tematikus szakaszát vizsgálva a visszautalás módja az esetek 70,95%-ban explicit volt, és csak 29,5%-ban implicit. Az orosz szövegcsoport 56,25%-ban találtam explicit és 43,75%-ban implicit visszautalást. Az impliciten visszautaló tematikus szakaszok tehát az orosz szövegcsoportban gyakoribbak voltak, mint a magyar szövegcsoportban.

A fordító ilyenkor két dolgot tehet. Ha meghagyja a szintaktikai formát, és a magyar főnevek bővítési lehetőségeit a végsőkig kihasználva főnévi szerkezettel fordít, akkor az olvasó számára túlságosan hosszú tematikus szakasszal kezdődő és első olvasásra nehezen érthető mondatot kap (5. példa). A másik lehetőség a szintemelő fordítás (6. példa). A mondat szintje alatt álló elemek a célnyelvben a mondat szintjére kerülnek, és ezzel új lehetőséget teremt a fordító a bővítmények elhelyezésére. De ezzel megváltoztatja a mondat kommunikatív szerkezetét is. Ennek elkerülésére szokták a fordítók „visszatematizálni" a felemelt szerkezeteket, és megjelennek, az *Az, hogy…, Az a tény, hogy…* kezdetű mondatok. Léteznek a magyarban is, de a fordított szövegekben nagyobb gyakorisággal fordulnak elő.

(5) Szintemelés nélküli fordítás: A közvetlenül a vállalati számláról finanszírozott nagyjavítások ellenőrzésének a felügyeleti szervek által történő elmulasztása…

(6) Fordítás szintemeléssel: Az a tény, hogy a felügyeleti szervek nem ellenőrzik, mire költi a vállalat azt az összeget, melyet az elszámolási számláról a nagyjavításokra közvetlenül felvett…

Ebben az esetben 500 mondatos mintákat vettem 3 OM és 3 M szövegből, és az OM szövegeinkben 2-5-7 *Az, hogy* … kezdetű mondatot találtam, míg a három autentikus magyar mintában egyet sem. A szintemelő fordítás elé kitett „tematikus előjel" tipikus példája a forrásnyelvi szöveg közvetett interferenciájának.

4.1.2. Az informatív téma

Az informatív téma olyan mondatkezdő tematikus szakasz, amely nem visszautaló funkciót tölt be, hanem új információt hordoz. Informatív súlya ellenére sem tekinthető rémának, mert a szerző ugyanabban a mondatban valami még fontosabbat is akar közölni.

Orosz szövegcsoportunk TR tagolású mondatainak 27,3%-a kezdődött informatív témával, átlagos hosszúságuk 6,58 szó, mondatszint alatti bővítettségük 2,39 szó volt. Magyar szövegcsoportunkban az informatív témával kezdődő mondatok aránya hasonló volt (23,3%), de átlagos hosszúságuk csak 3,5 szó, mondatszint alatti bővítettségük pedig csak 0,94 szó. Más a szintaktikai megformálás is. Az informatív téma a magyar mondatok elején inkább idő vagy helyhatározó (45,7%), és csak kisebb arányban alany 31,4%, míg az oroszban az informatív téma 52,2%-ban alanyi.

A hely- és időhatározós mondatkezdések az informatív téma természetes fajtái mindkét nyelvben, fordításuk semmilyen problémát nem okoz. Az orosz tudományos szövegekben azonban gyakran előfordul, hogy a szerzők a mondatkezdő alanyt is új információval terhelik meg, és ez a fordítókat nehéz helyzetbe hozza, különösen akkor, ha halmozott és bővített tematikus alanyokat kell lefordítania. Ezt szemlélteti a (7) példában az alábbi informatív tematikus alany:

(7) a valóságnak az ember által különböző módszerekkel és különböző absztrakciós szinteken megismert különféle oldalai között fennálló elvi jellegű feltételezettség és összefüggés… (OM 1.3)

E mögött a nominális szerkezet mögött minimálisan két önálló mondategység rejlik. *Az ember* (A) *megismeri* (Áll) *a valóságot* (T) *különböző módszerek segítségével* (H), és *különböző absztrakciós szinteken* (H), *s az így megismert valóság különböző oldalai* (A) *összefüggnek* (Áll), és *kölcsönösen* (H) *feltételezik* (Áll) *egymást* (T). Az informatív tematikus alanyok esetén a szintemelő fordítás nemcsak kényszermegoldás, mint a visszautaló téma esetében volt, hanem egyenesen kívánatos. Csak így tudja kiemelni a fordító a szerző által közölni akart új információt a névszói szerkezetek rabságából. És így van remény rá, hogy az olvasó is megérti, miről van szó. Hiszen, amint szintaktikai mutatóink számai is jelezték, az autentikus magyar szövegek szerzői az egy mondatra jutó információmennyiség növelését nem a névszói szerkezetek mondatszint alatti bővítettségének fokozásával, hanem az önálló predikátummal rendelkező, önálló mondategységek számának növelésével oldják meg.

Természetesen a szintemelő fordításnak ebben az esetben is vannak veszélyei, hiszen sokszor a megértett gondolat nem maradhat így kibontva, mégiscsak a mondat témájáról van szó, tehát valahogy vissza kell csomagolni. Ha meghagyjuk az önálló mondategységeket, úgy növeljük a kötőszók és utalószók számát is. Mivel pedig gyakran fordulnak elő pótlólagosan betoldott kötőszós és utalószós megoldások a rematikus szakasz fordításakor, a kötőszók és utalószók számának növekedése szintén a fordított szövegek egyik csak szövegszinten nyomon követhető sajátosságává válik. A *hogy* kötőszós mondategységek számának növekedését a fordításban később nagyobb korpuszon és angol–magyar irányban is kimutattam (Klaudy 2009, 2017). Ezt az akkor kiszámolt eredményünket később nagy fordítási korpuszban is kimutatta Olohan és Baker (2000), amikor megállapították, hogy a *that* kötőszó nagyobb számban fordul elő az angolra fordított szövegekben, mint az autentikus angol szövegekben. A (8) példában az utalószók halmozását szemléltetjük egy OM szövegcsoportból vett példán.

(8) *Az a helyzet*, amely a világon az utóbbi három év során kialakult, teljesen nyilvánvalóan igazolja *azt a tényt, hogy*… (OM 3.3)

4.1.3. A retorikus téma

Retorikus témának az olyan mondatkezdő tematikus szakaszokat nevezzük, amelyek lehetnek akár visszautalók, akár informatívak, a mondatban általában alanyi funkciójú, hosszú és bővített esetleg halmozott névszói szerkezettel vannak kifejezve, és utánuk olyan rematikus szakasz következik, amely nem hordoz fontosabb információt, mint a tematikus szakasz, csak összefoglalja a témában mondottakat, leszögezi fontosságukat. Ilyen mondatokat a magyar szónoki nyelvben is találunk, de szakszövegekben ritkán. Ez az a pátosz, amely az akkori orosz nyelvű szakmai és tudományos szövegeket jellemezte. Az autentikus magyar szövegcsoport ilyen szerkesztésű mondatot nem tartalmazott, utoljára Kossuth Lajos beszédeiben találkoztam hasonlókkal. A fordító lehetőségeit szemléltessük az orosz nyelvű szövegcsoport (O 4.2.) egy mondatának három fordítási változatával: (9a) megtartja az orosz retorikus témát, (9b) RT tagolású mondattal fordít, (9c) RT tagolású mondattal és szintemeléssel fordít. Ez utóbbi érthető a legjobban:

(9a) Nagy teljesítőképeségű, pontosan működő nemzetközi szállítmányozási rendszer létrehozása, a szállítások összehangolására szolgáló korszerű módszerek bevezetése, valamint új együttműködési területek kijelölése – ilyen feladatokat tűzött a program a szállítmányozási vállalatok elé.

(9b) A program a következő feladatokat tűzi szállítmányozási vállalatok elé: nagy teljesítőképeségű, pontosan működő nemzetközi szállítmányozási rendszer létrehozása, a szállítások összehangolására szolgáló korszerű módszerek bevezetése, valamint új együttműködési területek kijelölése.

(9c) A program a következő feladatokat tűzi szállítmányozási vállalatok elé: hozzanak létre nagy teljesítőképeségű, pontosan működő nemzetközi szállítmányozási rendszert, vezessenek be korszerű módszereket a szállítások összehangolására, valamint jelöljenek ki új együttműködési területeket.

4.2. Kvázi-helyesség a fordított szövegek rematikus szakaszában

Amint említettük, a fő aktuális tagolási típusok (TR, RT, KR, KT) megoszlásában nincs lényeges számszerű különbség a három vizsgált szövegcsoport között: a szakmai és a tudományos szövegekre jellemző módon, a mondategységek többsége objektív szórendű, azaz az információ adagolása a témától halad a réma felé. A különbséget a tematikus és rematikus szakaszok típusainak megoszlása jelenti.

A rematikus szakaszok szintaktikai megformálását alapvetően befolyásolja az a tény, hogy az orosz mondatokra dominánsan az SVO (szubjektum–verbum–objektum), a magyar mondatokra dominánsan az SOV (szubjektum–objektum–verbum) szórend jellemző. Ennek megfelelően az orosz mondatokban a névszói tematikus szakasz (alany és/vagy határozó és/vagy tárgy) után szinte mindig egy igei állítmány jelöli a téma és a réma közötti szakaszhatárt. A magyar mondatokban viszont gyakran előfordul, hogy a tematikus szakasz után nem ige következik, hanem az igei állítmány hangsúlyos bővítménye, a rematikus csúcs, amely szintén névszói megformálású. Így alakulhat ki, hogy az OM mondatok elején hosszan sorakoznak a névszói szerkezetek, és csak nagyon későn jutunk el a mondat szerkezetét világossá tevő állítmányig.

A rematikus szakasz szintaktikai megformálása alapján az orosz mondatokban négyféle rématípust különítettünk el: R1 = állítmány, R2 = alany és /vagy tárgy és/vagy határozó, R3 = gyenge igei állítmány + alany és /vagy tárgy és/vagy határozó, R4 = erős igei állítmány + alany és /vagy tárgy és/vagy határozó. Az oroszban nincs elváló igekötő, tehát az igei állítmány súlya szemantikai alapon dől el.

Ha a fenti rématípusokat a magyar szövegcsoportban próbáljuk megtalálni, látjuk, hogy az R1, R2 és R4 a magyarban is megvan, az R3 viszont nem létezik. A gyenge igei állítmány a magyarban nem állhat a rematikus szakasz elején, hiszen akkor az előtte álló bővítményt regresszíve rematikus csúccsá tenné. Fel kellett vennem egy ötödik rématípust (R5), amelyben a rematikus szakasz élén hangsúlyos bővítmény áll, ez a rematikus csúcs. A *8. táblázat* mutatja a rématípusok megoszlását az orosz és a magyar szövegcsoport mondategységeiben. Felvettünk még egy hatodik típust is azokra a ritka esetekre (R6), amelyekben a rematikus szakaszon belül két rémacsúcs található.

8. táblázat

A rematikus szakaszok fő típusainak számszerű megoszlása az O és az M szövegcsoportban

RSZ típusa	O		M	
	abszolút szám	**%**	**abszolút szám**	**%**
R1.	35	10,41	25	5, 10
R2.	62	18,45	37	7,55
R3.	122	36,30	---	----
R4.	113	33,63	178	36,32
R5.	---	---	235	47,95
R6.	4	1,19	15	3,06
Összesen	336	100,00	490	100,00

A *8. táblázatból* látható, hogy az R3 és az R5 a két legnépesebb csoport, vagyis éppen abból a rématípusból van a legtöbb az orosz tudományos nyelvben, amelyik nem létezik a magyarban, és fordítva. Tehát akármelyik irányban fordítunk, a rematikus szakaszok nagy részének meg kell változtatni a típusát, és ez az egyszerű szórendi átváltási műveletnek tűnő VO→OV váltás nem végezhető el mechanikusan, mivel az orosz gyenge ige után rendszerint nem egy, hanem több bővítmény áll, és a fordítónak el kell döntenie, mit hoz be a magyar ige elé a rematikus csúcsra.

4.2.1. Emelkedő (R3) és ereszkedő réma (R5)

Ebben a rövid visszatekintésben nincs mód az összes rématípus fordításával kapcsolatos megfigyelések részletes leírására, csak a fenti két legfontosabbra. Ehhez szükségünk van Firbas (1964) „kommunikatív dinamizmus" fogalmának használatára, és négy saját fogalom bevezetésére, amelyeket annak idején a dolgozatban részletesen kifejtettem: a progresszív és regresszív rémajelölés, valamint az ereszkedő és emelkedő réma fogalmára (Klaudy 1987).

Az orosz R3-at azért nevezzük emelkedő rémának, mert élén gyenge szemantikai töltésű ige áll, és a kommunikatív dinamizmus a mondat vége felé egyre nő. A magyar R5-öt azért nevezzük ereszkedő rémának, mert a gyenge vagy fordított szórendű igei állítmány a tőle balra álló bővítményekre helyezi a hangsúlyt, és az utána következő bővítmények kevesebb hangsúlyt kapnak.

Az orosz emelkedő rémában a rematikus csúcs jelölése progresszív (előre mutató), mert már a rematikus szakasz elején tudjuk, hogy a gyenge szemantikai töltésű igei állítmány után egyre hangsúlyosabb bővítmények következhetnek. A magyar ereszkedő rémában a réma jelölése regresszív (hátra mutató), mert a fordított szórendű vagy gyenge szemantikai töltésű igék utólag jelölik ki a tőlük balra álló rematikus csúcsot, és ha a fordító minden bővítményt behoz az ige elé, akkor csak a mondat végén világosodik meg a mondat szerkezete. Ezt mutatja az alábbi mondat az OM szövegcsoportból:

(10) [TSZ # A gazdasági fejlődés jelenlegi szakaszában tökéletesítésük szükségességét] # [RSZ (RCS a társadalmi termelés méreteinek növekedése, a népgazdasági kapcsolatok növekvő bonyolultsága és fokozódó intenzifikálása, a tudományos-technikai haladás ütemének meggyorsulása, az áru- és pénzviszonyok további fejlődése, és a tervezés módszereinek és a központi szervek szerepének megváltozása) határozza meg].

A (10) példában látható, hogy a fordított szórendű igei állítmány előtt (*határozza meg*) 28 szó áll a rematikus csúcson. De a tematikus szakasz utolsó szava után (*szükségességét*) semmi nem jelezte, hogy itt most a rematikus szakasz következik, ezt csak a mondat végén tudjuk meg. Azt, hogy a fordításokban az állítmány gyakran a mondat végére csúszik, a lektorok nyilván gyakran észreveszik és javítják. Nem tudatosan, de megpróbálják progresszívvé tenni a regresszív rémajelölést. Mondhatnák azt, hogy ne hozzunk be mindent az ige elé a rematikus csúcsra, használjuk ki a fordított szórendű igei állítmány utáni semleges zónát. Itt is megtehette volna a fordító, hogy a fordított szórendű igei állítmányt előreviszi, és a felsorolás első tagja után helyezi el (*a társadalmi termelés*

méreteinek növekedése határozza meg), a többi bővítményt pedig kiviszi az ige utáni semleges zónába, például a *valamint* beszúrásával.

Mindezt számszerűen is ki tudtuk mutatni. A *9. táblázatból* látható, hogy a rematikus csúcs hossza és jelölésének módja eltér az autentikus magyar és a fordított magyar szövegekben.

9. táblázat

A rematikus csúcs (RCS) hossza és jelölésének módja az M és az OM szövegcsoportban

Az összevetés egységei	**M**	**OM**
RCS átlagos hossza	3,2 szó	3,8 szó
progresszíven jelölt RCS aránya	47,21%	30,55%
regresszíven jelölt RCS aránya	52,78%	69,44%
progresszíven jelölt RCS átlagos hossza	3,03 szó	3,32 szó
regresszíven jelölt RCS átlagos hossza	3,13 szó	4,10 szó

A *9. táblázatból* látható, hogy az OM szövegekben hosszabbak a rematikus csúcsok, és több a csak regresszíven jelölt rematikus csúcs, mint az M szövegcsoportban. Mivel a magyar R5-re jellemző regresszív rémajelölés megértési nehézséget okozhat, az autentikus magyar szövegekben a regresszíven jelölt rematikus csúcsok nem szoktak túl hosszúak lenni, ritka bennük a halmozás, és a sok mondatszint alatti bővítmény elhelyezése. Ha a rematikus csúcson több halmozott és/vagy bővített névszói szerkezet van, az bizonytalanná teszi a TR szakaszhatárt. Ezt mutatja a (12) példa:

(12) [$_{TSZ}$ # Az ily módon kialakuló negatív emocionális alap, a pesszimizmus] # [$_{RSZ}$ ($_{RCS}$ elmaradott egészségtelen nézetek, szokások, gondolatok) jó talajává válhat]. (OM 11.5)

Ha nem lennének zárójelek, nem tudnánk, hogy a pesszimizmus után vége a tematikus szakasznak, és a rematikus csúcs következik, inkább felsorolásnak, illetve halmo-

zásnak gondolnánk, hiszen a grammatikai homonímia miatt az alany csak utólag válik el a hasonló szintaktikai formájú birtokos jelzőtől.

A regresszív rémajelölés progresszívvé tételét úgy is meg lehet oldani, hogy a rematikus csúcsnak csak a képviselőjét hozza be a fordító a gyenge igei állítmány elé. Ezt a műveletet a kezdő fordítók nem végzik el, a rutinosabbak elvégzik. Ezt illusztrálja a (12) példa két változata:

(12a) A tudásnak ez a két ága egymással szorosan összefüggő céljait a tudomány tényeinek minőségi elemzése alapján éri el. (hallgatói fordítás)

(12b) A tudásnak ez a két ága úgy közelíti meg egymással szorosan összefüggő céljait, hogy minőségileg elemzi a tudomány tényeit. (OM 1.11)

Mint a (12a) példából látható, a kezdő fordító minden bővítményt a gyenge igei állítmány (*éri el*) előtti rematikus csúcson helyezett el, míg a képzett fordító mondatában (12b) a rematikus csúcson csak a határozó bővítmény képviselője, az úgy utalószó található, a tárgyi bővítményt a fordított szórendű igei állítmány utáni semleges zónában találjuk, a határozói bővítményt pedig a külön mondategységben helyezi el a fordító.

4.2.2. A többi rématípus

Az R1-ben a rematikus szakaszban egyetlen mondatszintű elem van: az igei vagy névszói állítmány. Mindkét nyelvre jellemző rématípus. Fordítása csak akkor okoz problémát, ha az orosz névszói állítmány szemantikailag gyenge, és utána sok posztpozitív bővítmény sorakozik. Ha a fordító az orosz mondat összes jobbra álló bővítményét behozza a magyar főnév elé, megint elmosódik a TR szakaszhatár, ahogyan a (13a) példa mutatja.

(13a) Az enyhülés megvalósítása az államok közötti kapcsolatok kölcsönösen egyeztetett és szerződésben rögzített elveinek gyakorlati tartalommal való megtöltésére hivatott nemzetközi intézkedések széles körét felölelő fogalom. (hallgatói fordítás)

(13b) Az enyhülés megvalósítása nemzetközi intézkedések széles körét felölelő fogalom, ezek az intézkedések arra vannak hivatva, hogy… (OM 3.3)

A (13a) mondat grammatikailag helyes magyar mondat, de csak névszói szerkezetekből áll, és nehéz megérteni. A megjelent fordítás (13b) rutinosabb fordítója, előrehozta a támpontot, azaz a névszói állítmányt egyetlen bővítménnyel, és a többi bővítményt önálló mondategységben helyezte el. Itt tulajdonképpen új mondatot is kezdhetett volna, de nyilván nem akarta megváltoztatni a mondathatárokat.

Az R2-ben a rematikus szakaszban szintén csak egyetlen mondatszintű elem van, de nem az állítmány, hanem valamelyik másik mondatszintű bővítmény, az alany, a tárgy vagy a határozó. Amikor a grammatikai tagolás és az aktuális tagolás különbségeiről beszélünk, mindig hangsúlyozzuk, hogy a téma nem mindig az alany, a réma nem mindig az állítmány, ezért érdemes egyáltalán aktuális tagolásról beszélni. A *10.* és *11. táblázat* mutatja a rematikus alanyok arányát és mondatszint alatti bővítettségét a három szövegcsoportban.

A *10. táblázatból* leolvasható, hogy az alanyra valóban elsősorban a témaszerep a jellemző, hiszen az alanyok nagy része mindhárom szövegcsoportban a tematikus szakaszban található. A *11. táblázat* azt mutatja, hogy a rematikus alanyok bővítettsége mindhárom szövegcsoportban nagyobb, mint a tematikus alanyoké. Az is látszik, hogy az orosz szövegcsoportban található a legtöbb rematikus alany.

Ezek az adatok alátámasztják tapasztalati úton nyert feltételezésemet, hogy az orosz szakmai és tudományos nyelvben valóban létező tendencia az alanyok rematikus helyzetbe való juttatása és a lényeges mondanivaló rematikus alany formájában való kifejezése, ami a főnevesülési tendencia jele.

Ezekben az oroszból fordított magyar mondatokban nincs ige, amely segítené a mondat szerkezetének megértését.

10. táblázat

Az alanyok megoszlás a téma és a réma között

Alanyi megformálás	**O**	**OM**	**M**
A tematikus szakaszban	68,07%	70,1%	74,75%
A rematikus szakaszban	31,93%	29,9%	25,25%

11. táblázat

Az alanyok átlagos mondatszint alatti bővítettsége a témában és a rémában

Az alanyok átlagos mondatszint alatti bővítettsége	**O**	**OM**	**M**
A tematikus szakaszban	2, 15 szó	1,25 szó	1,18 szó
A rematikus szakaszban	4, 58 szó	2,88 szó	1, 83 szó

Az orosz eredetiben ilyenkor gyakran találunk является típusú üres igéket, de ezek a magyar fordításban nem jelennek meg.

(14) A fűtőanyag- és energiaprobléma megoldásának a célprogramokban megjelölt fő útjai: a szilárd fűtőanyagok termelésének növelése, a kőolaj-feldolgozás bővítése, az atomenergia-ipar fejlesztése, és az energiaigényes iparágaknak a fűtóanyag-termelő területeken való elhelyezése. (OM 5.10)

A (14) példamondathoz hasonló halmozott és bővített alanyi rémák számszerűen is kimutatható gyakorisága igazolja azt az érzésemet, hogy az oroszból fordított magyar szövegek nominálisabbak, mint az eredeti magyar szövegek.

Egyenletes rémáról (R4) akkor beszélünk, ha a rematikus szakasz elején erős igei állítmány áll, amely az egész rematikus szakasznak egyenletes kommunikatív töltést ad.

Ezt a magyarban az egyenes szórendű igei állítmány vagy erős szemantikai töltésű ige teszi lehetővé. Mivel az oroszban nincs elváló igekötő, ott csak az ige szemantikai töltése játszik szerepet. Az erős igei állítmány balra tematikus szakaszt jelöl, jobbra egyenletes rémát. Ez a rématípus mindkét nyelvben előfordul, fordítása nem jelent nehézséget.

5. Záró gondolatok

Az elmúlt 40 évben nemcsak a korpusznyelvészet fejlődött, hanem jelentős társadalmi és politikai átalakulások is történtek. Ezeknek az orosz szövegeknek nagy részét ma már nem fordítanák le magyarra, míg a magyar szövegcsoportom szerzőit (pl. Hankiss Elemér, Huszár Tibor, Kulcsár Kálmán, Pataki Ferenc, Hanák Péter stb., vö. Klaudy 1987) ma is olvassák. Ezért is kell vigyázni a korpusznyelvészeti munkákban a szövegválasztással. A vizsgált korpusz orosz és oroszból fordított magyar szövegein nagyon érezni lehet a kort, amelyben születtek. Ez nem érvényteleníti a dolgozat alaptéziseit, hogy a professzionális fordítók által készített szövegek jellegzetességeit szövegszinten érdemes vizsgálni, hogy a fordított szövegeket autentikus célnyelvi szövegekkel érdemes egybevetni, és hogy disztribúciós különbségekről van szó, amelyeket számszerűleg is ki lehet mutatni.

Lett-e mindennek valami folytatása? A manuális elemzésnek és számolásnak nagy előnye a szövegközeliség. A szövegek aprólékos vizsgálata közben sok mindenre felfigyel az ember, amit az automatikus elemzés nem észlel. Például valaminek a hiányára. Ilyen volt a TR szakaszhatár jelölésnek hiánya a fordított szövegekben, illetve ennek a kommunikatív szakaszhatárnak az elmosódása, amit később európai uniós szövegek anyagán, angol end-focus jellegű (R3) mondategységek magyarra való fordításában is kimutattam (Klaudy 2004). Az orosz–magyar fordításokon végzett megfigyeléseimet beépítettem az indoeurópai nyelvekről magyarra és magyarról indoeurópai nyelvekre való fordítás átváltási műveleteinek tipologizálásába (1994, 2003). Építettem később kisebb automatikusan lekérdezhető angolból fordított magyar és autentikus magyar korpuszt is a fordításokat

leleplező magyar nyelvi elemek kimutatására (2018). Két kutatót említek, aki folytatta a fordított szövegek aktuális tagolásának ilyen szellemű vizsgálatát: Aradi András (2019) és Pásztor Kicsi Mária (2012). A Fordítástudományi Doktori Program hallgatói közül Nagy János (2016) vágott bele angolról magyarra fordított szépirodalmi szövegek aktuális tagolásának vizsgálatába nagyon érdekes eredményekkel. Az aktuális (értelmi vagy kommunikatív) tagolás kutatásának korpusznyelvészeti eszközökkel való összekapcsolása azonban még mindig várat magára.

Amikor a nyolcvanas évek elején befejeztem a dolgozatomat, a Budapesti Műszaki Egyetem Számítóközpontjának egy munkatársa a következő megjegyzést tette: kár, hogy nem vártam a disszertációm megírásával egy-két évet, mert nemsokára az általam vizsgált 60 könyvet teljes egészében gépre lehet majd vinni, és automatikusan elemezni. Ha nem is egy-két év alatt, de valóra vált az álom. Mint a korpusznyelvészet későbbi fejlődése és e kötet további tanulmányai is mutatják, most már több millió szavas korpuszok automatikus elemzése folyik, cáfolva vagy igazolva a korai, kis korpuszokon manuálisan végzett elemzések megállapításait.

Irodalom

Aradi A. 2019. A mondatok információs szerkezetének és szintaktikai formájának viszonya a fordításban. *Fordítástudomány* 21. évf. 1. szám. 5–23.

Baker, M. 1993. Corpus Linguistics and Translation Studies. Implications and Applications. In: Baker, M., Francis, G. and Tognini-Bonelli, E. (eds) *Text and Technology: In honour of John Sinclair.* Amsterdam: Benjamins. 233–250.

Baker, M. 1995. Corpora in Translation Studies. An Overview and Suggestions for Future Research. *Target* Vol. 7. No. 2. 223–245.

Deme L. 1971. *Mondatszerkezeti sajátosságok gyakorisági vizsgálata.* Budapest: Akadémiai Kiadó.

É. Kiss K. 1978. A magyar mondatok egy szintaktikai modellje. *Nyelvtudományi Közlemények* 80. évf. 2. szám. 261–265.

Firbas, J. 1964. On Definig the Theme in Functional Sentence Analysis. *Travaux linguistiques de Prague* Vol. 1. 267–280.

Klaudy K. 1987. Fordítás és aktuális tagolás. *Nyelvtudományi értekezések* 123. Budapest: Akadémiai Kiadó.

Klaudy K. 1994. *A fordítás elmélete és gyakorlata*. Budapest: Scholastica.

Klaudy K. 2003. *Languges in Translation*. Budapest: Scholastica.

Klaudy K. 2004. A kommunikatív szakaszhatárok eltűnése a magyarra fordított európai uniós szövegekben *Magyar Nyelvőr* 128. évf. 4. szám. 389–407.

Klaudy K. 2018. Eredeti magyar szöveg vs. fordított magyar szöveg. In: Bódi Z., Hegedüs, R., Szöllősy-Sebestyén A. (szerk.) *Magyar szaknyelvek a Kárpát-medencében*. Budapest: Tinta Könyvkiadó.

Nagy J. 2012. *A kommunikatív dinamizmus (relatív) egyensúlya a fordításban*. Doktori értekezés. Kézirat. Budapest: ELTE.

Olohan, M., Baker, M. 2000. Reporting *that* in Translated English. Evidence of subconscious processes of explicitation? *Across Langugaes and Cultures* Vol. 1. No. 2. 141–158.

Papp F. 1972. Okoncsatyelnaja redakcija tyeksztovih jegyinyic dlinnyee predlozsenyija. *Slavica* XII. 27–41.

Papp F. 2005. A mondatnál hosszabb szövegegységek végső szerkesztése, avagy az idegen nyelvű beszéd kvázi helyessége In: Klaudy K. (szerk.) *Papp Ferenc olvasókönyv.* Budapest: Tinta. Fordította: Répási Györgyné. 122–135.

Pásztor Kicsi M. 2012. *A mai vajdasági magyar napi sajtó és elektronikus média informatív szövegeinek szintaktikai, intonációs és kommunikatív jellemzői.* Újvidék: Újvidéki Egyetem, Bölcsészettudományi Kar.

Yngve, V. 1973. A mélységhipotézis. In: Szépe Gy. (szerk.) *A nyelvtudomány ma.* Budapest: Gondolat. 443–458.

Korpuszalapú fordítástudomány, és ami mögötte van

Balaskó Mária

balasko.maria@sek.elte.hu

ELTE Savaria Egyetemi Központ

Angol Nyelv és Irodalom Tanszék

Kivonat: A korpusznyelvészetet és a leíró fordítástudományt egyaránt meghatározó szemléletmód értelmében a nyelvet úgy kell leírni, ahogyan az a természetes nyelvhasználatban létezik, azaz valódi, nem művi nyelvi adatok alapján. Ezzel az elméleti háttérrel a két tudományterület egymásra találása törvényszerű volt miután a fordítástudományban bekövetkezett paradigmaváltásnak köszönhetően a fordítás eredményeképpen keletkezett szöveget már nem másodrangú, hanem egyenértékű szövegnek tekintették. Az alábbiakban a brit nyelvészeti tradíció néhány meghatározó elvét ismertetem és mutatom be, hogyan járult hozzá a korpuszalapú fordításkutatás kialakulásához, majd a mellett érvelek, hogy a korpusznyelvészetben kidolgozott analitikus eszközöket alkalmazni lehet a fordított szöveg nyelvi mintázatainak vizsgálatára.

Kulcsszavak: brit nyelvészeti tradíció, fordítási univerzálék, kollokáció, korpusz, nyelvi mintázatok

1. Bevezetés

Huszonöt év telt el, amióta megjelent az első tanulmány, amely felvázolta a korpuszalapú fordításkutatásban rejlő lehetőségeket, mégpedig Mona Baker (1993) tollából. A cikk két önálló kutatási terület, a korpusznyelvészet és a fordítástudomány lehetséges kapcsolódási pontjait körvonalazta. Nem mellesleg e sorok írásának időpontjában az említett tanulmány több mint 1.700 független hivatkozással rendelkezik; tudománytörténeti jelentőségén túl

Balaskó Mária: Korpuszalapú fordítástudomány, és ami mögötte van. In: Robin E., Seidl-Péch O. (szerk.) 2020. *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordításkutatás Magyarországon.* Segédkönyvek a nyelvi közvetítésről I. Budapest: ELTE BTK Fordítástudományi Doktori Program, MANYE Fordítástudományi Szakosztály. DOI: https://doi.org/10.36252/Nyelvikozvsegedkonyv1.3

annak is köszönheti ismertségét, hogy a 20. század második felének egyik legjelentősebb, iskolateremtő nyelvésze, John Sinclair tiszteletére szerkesztett ünnepi könyvben jelent meg (Baker, Francis és Tognini-Bonelli 1993). Mona Baker John Sinclair-tanítvány volt egy olyan időszakban, amikor a lehető legkedvezőbb együttállás segítette a korpuszalapú fordításkutatás megszületését: a számítógépes technológia elegendően fejlett és felhasználóbaráttá vált, a Sinclair-iskolában jelentős tapasztalat és tudás halmozódott fel a számítógépes korpuszok feldolgozásáról, és szintet lépett a fordítástudomány, amikor a leíró fordítástudomány (Toury 1995) felé fordult.

A korpusznyelvészet és a leíró fordítástudomány egymásra találása törvényszerű volt, ugyanis mindkét tudományterület arra a szemléletmódra támaszkodik, hogy a nyelvet úgy kell leírni, ahogyan az a nyelvhasználatban létezik, mégpedig valódi, nem mesterségesen előállított, kitalált nyelvi adatok alapján. Ennek a szemléletnek az elfogadásához a fordítástudományban szükség volt arra a paradigmaváltásra, amelynek köszönhetően a fordítás eredményeképpen keletkezett szövegre már nem másodrangú, hanem egyenértékű szövegként tekintettek.

A korpusznyelvészet fejlődését eleinte nemcsak a technológia kezdetlegessége akadályozta, hanem a nyelvtudomány fősodrát a 20. század második felében uraló racionalista megközelítésmód is. Ezzel egy időben, ennek a szemléletmódnak az ellenpólusán működött egy szellemi műhely, amelynek alapjait az 1930-as években J. R. Firth fektette le (1935, 1957a, 1957b) a korábbi brit nyelvészeti hagyományokat követve, majd tanítványainak egymást követő generációi, M.A.K. Halliday és John Sinclair fejlesztettek tovább.

Az úgynevezett brit tradíció egyik alaptétele, hogy a nyelvészetet társadalomtudománynak és alkalmazott tudománynak tekinti, ami teljesen szemben áll a racionalista felfogással. A másik fontos ismérv az előzőből következik: amennyiben a nyelv társadalmi jelenség, azt természetes módon keletkezett, létező szövegekből kiindulva kell vizsgálni. Megkülönböztetett figyelmet kap a jelentés, ami idővel arra a felismerésre vezette a neo-firthiánusokat, hogy nyelvi megformálás és jelentés elválaszthatatlanok egymástól, és hogy nyelvtan és szókészlet szorosan összefüggenek egymással. Ezek a gondolatok hatá-

rozták meg John Sinclair teljes munkásságát, benne a Sinclair által képviselt korpusznyelvészeti irányzatot, és hatottak közvetve Mona Baker fordításról való gondolkodására is.

Noha Mona Baker a fordítási univerzálék (1993; 1996) meghatározásával alapozta meg a korpuszalapú fordításkutatást, ezek egy tágabb konceptuális keretben helyezkednek el, amelynek értelmében a fordítás eredményeként létrejött szöveg természetében más, mint az eredetileg adott célnyelven alkotott szövegek. A fordítás a nyelvhasználat egyedi változata, ugyanis a szövegprodukció minden egyéb fajtájától eltérően korlátozóan hat rá egy másik nyelven megfogalmazott szöveg (Baker 2005). Azaz a fordítás eredményeképpen létrejött szöveg egy olyan folyamat produktuma, amelynek során a fordító más gondolatait más szavaival más nyelven fogalmazza újra. Ebből természetesen következik, hogy a fordítás saját jellegzetes nyelvi mintázatokba rendeződik.

2. A nyelvi mintázatokról

A nyelvszerkezetet a szókészlettől független rendszernek tekintő felfogás a múlté. Számos önálló tudományterületen végzett kutatás jutott egymástól függetlenül ugyanarra az eredményre, hogy az élő nyelv morfémák többé-kevésbé rögzített sorozatából áll. Ezeket a sorozatokat elméleti háttértől függően nevezik lexikális kifejezéseknek (lexical phrases) (Nattinger és DeCarrico 1989; 1992), kompozitáknak (composites) (Mitchell idézi Cowie 1988), nyitólépéseknek (gambits) (Keller idézi Cowie 1988), rutinszerű képleteknek (routine formulae) (Coulmas idézi Cowie 1988), frazémáknak (phrasemes) (Melčuk 1988; 1995), előre gyártott rutinoknak és mintázatoknak (prefabricated routines and patterns) (Krashen 1981), és még sorolhatnánk. Más szóval, az élő nyelv előre gyártott elemekből, panelekből építkezik. Ebben a kontextusban nyer értelmet Sinclair (1991a) frazeológiai elve (idiom principle), amely korpusznyelvész és lexikográfusi tapasztalata szerint mélyen áthatja az írott és a beszélt nyelvet. Noha megállapításait a világviszonylatban is legtöbbet kutatott angol nyelv vonatkozásában fogalmazta meg, feltételezhetjük, hogy a többi nyelv esetében is ez a működési elv érvényesül.

A frazeológiai elv szerint a szöveg nem véletlenszerűen kiválasztott egymást követő szavakból keletkezik, hanem számos előre vagy félig előre gyártott szerkezetből áll össze, amelyek adott helyzetben az egyetlen választási lehetőséget jelentik akkor is, ha látszólag ezek a szerkezetek szétválaszthatók. A frazeológiai elv működésének egyik klasszikus példája a két szóból álló angol of course. Ez a szókapcsolat gyakorlatilag egyetlen szóként működik és a szóköz idővel ki fog kopni, mint ahogy az történt a maybe, anyway, another szavak esetében. Az of ebben a szerkezetben nem a nyelvtankönyvekben megtalálható of elöljáró, amely általában egy főnévi csoport főneve után következik, vagy egy mennyiségjelzőben, mint a pint of. Hasonlóképpen a course sem az a megszámlálható főnév, amely a szótárakban szerepel, jelentése pedig nem a szó tulajdonsága, hanem a szókapcsolaté (vö. Sinclair 1991a).

A frazeológiai elvvel szemben áll a szabad választás elve (open-choice principle) (Sinclair i.m.), amely a szöveget egy bonyolult kiválasztási folyamat eredményének tekinti, amit egyedül a nyelvtani helyesség korlátoz. Azaz minden szót, szókapcsolatot vagy mondatot bármilyen más szó, szókapcsolat, mondat követhet, ha az megfelel a nyelvtani helyesség szabályainak. Könnyen belátható, hogy ha a nyelvben csupán a szabad választás elve működne, nem tudnánk elfogadható szöveget alkotni.

A korpusznyelvészeti kutatások (Sinclair (1991a), Kjellmer (1994), Altenberg és Olofsson (1990) és Moon (1998)) kiterjesztették a frazeológia határait, hogy az magába foglalja a klasszikus, ámde nyelvhasználati gyakoriságukat tekintve meglehetősen ritka, átvitt értelmű kötött frazeológiai egységeken kívül, a szavak számos olyan laza kombinációját, amelyekben gyakori az együttes előfordulás. Mivel ezek a többszavas kifejezések gyakran ismétlődnek és szokványosak, a hétköznapi nyelvhasználatban észre sem vesszük, hogy előre gyártott elemeket használunk. Mindezek ellenére a több szóból álló egységek képezik az írott és a beszélt nyelv tipikus és előnyben részesített építőelemeit. A több szóból álló egységek együttes kiválasztás (koszelekció) eredményeként jönnek létre. A koszelekció azt jelenti, hogy a nyelvhasználat két vagy több elemet rendszeresen, szokásszerűen „csomagként” választ ki. A szavak koszelekciós hajlama kiemelkedően erős.

A több szóból álló előre gyártott elemek építőkövei a kollokáció és a kolligáció. A fogalmak, melyek közül az előbbi széles körben ismert, Firthtől származnak, aki egy népszerű közmondást módosítva írja körül a jelenségeket: „You shall know a word by the company it keeps […]" (Firth 1957b: 11), azaz madarat tolláról, szót barátjáról. Fontos hangsúlyozni, hogy a szóban forgó szintagmatikus viszonyok konkrét szavak között, és nem kategóriák között keletkeznek. A kollokáció két vagy több szó ismétlődő együttes előfordulása, míg a kolligáció az a grammatikai mintázat, amelybe egy állandósult szókapcsolat illeszkedik. Másképpen: bizonyos szavak és szóalakok vagy vonzzák, vagy kerülik egymás társaságát. A kollokációkat és a kolligációkat korpuszokból készített konkordanciák hozzák felszínre, amelyek függőleges oszlopaiban kirajzolódnak az ismétlődő mintázatok vagy panelek.

A Sinclair vezetésével megvalósult két korszakalkotó korpuszalapú munka, a Collins Cobuild English Dictionary (Sinclair 1987) és a Collins Cobuild English Grammar (Sinclair 1990) tanulságait Stubbs (1993) összegzi. Az alábbiakat tartom közülük ezen a helyen fontosnak kiemelni:

1. Minden nyelvtani szerkezet megszabja, milyen szókészleti elemeket lehet benne felhasználni, és viszont: minden szókészleti elem meghatározható azoknak a nyelvtani szerkezeteknek az alapján, amelyekben előfordul.
2. A ragozási paradigmában nem állandó, azaz változik a lemma jelentése.
3. Egy szó minden külön jelentésének megvan a saját grammatikája; mindegyik jelentéshez egyedi, jellegzetes nyelvi mintázat kapcsolódik. Nyelvi megformálás és jelentés nem választhatók el egymástól.
4. A nyelv szokásszerű használatát a koszelekció, vagyis az együttes kiválasztás irányítja.

A fentiek szemléltetésére álljon itt néhány példa.

Az első pont érvényességét támasztja alá az információs szerkezetről szóló fejezet fókuszt tárgyaló részénél a *Collins Cobuild English Grammar* leírásában (Sinclair 1990: 410) található alábbi kolligációs mintázat:

all / *what* + PRONOUN + VERB + BE + NOUN

Mint például *All you need is love*. A 7,2 millió szövegszóból álló korpusz alapján végzett elemzés kimutatta, hogy a szerkezetben előforduló igék száma egynéhányra korlátozódik, mi több, azok szemantikailag egymáshoz kapcsolódnak: *adore*, *dislike*, *enjoy*, *hate*, *like*, *loathe*, *love*, *need*, *prefer*, *want*. Más igék nem fordulnak elő ebben a mintázatban.

Egy másik megállapítást igazol a *set in* többszavas ige elemzése. Egyik alakja, a *set*, sokkal többször fordul elő, mint a ragozott *sets* vagy a toldalékolt *setting* alak. Konkordanciavizsgálatok kimutatták, hogy az ige leggyakoribb használata múlt idejű, általában (tag)mondat vége felé helyezkedik el, és alanyai döntően elvont főnevek, mégpedig a kellemetlen fajtából: *rot*, *disillution* (Sinclair 1991a).

Ellenőrzésképpen lekérdeztem a többszavas ige konkordanciáját a *MicroConcord* (Scott és Johns 1993) kb. 2 millió szövegszóból álló korpuszából. Az önálló *set* ige összesen 41-szer fordult elő, és ebből 32 esetben követte az *in* prepozíció. Terjedelmi okokból alább csak a többszavas igét tartalmazó maradék konkordanciasorokat közlöm:

Ez a néhány példa is szemléletesen mutatja, hogyan rendeződik a kulcsszó (*set in*) a Sinclair által leírt mintázatba. Az alanyok (ahol láthatóak) különösen figyelemre méltóak: *deterioration*, *unmitigated disaster*, *error and madness*, *deterioration in mood*, *complications*, *foot-rot*, *loss/decline in confidence*.

Az a tény, miszerint egy kisméretű korpuszban az összes előfordulás a nagyságrendekkel nagyobb korpusz eredményeit tükrözi, megerősíti, hogy a leírt nyelvhasználati mintázat tipikus.

1. ábra

A set in konkordanciája a MicroConcord korpuszban (Scott és Johns 1993)

we expected. Clearly some deterioration has set in." <p> He said that Sky `
to unmitigated disaster only seems to have set in from 1942 onwards. <p> F
as involvement of the fallopian tubes, have set in. Rectal discharge may oc
on remained passive until error and madness set in near the end; then she
ntinental ally) and a deterioration in mood set in owing to the threatening
our. These complications are only likely to set in if the infection has bee
eal with the foot-rot which was starting to set in. Another problem was the
imself. <p> The loss of confidence which set in after Stalingrad was not
the inevitable decline in confidence which set in during the following wee

Az utolsó példa egy hétköznapi főnév egyes és többes számú alakjainak eltérő viselkedését mutatja be. Az *eye*, illetve az *eyes* alak teljesen különböző mintázatokban jelenik meg. A többes számú alak például a *blue*, *brown*, *covetous*, *manic* melléknevekkel együtt fordul elő, míg az egyes számú alak szinte soha nem utal a látószervre. Ugyanakkor a főnév tipikus használatára egyes és többes számban egyaránt az idiomatikus szerkezet és az átvitt értelmű jelentés jellemző: *all eyes will be on*, *rolling their eyes*, *turn a blind eye*, *keep an eye on* (Sinclair 1991b).

A fenti példák látszólag annyira triviálisak, hogy könnyű alábecsülni a megállapítások jelentőségét. Hétköznapiságuk ellenére – vagy épp miatta – ezeket és a hozzájuk hasonló nyelvi jelenségeket korpuszok és belőlük készített konkordanciák nélkül nem lehetett volna azonosítani. Az élő nyelv természetének és viselkedésének megismerésére az intuíció nem annyira alkalmas – még a nyelvészé sem –, mint azt feltételezték; több engedményt tesz a *lehetséges* számára és elsiklik a *tipikus* felett. Ha valaki kérte már ki nem anyanyelvűként anyanyelvi beszélő véleményét, annak ismerősnek hangozhat az egyenválasz: „igen, helyes; de várjunk csak, most, hogy alaposan belegondolok, nem *szokták* így mondani."

A korpusznyelvészeti vizsgálatok a nyelvi valóság semmilyen más eszközzel nem megfigyelhető részletgazdagságát tárják fel, és korábbi ismereteinkhez képest eredményei egy teljesen újszerű, esetenként meglepő képet alakítanak ki nyelvről és nyelvhasználatról. A Firth által elméleti alapokra helyezett, majd kutatási eredményeik alapján követői által továbbfejlesztett brit nyelvészeti tradíció elsőként egy lexikográfiai munkában teljesedett ki. Ez volt a John Sinclair főszerkesztésében kiadott, 2018-ban 9-ik kiadásánál tartó *Collins Cobuild English Dictionary* (1987), amely forradalmasította a szótártudományt, és követendő mintául szolgált minden jelentős, elsősorban nyelvtanulói szótárakban érdekelt brit könyvkiadó számára – a Longmantől az Oxfordig – a rákövetkező években megjelenő megújult kiadásaik elkészítéséhez. A szótárkészítés tapasztalata alapján, és ugyanannak a közben folyamatosan bővülő korpusznak a felhasználásával jelent meg szintén John Sinclair irányítása alatt a *Collins Cobuild English Grammar* (1990) funkcionális grammatika, amelyben a szó jelentése kiemelt szerephez jut. A grammatika segédszerkesztői között megtaláljuk Mona Baker nevét is. Ennek a két főmunkának „melléktermékeként" jelent meg tíz további segédkönyv *Collins Cobuild English Guides* 1–10 gyűjtőcím alatt, amelyek elsősorban nyelvtanulóknak nyújtanak segítséget többek között a névelők, kötőszók, elöljárószók használatához. A Cobuild-műhely korpusznyelvészeti módszere és az ott képviselt nyelvészeti szemlélet átterjedt egyéb nyelvtudományi és alkalmazott nyelvészeti területekre, amelyeket ma már lehetetlen mind számba venni. A teljesség igénye nélkül: frazeológia (Moon 1998), szókészlet (Hoey 2005), lexikális szemantika (Stubbs 2001), diskurzuselemzés (Tognini-Bonelli és Del Lungo Camiciotti 2005), idegennyelv oktatás (Tognini-Bonelli 2001), fordítástudomány (Olohan 2004). A mintázatok vizsgálata egyre nagyobb érdeklődést vált ki, és elsősorban – ismét csak néhány témát kiemelve – a lexikai elemek grammatikai mintázatainak (Hunston és Francis 2000) vagy a szövegszerveződés mintázatainak a leírására (Hoey 1991), a nyelv és kultúra összefüggéseinek vizsgálatára (Stubbs 1996) és a kontextusnak a lexikai egység megértésében játszott szerepére (Hanks 2013) fókuszál.

3. Korpuszok a fordításkutatásban

Függetlenül attól, hogy egy- vagy többnyelvűek, a fordítás kutatására összeállított korpuszok elsődleges tulajdonsága, hogy két komponensből vagy modulból állnak. Ez természetes, hiszen minden elemzésnek, de még a leírásnak is az összehasonlítás a módszertani alapja: célnyelvi szöveget hasonlítunk össze forrásnyelvi szöveggel, célnyelvi szöveget azonos célnyelven keletkezett elsődleges szöveggel, egy adott célnyelvű szöveget egy másik célnyelvű szöveggel, eltérő forrásnyelvekből származó célnyelvi szövegeket egymással és így tovább.

Más fiatal (rész)diszciplínákhoz hasonlóan a korpuszalapú fordítástudományban is vannak megoldásra váró terminológiai kérdések. Leginkább a korpusz-típusok elnevezése okoz következetlenségeivel nem kevés zavart. Különböző szerzőknél, kutatóknál más elnevezés alatt találjuk meg ugyanazt a korpusz típust, illetve azonos elnevezéssel más típusú korpuszra utalnak (Seidl-Péch 2018:183–184). Ezért mindig fontos tisztázni, melyik szerző pontosan mit ért az általa használt fogalmon.

Az alábbiakban három fordításhoz kapcsolódó korpuszról lesz szó: a párhuzamos, a fordítási és az összehasonlítható korpuszról. A párhuzamos korpusz azonos kritériumok szerint válogatott két vagy több forrásnyelvű szövegeket tartalmaz; a korpuszban szereplő szövegek nem fordításai egymásnak, hanem hasonló tartalmú forrásnyelvi szövegek, mint például a *Council of Europe Multilingual Lexicography Project* vagy *The Aarhus Corpus of Contract Law*. Látszólag nincs sok közük a fordításhoz, mégis értékes adatforrások azáltal, hogy eredeti célnyelvi mintákat mutatnak a fordítók, a fordító képzésben részt vevők és a fordításkutatók számára egyaránt.

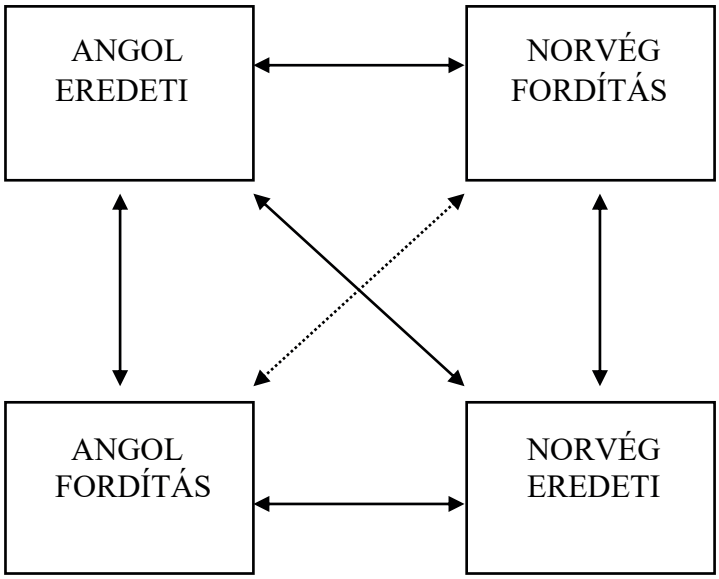
A fordítási korpusz forrásnyelvi szövegeket és azok célnyelvi megfelelőit tartalmazza. Nem meglepő módon ez a leggyakoribb korpusz a fordítói gyakorlatban, különösen, ha figyelembe vesszük, hogy minden fordítómemória gyakorlatilag folyamatosan új szövegpárokkal bővülő fordítási korpusz, de a kezdeteknél ott találjuk a kanadai francia–angol *Hansard Corpust*. Gyakorló fordítóknak, leendő fordítóknak segítenek gyorsabban megtalálni a célnyelvi megoldásokat vagy profi fordítók mintaként használható megoldásait, fordításkutatóknak pedig lehetővé teszik a fordítási normák vizsgálatát.

Az összehasonlítható korpusz eredetileg célnyelven keletkezett, valamint azonos célnyelvű, de fordítás eredményeképp keletkezett szövegeket tartalmaz. Ideális esetben az egyetlen változó a korpuszban a szöveg keletkezésének körülménye, azaz hogy fordítás vagy nem fordítás. Ezen túl minden egyéb paraméter mindkét korpuszkomponensben azonos (szövegfajta, téma, hossz, keletkezés ideje és így tovább). Az összehasonlítható korpuszok betekintést nyújtanak a fordítás eredményeképpen létrejött szöveg természetébe, illetve biztosítják a terepet a fordítási univerzálék kutatásához. Ezen oknál fogva elsősorban a fordításkutatás számára fontosak ezek a korpuszok, ugyanakkor a vizsgálati eredmények a fordításoktatásban is felhasználhatók az eredeti és a fordítás eredményeképpen létrejött célnyelvi szövegek közötti különbségek tudatosítására.

Az ideális korpusz minden fordításkutató számára a kétirányú (bidirekcionális) korpusz. Az *English Norwegian Parallel Corpus* (Johansson 1998) mintájára több skandináv ország elkészítette saját kétirányú korpuszát (vö. Aijmer et al. 1996, Aijmer és Altenberg 2000, Mauranen 2000), amely magában foglalja az előzőekben bemutatott korpuszfajtákat. A korpuszmodulok összepárosításától függően a korpuszt lehet fordítási korpuszként két irányban, összehasonlítható korpuszként, párhuzamos korpuszként, és fordított szövegek párhuzamos korpuszaként használni. Könnyen belátható, milyen széles körű egy ilyen korpusz felhasználhatósága, és mennyi lehetőséget rejt magában fordító és kutató számára egyaránt. Egy bidirekcionális korpusz létrehozásához azonban szerzői jogi kérdéseken túl egyéb akadályokat is le kell küzdeni. Az összehasonlíthatóság érdekében az egyes modulok méretének és tartalmának meg kell egyeznie, ami nyelvpároktól függően könnyebben vagy nehezebben valósítható meg. Hogy mit és mennyit fordítanak egy adott korban és egy adott kultúrában függ a forrásnyelv és a forrásnyelvi kultúra célnyelvi kultúrában elfoglalt helyétől, státuszától. Nyilvánvaló aránytalanságok vannak „nagy” és „kis” nyelvek között, ugyanis lényegesen kevesebb fordítás áramlik „nagy” nyelvek felé, mint fordítva, ami korlátozza a merítést. Mégis, amit kutatóként, fordítóként nyerünk messze felülírja a korpuszkészítés nehézségeit, ezért minél több nyelvpár viszonylatában ajánlatos kétirányú korpuszokat létrehozni.

2. ábra

Az ENPC szerkezete (Johansson 1998: 8)



4. Amit korpuszokból lehet megtudni a fordításról

A korpuszok megjelenésével kutathatóvá vált számos téma közül a fordítási univerzálékra irányult a legnagyobb figyelem (vö. Hansen 2000, Jantunen 2000, Laviosa-Braithwaite 1996, Laviosa 1997; 1998, Pápai 2001; 2002, Tirkkonen-Condit 2000, Zanettin 2000). Már korábban is felmerült, hogy a fordítás eredményeképpen keletkezett szövegeknek olyan jegyeik vannak, amelyek semmilyen egyéb körülmények között keletkezett szövegekre nem jellemzők (vö. Blum-Kulka 1986, Toury 1995). Ahhoz azonban, hogy a kutatások szintet léphessenek, fordítás eredményeképpen keletkezett szövegek nagy mennyiségére volt szükség, továbbá korpusznyelvészeti módszerekre és eszközökre az adatok vizsgálatához (vö. Kohn 1996). A fordított szövegek jellegzetességeinek korpuszok segítségével történő vizsgálatát Baker (1993; 1996) vetette fel elsőként. A fordított szövegekben nyel-

vpároktól, fordítási iránytól, szövegtípustól, a fordítási folyamat egyéb körülményeitől függetlenül megjelenő jegyeket fordítási univerzáléknak nevezi. A tudományos diskurzusban részt vevő kutatók egy része tényként fogadja el a fordítási univerzálék létezését, és különböző hipotéziseket tesztelnek (például Laviosa 2002, Olohan 2004, Zanettin 2012), kutatásmódszertani kérdéseket vetnek fel (Bernardini és Zanettin 2004, Zanettin 2012), konceptuális kérdéseket tisztáznak (Chesterman 2004), vagy kognitív nyelvészeti alapon magyarázzák működésüket (Halverson (2003; 2010). Ugyanakkor kritikus hangok is megjelentek, amelyek például a fogalom tartalmi homályosságára mutatnak rá (Malmkjaer 2008), létezésüket megkérdőjelezik (vö. Tymoczko 1998), vagy határozottan visszautasítják (House 2008, Becher 2010).

Annak ellenére, hogy sok kérdőjel merül fel a fordítási univerzálék körül, és vita tárgyát képezi az is, hogy pontosan mi minősül univerzálénak (vö. Tirkkonen-Condit 2002; 2004), a kutatások többsége számára Baker (1996) osztályozása jelenti a kiindulópontot, az univerzálék négy csoportját különítve el: az explicitáció, a szimplifikáció, a normalizáció (más néven konzervativizmus vagy standardizáció) és a kiegyensúlyozás (levelling out).

Tekintettel arra, hogy a tanulmánygyűjtemény egyik írása részletesen foglalkozik a fordítási univerzálékkal, a következőkben csak néhány észrevétel erejéig térek ki a témára. Csakúgy, mint az univerzálék kérdésénél általában, az egyes fordításiuniverzálé-jelöltek körül sem kevés az ellentmondás, homályosság és átfedés. Az explicitációval, például, jelentős számú kutatás foglalkozik a korpuszok megjelenése előtt és után (vö. Baker 1995; 1996, Olohan és Baker 2000, Blum-Kulka 1986, Heltai 2003, Klaudy 1993; 1996; 1998; 1999, Laviosa-Braithwaite 1996, Laviosa 1998, Pápai 2001; 2002, Séguinot 1988, Weissbrod 1992), elkészült több tipológia (Klaudy 1999; 2008, Englund Dimitrova 2005, Robin 2015; 2018), a fogalom meghatározása mégis ingadozik jelenség, tendencia vagy művelet között, és máig is folyik annak tisztázása, mi tekinthető explicitációnak (vö. Balaskó 2004), de az sem egyértelmű, hogy fordítási univerzálénak számít-e egyáltalán az explicitáció (Becher 2010). Ugyanezek a kérdések merülnek fel a szimplifikáció és a normalizáció, vagy más néven standardizáció kapcsán. A fordított szöveg nyelvezetének és/

vagy üzenetének egyszerűsítésére irányuló vizsgálatok eredményei például nem koherensek: egy részük igazolja (Laviosa-Braithwaite 1996, Laviosa 1997; 1998), más részük cáfolja a szimplifikáció jelenlétét a fordított szövegekben (Mauranen 2000, Jantunen 2001). Hasonlóan vitatható eredmények születtek a normalizációvizsgálatokban (vö. Kenny 2000, Mauranen 2008, Teich 2001, Hansen 2003), amelyek azt próbálták megállapítani, hogy a fordítások valóban a célnyelv tipikus nyelvhasználati mintáihoz igazodnak, illetve hogy a fordítók valóban kihangsúlyozzák-e a célnyelv jegyeit. A legkevesebb kérdőjel a kiegyensúlyozás körül merül fel, de ez elképzelhetően annak köszönhető, hogy mindezidáig erre az univerzáléjelöltre irányult a legkevesebb figyelem. A vizsgálatok azt a hipotézist próbálják alátámasztani, hogy a fordított szövegek semlegesebbek, homogénebbek, kevésbé egyediek, mint az eredeti szövegek, ezért egy fordítási korpusz egyes szövegei között kisebb a variáció, mint egy eredeti szövegekből álló korpusz szövegei között.

Az univerzálékutatás továbbra is izgalmas területe a fordítástudománynak, hiszen magában hordozza a törvényszerűségek felfedezésének ígéretét. Még ha nem is azokra a kérdésekre kaptunk válaszokat, amelyeket feltettünk, a kapott válaszok rendkívüli mértékben gazdagítják tudásunkat a fordított szöveg természetéről. Az időszakos elakadások ellenére folytatni kell a kutatásokat, de közben ismétlődően fel kell tenni a kérdést, hogy megfelelő módon és megfelelő célra használjuk-e a korpuszokat.

5. Nyelvi mintázatok a fordításban

Míg a Sinclair által képviselt korpusznyelvészeti irányzat hagyományosan az alulról felfelé módszert alkalmazva jut el a nyelvi adattól az általánosításig, az alapjain kialakult korpuszalapú fordításkutatás áttért a felülről lefelé módszerre, hogy bizonyítékokkal támassza alá a fordítási univerzálékhoz kapcsolódó hipotéziseket, mint láttuk több-kevesebb sikerrel. Ha nem is kell teljesen elvetni ezt a fajta megközelítést, feltétlenül szükségesnek tartjuk kiegészíteni az alulról felfelé módszert alkalmazó vizsgálatokkal, amelyek nyelvi tényekből kiindulva segíthetnek azonosítani azokat a jegyeket, amelyek megkülönbözte-

tik a fordítás eredményeképp keletkezett célnyelvi szövegeket az eredetileg célnyelven keletkezett szövegektől.

Mint ahogyan arra korábban rámutattunk, a nyelv nagyrészt panelekből épül fel, és jellegzetes mintázatokba rendeződik. Ha a nyelvi transzfertevékenység nem hagyna nyomott a fordított szövegen, az azonos paraméterekkel (például azonos műfaj, téma, keletkezési idő) rendelkező szövegeknek azonos panelekből kellene építkezniük, függetlenül attól, hogy fordítások vagy nem-fordítások. Az összehasonlítható korpuszokkal végzett kutatások azonban azt mutatják, hogy keletkezésük sajátos körülményei miatt a fordított szövegek a szokásostól eltérő mintázatokba rendeződnek.

Egy igei kulcsszó konkordanciavizsgálatánál például eltérő mintázatok rajzolódtak ki a paradigmatikus és a szintagmatikus tengelyen egyaránt. Az ige paradigmájának más-más alakjai fordultak elő fordított és nem-fordított szövegekben, és rendeződtek egymástól teljesen különböző mintázatokba. A fordított szövegektől eltérően az eredeti célnyelvi szövegekben kevesebb és lazább mintázatok körvonalazódtak (Balaskó 2012). Hasonló eredményre jutott Baroni és Bernardini (2003), akik sokkal erősebb, stabilabb kollokációs kapcsolatokat találtak fordított szövegekben, mint nem-fordított szövegekben, ahol a kollokációs kapcsolatok nagyon gyengék voltak. Egyéb kollokáció-vizsgálatokból az derült ki, hogy fordított szövegekben gyakran fordulnak elő olyan szókapcsolatok, amelyek léteznek és elfogadottak ugyan a célnyelvben, eredeti célnyelvi szövegekben mégsem, vagy csak ritkán fordulnak elő (Mauranen 2008), Jantunen (2001) pedig teljesen eltérő kollokációs mintázatokat talált fordított, illetve nem-fordított szövegekben. A nyelvi mintázatokról szóló részben bemutatott konkordancia-elemzés módszerét alkalmazva, ugyancsak különböző mintázatokra derült fény egy főnévi kulcsszó vizsgálatánál. Nemcsak a mintázatok voltak különbözőek, hanem a teljes paradigma más-más alakjai fordultak elő fordított és nem-fordított szövegekben (Balaskó 2008). A nyelvi mintázatok vizsgálata messze túlmutat a mikroelemzések korlátain, és felveti többek között a fordítás egységének kérdését. A nyelvi mintázatokra irányuló vizsgálatok azt mutatják, hogy az erős forrásnyelvi mintázatok ugyanolyan erős célnyelvi mintázatokká állnak össze, amelyek a *lehetséges*, ámbár *nem megszokott* keretei között helyezkednek el, ha azok nem sértik

a célnyelvi használat normáit. A dominóhatásnak köszönhetően pedig az eltérő mintázatokba rendeződő fordított szöveg az eredeti célnyelvi szövegtől különböző szöveggé állnak össze, melyet a kohézió szintjén eltolódások jellemeznek (vö. Blum-Kulka 1986, Seidl-Péch 2013, Károly 2017).

6. Összegzés

Az 1980-as évek közepéig a korpusznyelvészet nem volt több néhány megszállott kutató kísérletezésénél a lehetetlennel, majd fejlődött fokozatosan önálló nyelvtudományi diszciplínává és egyúttal vizsgálati módszerré. A John Sinclair nevéhez fűződő korpusznyelvészeti irányzat módszerére alapozva vált a korpuszalapú fordításkutatás a fordításstudomány egyik marginális témájából új paradigmává, amely rövid időn belül bekerült a leíró fordítástudomány fősodrába. A fordított szöveg természetének feltárására irányuló univerzálékutatást jelentős mennyiségű megfigyelési és kísérleti adat támogatja, ezek azonban olykor ellenbizonyítékokat is szolgáltatnak, megkérdőjelezve az univerzálék meghatározásának érvényességét és megfelelő kategóriákba való sorolását. Ezért az alkalmazott fentről lefelé módszert ki kell egészíteni a nyelvet alkotó panelek mintázatainak alulról felfelé történő, valamint a mintázatokba rendeződés teljes hatásmechanizmusának vizsgálatával.

Irodalom

Aijmer, K., Altenberg, B., Johansson, M. 1996. Text-based contrastive studies in English. Presentation of a project. In: Aijmer, K., Altenberg, B., Johansson, M. (eds) *Languages in Contrast. Papers from a Symposium on Text-based Cross-linguistic studies. Lund 4-5 March 1994.* Lund: Lund University Press. 73–85.

Aijmer, K., Altenberg, B. 2000. The English-Swedish Parallel Corpus: A resource for contrastive research and translation studies. In: Mair, Ch., Hundt, M. (eds) *Corpus Linguistics and Linguistic Theory*. Papers from the Twentieth International Conference on English Language Research on Computerized Corpora. (ICAME 20) Freiburg im Breisgau 1999. Amsterdam-Atlanta GA: Rodopi. 15–33.

Altenberg, B., Olofsson, M. E. 1990. Phraseology in spoken and written English: Presentation of a project. In: Aarts, J., Meijs, W. (eds) *Theory and practice in corpus linguistics*. Amsterdam-Atlanta GA: Rodopi. 1–26.

Baker, M. 1993. Corpus Linguistics and Translation Studies. Implications and Applications. In: Baker, M., Francis, G., Tognin-Bonelli, E. (eds) *Text and Technology: In Honour of John Sinclair*. Amsterdam: John Benjamins. 233–250. https://doi.org/10.1075/z.64.15bak

Baker, M., Francis, G., Tognin-Bonelli, E. (eds) 1993. *Text and Technology: In Honour of John Sinclair*. Amsterdam: John Benjamins. https://doi.org/10.1075/z.64

Baker, M. 1995. Corpora in Translation Studies. An Overview and Suggestions for the Future Research. *Target* Vol. 7. No. 2. 223–245. https://doi.org/10.1075/target.7.2.03bak

Baker, M. 1996. Corpus-based translation studies: The challenges that lie ahead. In: Somers, H. (ed) *Terminology, LSP and Translation. Studies in language engineering in honour of Juan C. Sager*. Amsterdam/Philadelphia: John Benjamins. 175–186. https://doi.org/10.1075/btl.18.17bak

Baker, M. 2005. Linguistic models and methods in the study of translation. In: Kittel, H., Frank, A. P., Greiner, N., Hermans, T., Koller, W., Lambert, J., Paul, F. (eds) Übersetzung. *Translation*. *Traduction*. Berlin, New York: Walter de Gruyter. 285–294.

Balaskó M. 2004. Korpusznyelvészeti vizsgálatok és fordításnyelvi minták. Doktori disszertáció. Pécsi Tudományegyetem.

Balaskó, M. 2008. What Does the Figure Show? Patterns of Translationese in a Hungarian Comparable Corpus. *trans-kom* Vol. 1. No. 1. 58–73.

Balaskó, M. (2012) „Elective affinities" – or Do words keep different company i translation? In: Karabalić, V., Aleksa Varga, M., Pon, L. (eds) *Discourse and Dialogue*. Frankfurt am Main: Peter Lang. 155–168.

Baroni, M., Bernardini, S. 2003. A preliminary analysis of collocational differences in monolingual comparable corpora. In: Archer, D., Rayson, P., Wilson, A., McEnery, T. (eds) *Proceedings of the Corpus Linguistics 2003 Conference*. Technical Paper number 16, Special issue. Lancaster: UCREL, Lancaster University. 82–91.

Becher, V. 2010. Abandoning the notion of 'translation-inherent' explicitation: against a dogma of translation studies. *Across Languages and Cultures.* Vol. 11. No. 1. 1–28. https://doi.org/10.1556/Acr.11.2010.1.1

Bernardini, S., Zanettin, F. 2004. When is a universal not a universal? Some limits of current corpus-based methodologie for the investigation of tranlation universals. In: Mauranen, A., Kujamäki, P. (eds) *Translation Universals – Do they exist?* Amsterdam/Philadelphia: John Benjamins. 51–62. https://doi.org/10.1075/btl.48.05ber

Blum-Kulka, S. 1986. Shifts of Cohesion and Coherence in Translation. In: Blum-Kulka, S., House, J. (eds) *Interlingual and Intercultural Communication: Discourse and Cognition in Translation and Second Language Acquisition Studies*. Tübingen: Narr. 17–35.

Chesterman, A. 2004. Beyond the particular. In: Mauranen, A., Kuyamaki, P. (eds) *Translation universals: Do they exist?* Amsterdam/Philadelphia: John Benjamins. 33–49. https://doi.org/10.1075/btl.48.04che

Cowie, A. P. 1988. Stable and creative aspects of vocabulary use. In: Carter, R., McCarthy, M. (eds): *Vocabulary and Language Teaching*. London: Longman. 126–139.

Dimitrova, E. B. 2005. *Expertise and Explicitation in the Translation process*. Amsterdam: Benjamins. https://doi.org/10.1075/btl.64

Firth, J. R. 1935. The Technique of Semantics. *Transactions of the Philological Society*. 36–72. https://doi.org/10.1111/j.1467-968X.1935.tb01254.x

Firth, J. R. 1957a A Synopsis of Linguistic Theory, 1930-1955. *Studies in Linguistic Analysis*. Special Volume of the Philological Society. 1–32.

Firth, J. R. 1957b *Papers in Linguistics*. London: Oxford University Press.

Halverson, S. 2003. The cognitive basis of translation universals. *Target* Vol. 15. No. 2. 197–241. https://doi.org/10.1075/target.15.2.02hal

Halverson, S. 2010. Cognitive translation studies: developments in theory and method. In: Shreve, G. M., Angelone, E. (eds) *Translation and Cognition*. Amsterdam/Philadelphia: John Benjamins. 349–369. https://doi.org/10.1075/ata.xv.18hal

Hanks, P. 2013. *Lexical analysis: norms and exploitations*. Cambridge, Mass.: The MIT Press. https://doi.org/10.7551/mitpress/9780262018579.001.0001

Hansen, S. 2000. A Contrastive Analysis of Parallel and Comparable Corpora. Paper presented at the UMIST/UCL Research Models in Translation Studies Conference, Manchester, 28–30 April, 2000.

Hansen, S. 2003. *The Nature of Translated Text: an Interdisciplinary Methodology for the Inspection of the Specific Properties of Translation*. Saarbruecken Dissertations. Computational Linguistics and Language Technology. Vol. 13. Saarbruecken: University of Saarbruecken.

Heltai P. 2003. Az explicitáció egyes kérdései az angol–magyar szakfordításban. In: Feketéné Silye M. (szerk.) *Porta Lingua: Szaknyelvoktatásunk az EU kapujában*. Debrecen: DE ATC. 173–198.

Hoey, M. 1991. *Patterns of Lexis in Text*. Oxford: Oxford University Press.

Hoey, M. 2005. *Lexical Priming*. London/New York: Routledge.

House, J. 2008. Beyond Intervention: Universals in Translation? *trans-kom* Vol. 1. No. 1. 6–19.

Hunston, S., Francis, G. 2000. *Pattern Grammar. A corpus-driven approach to the lexical grammar of English*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/scl.4

Jantunen, J. 2000. What Can Corpora Tell us about Translated Language? A Comparable Corpus of Finnish in Use for Making Hypotheses. Paper presented at the UMIST/ UCL Research Models in Translation Studies Conference, Manchester, 28–30 April, 2000.

Jantunen, J. 2001. Synonymity and lexical simplifi cation in translations: A corpus-based approach. *Across Languages and Cultures,* Vol. 2. No. 1. 97–112.

Johansson, S. 1998. On the role of corpora in cross-linguistic research. In: Johansson, S., Oksefjell, S. (eds) *Corpora and Cross-linguistic Research*. Amsterdam-Atlanta GA: Rodopi. 3–24.

Károly, K. 2017. *Aspects of coherence and cohesion in translation*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/btl.134

Kenny, D. 2000. Translators at play: Exploitations of collocational norms in German–English translation. In: Dodd, B. (ed.) *Working with German corpora.* Birmingham: Uniersity of Birmingham Press. 143–160.

Kjellmer, G. 1994. *A dictionary of English collocations*. Oxford: Clarendon Press.

Klaudy, K. 1993. On explicitation hypothesis. In: Klaudy, K., Kohn, J. (eds) *Transferre necesse est… Current issues of translation theory*. Szombathely: BDTF. 69–77.

Klaudy, K. 1996. Back-Translation as a Tool for Detecting Explicitation Strategies in Translation. In: Klaudy, K., Lambert, J., Sohár, A. (eds) *Translation Studies in Hungary*. Budapest: Scholastica. 99–114.

Klaudy, K. 1998. Explicitation. In: Baker, M. (ed.) *Routledge Encyclopedia of Translation Studies*. London: Routledge. 80–84.

Klaudy K. (1999) Az explicitációs hipotézisről. *Fordítástudomány* Vol. 1. No. 2. 5– 22.

Klaudy, K. (2008) Explicitation. In: Baker, M., Saldanha, G. (eds) *Routledge Encyclopedia of Translation Studies*. London: Routledge. 80–85.

Kohn, J. 1996. What Can (Corpus) Linguistics Do for Translation? In: Klaudy, K., Lambert, J., Sohár, A. (eds) *Translation Studies in Hungary*. Budapest: Scholastica. 39–52.

Krashen, S. D. 1981. *Second Language Acquisition and Second Language Learning*. Oxford: Pergamon Press.

Laviosa-Braithwaite, S. 1996. The English Comparable Corpus (ECC): A Resource and a Methodology for the Empirical Study of Translation. PhD Thesis. Manchester: Centre for Translation Studies, UMIST.

Laviosa, S. 1997. How Comparable Can 'Comparable Corpora' Be? *Target* Vol. 9. No. 2. 289–319. https://doi.org/10.1075/target.9.2.05lav

Laviosa, S. 1998. Core Patterns of Lexical Use in a Comparable Corpus of English Narrative Prose. *Meta* Vol 43. No. 4. 557–570. https://doi.org/10.7202/003425ar

Laviosa, S. 2002. *Corpus-based Translation Studies: Theory, Findings, Applications*. Amsterdam-Atlanta GA: Rodopi.

Mauranen, A. 2000. Strange strings in translated language: A study on corpora. In: Olohan, M. (ed), *Intercultural faultlines: Research models in translation studies 1: Textual and cognitive aspects*. Manchester: St. Jerome Publishing. 119–141.

Mauranen, A. 2008. Universal Tendencies of Translation. In: Anderman, G., Rogers, M. (eds) *Incorporating Corpora: The Linguist and the Translator*. Clevedon: Multilingual Matters. 32–48.

Malmkjaer, K. 2008. Norms and nature in translation studies. In: Anderman, G., Rogers, M. (eds) *Incorporating Corpora. The Linguist and the Translator.* Clevedon: Multilingual Matters. 49–59. https://doi.org/10.21832/9781853599873-007

Melčuk, I. 1988. Semantic description of lexical units in an Explanatory Combinatorial Dictionary: basic principles and heuristic criteria. *International Journal of Lexicography*. Vol. 1. No. 3. 165–188.

Melčuk, I. 1995. Phrasemes in language and phraseology in linguistics. In: Everaert, M., van der Linden, E-J., Schenk, A., Schreuder, R. (eds) *Idioms: structural and psychological perspectives*. Hillsdale, N. J.: Lawrence Erlbaum Associates. 167–232.

Moon, R. 1998 *Fixed expressions and idioms in English*. Oxford: Clarendon Press.

Nattinger, J., DeCarrico, J. 1989. Lexical phrases, speech acts and teaching conversation. In: Nation, P., Carter, R. (eds): *AILA Review 6: Vocabulary Acquisition*. Amsterdam: AILA. 118–139.

Nattinger, J., DeCarrico, J. 1992. *Lexical Phrases and Language Teaching*. Oxford: OUP.

Olohan, M., Baker, M. 2000. Reporting *that* in Translated English. Evidence for Subconscious Processes of Explicitation? *Across Languages and Cultures* Vol. 3. No. 2. 141–158. https://doi.org/10.1556/Acr.1.2000.2.1

Olohan, M. 2004. *Introducing Corpora in Translation Studies*. London/New York: Routledge. https://doi.org/10.4324/9780203640005

Pápai V. 2001. *Az explicitációs hipotézis vizsgálata*. Doktori disszertáció. Pécsi Tudományegyetem.

Pápai V. 2002. Fordítási univerzálék: az explicitáció. In: Fóris Á., Kárpáti E., Szűcs T. (szerk.) *A nyelv nevelő szerepe. A XI. Magyar Alkalmazott Nyelvészeti Kongresszus előadásainak válogatott gyűjteménye*. Pécs: Lingua Franca Csoport. 486–493.

Robin E. 2015. *Fordítási univerzálék a lektorált fordításokban*. Doktori értekezés. Budapest: ELTE.

Robin E. 2018. *Fordítási univerzálék és lektorálás*. Budapest: Eötvös József Kiadó.

Scott, M., Johns, T. 1993. *MicroConcord.* Oxford: Oxford University Press.

Séguinot, C. 1988. Pragmatics and the Explicitation Hypothesis. *TTR: Traduction, Terminologie, Rédaction* Vol. 1. No. 2. 106–114. https://doi.org/10.7202/037024ar

Seidl-Péch O. 2013. Célnyelvi szövegek nyelvtechnológiai eszközökkel támogatott lexikai kohéziós vizsgálata. In: Klaudy K. (szerk.) *Fordítás és tolmácsolás a harmadik évezred elején*. ELTE Eötvös Kiadó. 95–106.

Seidl-Péch O. 2018. Melyek a (szak)fordító és a fordításkutató munkáját segítő legfontosabb nyelvi korpuszok? In: Robin, E., Zachar, V. (szerk.) *Fordítástudomány ma és holnap.* Budapest: L'Harmattan Kiadó. 175–191.

Sinclair, J. 1987. *Collins Cobuild English Dictionary*. London: Collins.

Sinclair, J. 1990. *Collins Cobuild English Grammar*. London: Collins.

Sinclair, J. M. 1991a. *Corpus Concordance Collocation*. Oxford: Oxford University Press.

Sinclair, J. M. 1991b. *Council of Europe Multilingual Lexicography Project*. [Report submitted to the Council of Europe under contract no. 57/89.]

Stubbs, M. 1993. British Traditions in Text Analysis – from Firth to Sinclair. In: Baker, M., Francis, G., Tognin-Bonelli, E. (eds) *Text and Technology: In Honour of John Sinclair*. Amsterdam/Philadelphia: John Benjamins. 1–33. https://doi.org/10.1075/z.64.02stu

Stubbs, M. 1996. *Text and Corpus Analysis*. Oxford: Blackwell.

Stubbs, M. 2001. *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford: Blackwell.

Teich, E. 2001. Towards a model for the description of cross-linguistic divergence and commonality in translation. In: Steiner, E., Yallop, C. *Exploring translation and multilinual text production: Beyond content.* Berlin: Mouton de Gruyter. 191–227.

Tirkkonen-Condit, S. 2000. In Search of Translation Universals: Non-equivalence or 'Unique' Items in a Corpus Test. Paper presented at the UMIST/UCL Research Models in Translation Studies Conference, Manchester, 28-30 April, 2000.

Tirkkonen-Condit, S. 2002. Translationese – a Myth or an Empirical Fact? *Target* Vol. 14. No. 2. 207–220. https://doi.org/10.1075/target.14.2.02tir

Tirkkonen-Condit, S. 2004. Unique items – over- or under-represented in translated language? In: Mauranen, A., Kujamäki, P. (eds) *Translation Universals – Do they exist?* Amsterdam/Philadelphia: John Benjamins. 177–184. https://doi.org/10.1075/btl.48.14tir

Tognini-Bonelli, E. 2001. *Corpus Linguistics at Work*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/scl.6

Tognini-Bonelli, E., Del Lungo Camiciotti, G. (eds) 2005. *Strategies in Academic Discourse*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/scl.19

Toury, G. 1995. *Descriptive Translation Studies and Beyond*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/btl.4

Tymoczko, M. 1998 Computerized Corpora and the Future of Translation Studies. *Meta* Vol. 43. No. 4. 652–659. https://doi.org/10.7202/004515ar

Weissbrod, R. 1992. Explicitation in translations of prose-fiction from English to Hebrew as a function of norms. *Multilingua* Vol. 11. No. 2. 153–171. https://doi.org/10.1075/ata.v.22wei

Zanettin, F. 2000. Parallel Corpora in Translation Studies. Issues in Corpus Design and Analysis. In: Olohan, M. (ed.) *Intercultural Faultlines. Research Models in Translation Studies* I. *Textual and Cognitive Aspects*. Manchester, UK, Northampton MA.: St. Jerome. 105–118.

Zanettin, F. 2012. *Translation-Driven Corpora. Corpus Resources for Descriptive and Applied Translation Studies*. Manchester: St Jerome. https://doi.org/10.1016/j.sbspro.2013.10.618

Korpuszalapú fordításkutatás: lehetőségek és nehézségek
Fókuszban a korpuszépítés és a korpuszalapú elemzés

Seidl-Péch Olívia

olivia@inyk.bme.hu

Budapesti Műszaki és Gazdaságtudományi Egyetem

Idegen Nyelvi Központ

Kivonat: A fordítástudomány számára releváns kutatási módszerek között már az ezredfordulót megelőzően is hangsúlyosak voltak a korpuszalapú kutatások, amelyek szerepe napjainkra tovább növekedett. A fordítástudomány alkalmazott nyelvészeti beágyazottsága már a kezdetektől kezdve szükségessé tette a módszertani kérdések tisztázását (vö. Károly 2002), melyhez hasonlóan napjainkban a tudományágon belül a korpuszalapú kutatások előtérbe kerülését eredményező kutatási paradigmaváltás és az azt lehetővé tevő technológiai fejlődés miatt válik szükségessé a korpuszalapú módszer jellemzőinek és kritériumainak áttekintése. A hatékony és lekérdezési eredményeit tekintve megbízható kutatások alapvető tulajdonságai között kerülnek előtérbe a kutatási kérdéseknek megfelelő, valamint a vizsgált sokaságot kiegyensúlyozottan reprezentáló minta, illetve a minta adekvát rendezését megvalósító korpuszdesign. A korpuszalapú kutatások elméleti kereteit biztosító módszertani alapelvek (vö. Laviosa 1998a) mellett tehát a mintavételi kritériumok tisztázása válik elsődlegessé, amelynek jelentőségét növeli a könnyen hozzáférhető, nagy adatmennyiségek megjelenése. A kutatások számára tehát hangsúlyozottan elsődleges (a valóban óriási kínálatból) a reprezentatív minta összeállításának szükségessége. A mintavételi szempontrendszer összeállításához, illetve a kutatások tervezéséhez kínál egyfajta összefoglalást a jelen tanulmány.

Kulcsszavak: korpuszépítés, korpuszalapú fordításkutatás, kvantitatív kutatás, korpusztervezés, mintavétel

Seidl-Péch Olívia: Korpuszalapú fordításkutatás: lehetőségek és nehézségek. Fókuszban a korpuszépítés és a korpuszalapú elemzés. In: Robin E., Seidl-Péch O. (szerk.) 2020. *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordításkutatás Magyarországon.* Segédkönyvek a nyelvi közvetítésről I. Budapest: ELTE BTK Fordítástudományi Doktori Program, MANYE Fordítástudományi Szakosztály. DOI: https://doi.org/10.36252/Nyelvikozvsegedkonyv1.4

1. Bevezetés

A korpuszalapú kutatások megjelenése óta eltelt évtizedek tapasztalatai alapján ma már bátran megállapíthatjuk, hogy a szakterület méltán foglal el kiemelkedően fontos helyet a nyelvészeti kutatások között (McCarthy és O'Keefee 2010). A korpuszok legfontosabb ismérve, miszerint a korpuszban összegyűjtött szövegek a természetes nyelvhasználat gépileg olvasható formában tárolt mintái (vö. Bowker és Pearson 2002: 9), meghatározó jelentőséggel bír az empirikus nyelvészeti kutatások számára.

A kutatók a kutatási célokat szolgáló tudatos gyűjtés és mintavétel során az adott nyelvre vagy nyelvváltozatra jellemző szövegeket rendezik korpuszokba és azok alkorpuszaiba, mindig ügyelve arra, hogy a vizsgálatokat és a lekérdezéseket a lehető legkiegyensúlyozottabb korpuszok segítségével végezzék. Az mára már nyilvánvaló, hogy – a nagyon kevés és csak bizonyos kutatási témákra jellemző kivételtől eltekintve (például egy szerző munkásságának vizsgálata) – a számítástechnikai eszközök megnövekedett kapacitása ellenére sem lehetséges a korpuszban az adott nyelvre vagy nyelvváltozatra jellemző összes nyelvi minta rögzítése. Ennek következtében egyre inkább előtérbe kerül a korpuszok kiegyensúlyozottságának és reprezentativitásának kérdése (Seidl-Péch 2018), mivel csak ezek figyelembevételével zárható ki a kutatások eredményeinek véletlenszerűsége. Másképpen fogalmazva, csak a körültekintően meghatározott kritériumok alapján válogatott mintán végzett korpuszalapú kutatások eredményei tekinthetők tudományosan megalapozottnak. Ellenkező esetben a lekérdezések validitása megkérdőjelezhető akkor is, ha a kutatás esetleg egy több millió szövegszavas korpusz lekérdezésével valósul meg. A korpuszok kiegyensúlyozottsága és reprezentativitása tehát elengedhetetlen feltétele a kutatási eredmények érvényességének és megbízhatóságának. Megállapíthatjuk továbbá, hogy a nyelvi korpusz kiegyensúlyozottsága és reprezentativitása teszi a korpuszt alkalmassá arra, hogy segítségével információkat gyűjtsünk egy-egy kifejezés, illetve szókapcsolat forrás- vagy célnyelvi környezetben történő műfajra és/vagy regiszterre jellemző használatáról, vagy tágabban értelmezve az adott nyelv térben, időben és modalitásban behatárolt változatáról, amely megmutatja a korpusz összetételének megfelelő, műfajra,

szövegtípusra, regiszterre, illetve az autentikus és/vagy célnyelvi nyelvhasználatra jellemző használati mintákat (vö. Seidl-Péch 2018).

2. A korpuszalapú módszer jelentőségének növekedése a fordításkutatásban

Már az 1950-60-as években, a nyelvészeti fordítástudomány megjelenésével (vö. Klaudy 2004) megkezdődött a fordítási folyamat **összes nyelvi és nem nyelvi** változójának függvényében a fordítás eredményének, folyamatának és funkciójának leírása (vö. Holmes 1975). Az irodalmi paradigmával való szakítás következtében megerősödött a fordítástudomány nyelvészeti beágyazottsága, ezzel együtt pedig megjelentek a társadalomtudományi kutatásokat jellemző módszerek és eredmények. Az alkalmazott nyelvészeti kutatások sorába illeszkedő leíró fordítástudományi vizsgálatok között a korpuszalapú paradigma kezdetekben a fordított és a nem fordított (autentikus), illetve a forrásnyelvi és a célnyelvi szövegek nyelvi jellemzőinek kontrasztív leírását részesítette előnyben, mely kutatások kedveztek a fordítási univerzálék megfogalmazásának és a nyelvpárspecifikus fordítói magatartás vizsgálatának (vö. Károly 2003). Klaudy korszakolása alapján (Klaudy 2004, 2005) az 1970-80-as éveket a nyelvészeten belüli (például szövegnyelvészet, pragmatika, számítógépes nyelvészet) és kívüli (például pszichológia, szociológia, irodalomtudomány, filozófia) határtudományokkal való összefogás, míg az 1990-es éveket az európai integráció következtében a fordítások iránt megnövekedett társadalmi igény megjelenése és felerősödése (például az EU nyelvpolitikája, multinacionális vállalatok igényei, lokalizáció) jellemezte.

Az ezredfordulótól kezdve a fordítandó dokumentumok számának robbanásszerű növekedése szükségessé tette a nyelvtechnológiai alkalmazások integrálását a fordítási folyamatba (például gépi adatbázisok, párhuzamos korpuszok, a fordító számítógépes segédeszközei). Napjaink fordítója munkája során lépten nyomon találkozik korpuszokkal (például online korpuszok, gépi fordítás), illetve maga is hoz létre speciális korpuszokat (például terminológiai adatbázis, fordítómemória, párhuzamos dokumentumok). Mindezek hozzájárultak ahhoz, hogy a 2000-es évektől kezdve egy olyan **újabb paradigmavál-**

tás következzen be a fordításkutatásban, amelyet az empirikus alapokon nyugvó korpuszalapú tudományos megközelítés jellemez (Seidl-Péch 2019). A korpuszalapú kutatások előtérbe kerülését a nagy adatállományok tárolásának és kezelésének egyszerűbbé válása tette lehetővé. Az eredményközpontú kutatások fókusza így már igen széles skálán mozoghat a célnyelvi szövegtulajdonságok leírásától a fordítói stílus vizsgálatáig, illetve a fordítóképzés tapasztalataitól a tolmácsoláskutatásig. A korszakra ugyanakkor egyre inkább jellemző a költségigényesebb folyamatközpontú kutatások számának növekedése is (például tolmácsolásnál és fordításnál egér- és/vagy szemmozgáskövető vizsgálat, tolmácsolásnál levegővétel/szünet- vagy elakadásvizsgálat), amelyekre a technológiailag teljesen megújult tolmácsszakmának (vö. Fantinuoli 2018.) óriási szüksége is van. A korpuszalapú kutatások kvantitatív eredményeit ugyanakkor a kutatók mindinkább szükségesnek érzik kvalitatív vizsgálatokkal kiegészíteni (lásd például Robin 2018), amely gyakorlatot egyrészt a korpuszok kiegyensúlyozottságának és reprezentativitásának megkérdőjelezhetősége teszi szükségessé, másrészt annak felismerése, hogy a kvantitatív eredmények sok esetben csak az intuíció megerősítésére elegendőek, de nem adnak valós válaszokat a felmerülő kutatási kérdésekre (Seidl-Péch és Robin 2019).

Az új módszertan a kutatási témakínálat bővülését is elősegítette. Egyes kutatások a fordítási normát kísérelik meg körüljárni, és a fordítási mintázatot (például Balaskó 2005) írják le, továbbá előtérbe került a fordítások szociokulturális jellemzőinek nyelvészeti és interkulturális beágyazottságának vizsgálata is. A kutatások egy másik csoportja az új technológiák mentén végez vizsgálatokat (például audiovizuális fordítás: Bańos et al. 2013; közösségi fordítás: Malaczkov 2020; távtolmácsolás: Castagnoli–Niemants 2018, Devaux 2018, Seresi 2016) , felerősödött a szaknyelvi fókuszú szövegek iránti érdeklődés (például jogi szaknyelv: Vincze 2018; EU-s kontextus: Jablonkai 2009; műszaki szaknyelv: Nagy 2020; orvosi szaknyelv: Mány 2020), illetve általánossá vált az angolon kívüli nyelvek fordításközpontú vizsgálatának integrálása (például ázsiai és arab nyelvek: Bakaja 2020; török: Aksan és Aksan 2018; kis nyelvek: Karakanta et al. 2018).

A korpuszalapú fordításkutatás eredményeinek hasznosítása egyre több szakterület számára kínál releváns tartalmakat. A kutatási eredmények fontosak mind az autentikus,

mind a célnyelvi szövegprodukció leírása, az anyanyelvhasználat tudatosítása, a fordításnyelv sajátosságainak feltárása (Balaskó 2007) és a célnyelvi jellemzők célnyelvi normától eltérő használatának megvilágítása szempontjából(Seidl-Péch 2011). Az eredmények közvetlenül és közvetve is felhasználhatók a fordításoktatásban és a nyelvtechnológiai alkalmazások használatakor (például gépi fordítás, online szótárak).

3. Korpuszépítés

A korpuszalapú kutatások eredményeinek érvényessége és megbízhatósága szempontjából a korpusztervezés kiemelkedő jelentőséggel bír (vö. Robin et al. 2017), mivel a korpusz(ok)ba rendezett szövegek válogatásának szisztematikussága alapján dönthető el, hogy a korpuszban tárolt mintára vonatkozó megállapítások általánosíthatók-e a vizsgálni kívánt szövegtípusra. A mintavétel tekintetében elvárható tehát az egyes kutatásoktól, hogy a vizsgált korpusz(ok)ba rendezett szövegek vagy szövegrészletek az adott nyelv vagy nyelvváltozat vertikális és/vagy horizontális rétegződését a teljesség igényével reprezentálják (vö. Seidl-Péch 2017). Természetesen a mintavétel nem zárja ki a kutatás leszűkítését egy bizonyos modalitás/szövegtípus/korszak/szerző vagy fordító vizsgálatára, ugyanakkor a kutató felelőssége egy olyan megbízható kritériumrendszer kidolgozása, amely alapján a korpuszba válogatott szövegek bekerülési esélye megegyezik az adott korpuszba éppen fel nem vett, de a korpusz felépítése (korpusztervezés) szempontjából szintén beválogatható szövegek bekerülési esélyeivel.

A szisztematikus, de egyidejűleg véletlenszerű mintavétel kritériumának teljesítése igazi kihívás elé állítja a kutatókat, hiszen a minta végessége miatt mindenképpen szükséges a módszertani lépések részletes igazolása. A korpuszalapú kutatások általánosíthatósága tehát nagy mértékben függ attól, hogy a tervezés során a kutatók eleget tesznek-e a minőségbiztosítási kritériumoknak. Az adatgyűjtés folyamán ennek megfelelően a kutatás elméleti kereteinek megfelelő szövegek sokaságából úgy kell kiválasztani a korpuszba felvett elemeket, hogy a szelekció során az összes potenciálisan választható szöveg egyenlő eséllyel szerepeljen a mintában. Erre még akkor is kiemelt figyelmet kell fordítani, ha

másrészről evidenciának tekintjük azt a tényt, hogy a minta sohasem képezheti le teljes mértékben azt a sokaságot, amelyet reprezentálni hivatott (vö. Dörnyei és Csizér 2012). A kutatónak ugyanakkor számolnia kell több olyan, a kutatás tervezését és lefolytatását befolyásoló tényezővel, amelyek korlátozzák a mintavétel véletlenszerűségét, mivel a vizsgált populációt érintő adottságok mindenképpen befolyásolják a mintavételt.

3.1. Mintavételi nehézségek

Amint az előzőekből is kitűnik, a kvantitatív nyelvészeti kutatások a korpuszba összegyűjtött mintára vonatkozó mennyiségileg is feldolgozható információk alapján kívánnak a minta által reprezentált sokaság tulajdonságaira rámutatni. Ennek következtében egy bizonyos nyelv vagy nyelvváltozat adott vertikális és/vagy horizontális rétegződésének megfelelő nyelvi adathalmaz, azaz sokaság vizsgálatának esetében elvárható, hogy a sokaságot reprezentáló korpusz elemzése alapján bemutatott megállapítások a korpuszon kívül is érvényesek maradjanak, azaz ne csak az adott mintát jellemezzék, hanem a minta által reprezentált sokaságot is. Ezen elvárásnak való megfelelés teljesülése jelenti a korpuszalapú kutatások esetében a legnagyobb mintavételi nehézséget.

Ezzel kapcsolatban már a téma tárgyalása elején le kell szögezni, hogy téves az a napjainkban egyre inkább terjedő felfogás, miszerint az egyre nagyobb minta egyre jobb mintavételt eredményez. Kétségtelen tény, hogy a korpuszok reprezentativitása szempontjából fontos szerepet játszik a kutatási kérdések nagy adatmennyiségeken való tesztelése, ugyanakkor nem lehet figyelmen kívül hagyni ezen nyelvi adatok összeválogatásának szempontjait sem. Ez utóbbiak hozhatók összefüggésbe a korpuszok kiegyensúlyozottságával.

Továbbá azt is figyelembe kell venni a korpuszalapú nyelvészeti kutatásoknál, hogy a minta kiválasztásának alapjául szolgáló nyelvi adatok száma a legtöbb vizsgálat esetében végtelen nagyságú, és ennek következtében nem beszélhetünk matematikai módszerekkel pontosan körülírható mintavételi eljárásról, hanem sokkal inkább a kutatási szempontrendszer alapján praktikus elvek mentén összeállított korpuszokról. A kutatónak

meg kell elégednie az ideális mintavétel helyett a kutatási céloknak megfelelő, a vizsgált szempontok alapján rétegzett mintával, amelynek hiányosságaira a mintavétel bemutatásánál mindenképpen reflektálnia kell.

A korpuszalapú kutatások bemutatásának amúgy is kiemelten fontos része a korpuszok összeállításának és a korpuszban tárolt szövegek és/vagy szövegrészletek kiválasztási metódusának leírása. Ennek hiányában az olvasóban számos kétely fogalmazódik meg a minta alapján kapott eredmények **érvényességére** vonatkozóan. Annak ellenére, hogy a korpuszalapú kutatások elsődleges célja a lekérdezési eredmények és a belőlük levonható következtetések tárgyalása, ezek a kutatások nem értelmezhetők az adatgyűjtő eszköz részletes bemutatása nélkül, amely az eredmények érvényességét támasztja alá a mért változók szakirodalmi áttekintéséhez hasonlóan. Az adatgyűjtési megfontolások és lépések részletes és alapos bemutatása biztosítja többek között a kutatás **megismételhetőségét**, illetve a mért eredmények **összevethetőségét**. Ha például a bemutatott mintavétel alapján valaki egy későbbi időpontban megismétli az adott kutatást, akkor az első kutatás eredményeivel megegyező eredmények igazolni tudják az előző kutatás **megbízhatóságát** (vö. Dörnyei és Csizér 2012).

A kutatási eredmények összevetésére akkor kerülhet sor, ha egy következő kutatás az előbbi mintavételére támaszkodva pusztán egyetlen kritérium alapján változtatja meg a minta összetételét. Ilyen lehet például egy újabb nyelvpár vagy egy másik szövegtípus esetében az adott vizsgálat megismétlése. Ugyanakkor igen gyakori problémát okoz, amikor egyes kutatók egy korábbi kutatás mintavételi kritériumai közül egyszerre többet is megváltoztatva kívánnak a korábbi kutatás eredményeire reflektálni, illetve amikor egy kutatáson belül az alkorpuszok összeállítása több tulajdonság esetében sem halad ugyanazon **mintavételi kritériumrendszer** szisztematikus végigvitele mentén. Ilyen esetekben nem állapítható meg teljes bizonyossággal, hogy az eltérések mely változók mentén jöttek létre, és ebből következően nem vonhatók le egyértelmű következtetések.

További bizonytalanságot okozhatnak a mintavétel során tapasztalható belső aránytalanságok, amikor egy-egy szövegtípus, nyelvpár, szerző, téma, műfaj stb. valamilyen okból kifolyólag (például könnyebben vagy nehezebben elérhető szövegek) felül-

vagy alulreprezentált a korpuszban. Ilyen esetben nehezen bizonyítható, hogy a korpusz összetételének esetleges megváltoztatásával nem módosulnának-e a lekérdezések eredményei, ezért a kutatás során feltárt összefüggések sem tekinthetők **érvényesnek**.

3.2. Kvantifikálható szempontrendszer

A kutatás másik lényeges jellemzője a kutatási kérdés(ek) megfogalmazásának szükségessége a kutatás megkezdése előtt, amelyek természetesen nem zárják ki, hogy a kutatás közben újabb és újabb feltárandó kérdések merüljenek fel. Ugyanakkor a vizsgálandó kérdések meghatározása a kutatás elején és ezzel párhuzamosan a feltételezett eredmények hipotézisek formájában való megfogalmazása elengedhetetlen annak számbavételéhez, hogy a tervezett kutatás valóban elvégezhető-e kvantitatív kutatási módszerekkel. Másként fogalmazva a kutatás tervezési szakaszában el kell dönteni, hogy a vizsgálni kívánt kérdés esetében a kvantitatív lekérdezés és az azt lehetővé tevő nagy adatmennyiség gyűjtése a célszerű módszer-e, vagy előnyösebb a kérdés kisebb mintán végzett kvalitatív vizsgálata. Míg a nagyszámú nyelvi minta vizsgálata alapján végzett kvantitatív kutatások többnyire hipotézisek megfogalmazásával, előfeltételezések alapján keresik a választ egy-egy nyelvi minta működésének jellemzőire és gyakoriságára, addig a kisebb mintán végzett kvalitatív kutatások az adott minta működésének ok–okozati összefüggéseit is fel tudják tárni.

A korpuszalapú vizsgálatok igen fontos jellemzője, hogy a vizsgálatok tárgyai a valós nyelvi előfordulások, így a lekérdezések eredményei a tényleges nyelvi produktum (például szövegkutatások) vagy folyamat (például elakadásvizsgálat a tolmácsoláskutatásban) elemzését és feltárását teszik lehetővé. Ugyanakkor a kutatónak a kutatási kérdéseket mindenképpen úgy kell megfogalmaznia, hogy az eredményeket számszerűsíthető adatok formájában tudja lekérdezni és elemezni. A fordítás-/tolmácsoláskutató vizsgálhatja például, hogy a fordítók/tolmácsok mely nyelvi jelenséget (például lexikai elemet, szókapcsolatot) használják gyakrabban vagy ritkábban, mint az anyanyelvi beszélők, vagy éppen

mely elemek használata nem jelenik meg a fordított/tolmácsolt célnyelvi szövegekben a forrásnyelvi stimulus hiányában (például egyedi nyelvi elemek, Dankó 2017).

A korpuszalapú fordítástudományi vizsgálatok esetében végzett leggyakoribb leíró statisztikai lekérdezések a **szövegszavak számát** (az összes szövegszó gyakoriságtól független számát), a korpuszban szereplő különböző **szótári szavak számát** (szótípus) és a **szótípus/szövegszó arányt** (a korpuszra jellemző lexikai változatosságot) vizsgálják (vö. Laviosa 1998b). Jellemzőek továbbá a **betűgyakorisági** listák (az adott nyelvre vagy szerzőre jellemző betűeloszlás), a **gyakoriság szerinti szólisták** (a szöveg(ek)ben gyakrabban és kevésbé gyakran – akár csak egyszer – előforduló szavak), a szöveg feldolgozhatósága és az egyszerűsítés szempontjából meghatározó **átlagos szó- és mondathossz** (a szövegben található összes betű/szó száma elosztva az összes szó/mondat számával) lekérdezései, a **kulcsszavak** szűrése (egy hosszabb szöveg szólistájához viszonyítva a vizsgált szöveg szólistájában gyakrabban előforduló szavak), illetve a klaszterek vagy **N-gramok** elemzése (a szövegben szereplő több egységből álló szerkezetek), ahol az N helyére kerülő szám határozza meg, hogy a szövegben szereplő szavak hány egységes előfordulását vizsgáljuk. Ez utóbbi segíthet például feltárni a terminusjelölteket, illetve a szövegben szereplő formulaszerű elemeket (Nagy 2019). A gyakorisági listák és a kvantitatív elemzések segítségével vizsgált felszíni jelenségek jó alapot kínálnak a mélyebb szerkezeti jellemzők feltárásához, a kvantitatív kutatások eredményei alapján megkezdett kvalitatív kutatások lefolytatásához.

Az egyszerű statisztikai lekérdezéseken túl a korpusz felszíni jegyei további elemzéseket is lehetővé tesznek, amennyiben a korpusz annotált, azaz a nyelvészeti elemzés számára érdekes jelenségek megjelölésére metanyelvi többletinformációt tartalmaz. A kézi vagy gépi annotálás céljára általában az úgynevezett jelölő (Mark-up) nyelveket használják, melyek közül a HTML, SGML, XML (Hyper Text Markup Language, Standard Generalized Markup Language, Extensible Markup Language) a legelterjedtebbek. Az annotálásra használt jelölőelemek (tagek) szabadon bővíthetők, de használatuk a TEI (The Text Encoding Initiative) által szabályozott szigorú szintaxishoz kötött. Az annotációként megjelenő többletinformációk (pl. bekezdések, mondathatárok, szótövek és szófajok jelö-

lése) alkalmassá teszik a vizsgált korpuszt/szöveget többek között grammatikai (például Sass et al. 2011) vagy szintaktikai (például Seidl-Péch 2011) összefüggések feltárására is.

4. Korpuszalapú elemzés

A korpuszalapú fordításkutatást jellemző kvantitatív vizsgálatok legfontosabb tulajdonsága talán abban összegezhető, hogy a hipotézisek megerősítése vagy cáfolata nagy szövegállományok korpuszalapú lekérdezésével történik. A lekérdezési eredmények elsősorban a korpuszban gyakran és jellemzően előforduló mintázatok kutatását támogatják, bár nem lehetetlen a ritkábban előforduló mintázatok vizsgálata sem. Ez utóbbiak esetében a kutatási kérdések megjelenése általában nem a korpusz tanulmányozásához köthető, hanem azok már korábban felmerültek, és a korpusz inkább csak a már meglévő hipotézisek tesztelésére szolgál (például egyedi nyelvi elemek vizsgálata, terminuseloszlás vizsgálata). Ugyanakkor a gyakran előforduló, jellemző mintázatok esetében egyáltalán nem ritka, hogy maguk a korpuszvezérelt lekérdezések során nyert kvantitatív eredmények hívják fel a kutatók figyelmét egy-egy tipikus nyelvi jelenség létére (például konkordancia vizsgálatok).

Az utóbbi időben elterjedt az a hibás felfogás is a kutatók körében, hogy a kvantitatív vizsgálatok eredményei annál megbízhatóbbak, minél nagyobb tokenszámú (szóközzel határolt szavak száma) korpuszon történnek a lekérdezések. Mivel a korpuszok méretéből még nem lehet egyértelműen következtetni a mintavétel pontosságára, vagyis hogy a korpusz mennyire pontosan reprezentálja a vizsgálni kívánt nyelvet/nyelvváltozatot, ezért a diszciplínát jellemző paradigmaváltás következtében általános elvárássá vált a korpuszalapú kvantitatív vizsgálatok eredményei tekintetében a **statisztikai szignifikancia vizsgálata** is (Bisiada 2017: 242). Ez utóbbi segít annak alátámasztásában, hogy a minta vizsgálata során észlelt mérési eredmények nem a véletlenek összejátszásából, a mintavétel hibájából vagy valamely mérési hibából következnek, hanem valóban jellemzik a vizsgálni kívánt nyelvet/nyelvváltozatot. Ha ugyanis a statisztikai szignifikancia vizsgálatok alapján a lekérdezések során kapott eredmények nem szignifikánsak, akkor

önmagukban ezen eredmények alapján még nem lehet a korpuszban reprezentált nyelvre/nyelvváltozatra egyértelmű következtetéseket levonni.

A kutatási céloknak megfelelően felépített reprezentatív és kiegyensúlyozott korpuszok önmagukban még nem garanciái a kutatás sikerének. A kutatás bemutatásánál azt is célszerű igazolni, hogy az egymást követő lépések milyen **módszertani elveket** követnek (vö. Dörnyei és Csizér 2012). A kutatáshoz kapcsolódó minőségbiztosítási folyamat támaszkodhat egy már korábban elvégzett sikeres kutatásra, amelynek eredményei már igazolódtak, vagy magának a kutatásnak egy kisebb mintán történő kipróbálására. Mindkét esetben fontos szerepet játszik a jelenlegi kutatás kontextusának részletes bemutatása, amely alapján eldönthető, hogy a példaként használt vagy előzetesen elvégzett pilot kutatás megfelel-e módszertanilag a jelenlegi kutatás kontextusának. Ebből arra is lehet következtetni, hogy az aktuális kutatás az alkalmazott módszertan segítségével valóban a vizsgálni kívánt kérdésekre ad-e választ. Már a pilot projekt során is mindenképpen meg kell arról győződni, hogy a kutatási kérdések és a mérési eszköz összhangban vannak-e, tehát a kutatás során valóban a feltett kérdésekre kapunk-e válaszokat, illetve hogy milyen a mérési eredmények minősége. A kvantitatív korpuszalapú kutatások **érvényessége** tehát arra vonatkozik, hogy a kapott eredmények valóban alátámasztják vagy cáfolják-e a hipotéziseket, míg a **megbízhatóságuk** az elemzés megismételhetőségére (vö. Károly 2002), a következetes osztályozásra és a számítások helyességére enged következtetni (vö. Dörnyei 2011: 48–78).

A kutatás **belső érvényessége**, azaz annak teljesülése, hogy a vizsgált függő változókban valóban a független változók okoznak-e változást, nagy mértékben függ attól, hogy a kutatás figyelembe veszi-e az összes olyan változót, amely befolyásolhatja a mért eredményeket. Az egyes lekérdezések során a kutatóknak ügyelniük kell arra is, hogy ezen változók közül mindig csak egyet változtassanak meg az eredmények kiértékelhetősége érdekében. A belső érvényességgel rendelkező kutatások esetében a reprezentatív és kiegyensúlyozott korpuszokból nyert adatok **külső érvényességet** is mutatnak, azaz a megállapítások érvényesek lesznek arra a nyelvre/nyelvváltozatra, amelyet ezek a korpuszok reprezentálnak.

5. Összegzés

Az ezredfordulót követő módszertani paradigmaváltás számos fordítástudományi kutatás esetében a korpuszalapú megközelítést helyezte előtérbe, amelyek között az eredményorientált kutatások mellett egyre nagyobb számban jellennek meg az előkészítés szempontjából sokkal nagyobb körültekintést igénylő folyamatorientált kutatások is. Az elmúlt évtizedek gyakorlatából kiindulva, a korpuszalapú fordítástudományi kutatások előkészítésének és lefolytatásának metódusa egyfajta purifikációs folyamaton esett át. Részben a rendelkezésre álló eszközök technikai kapacitásának növekedése, részben az újabban megjelenő kutatási szempontok integrálása irányította a kutatók figyelmét a körültekintőbb adatgyűjtés fontosságára és a lekérdezések statisztikai validálására. Az ezek hatására gyökeresen megújuló fordítástudományi korpuszalapú kutatások szakítani látszanak a területet jellemző kezdeti átgondolatlanságokkal és felszínességgel: a második generációs kutatásokat egyre inkább összeköti egyfajta következetes metodológiai koncepció, amely a kutatási eredmények megbízhatóságáért is felel.

Irodalom

Aksan, Y.; Aksan, M. 2018. Linguistic corpora: A view from Turkish. In: Oflazer, K., Saraçlar, M. (eds) *Turkish Natural Language Processing. Theory and Applications of Natural Language Processing.* Springer International Publishing. 301–327. https://doi.org/10.1007/978-3-319-90165-7_14

Bakaja Z. 2020. Az indiai vallási irodalomban használatos beszélőjelölő igéjének magyar fordítása. In: Robin E., Seidl-Péch O. (szerk.) *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordításkutatás Magyarországon.* Segédkönyvek a nyelvi közvetítésről I. Budapest: ELTE BTK Fordítástudományi Doktori Program, MANYE Fordítástudományi Szakosztály. https://doi.org/10.36252/Nyelvikozvsegedkonyv1.9

Balaskó M. 2005. Korpusznyelvészeti vizsgálatok és fordításnyelvi minták (angol és magyar tudományos szövegek anyaga alapján). Doktori értekezés. Budapest: ELTE.

Balaskó M. 2007. A fordításnyelvről, avagy a flamand szőnyeg láthatatlan szálairól. In: Heltai P. (szerk.) 2007. *Nyelvi modernizáció. Szaknyelv, fordítás, terminológia. A XVI. MANYE Kongresszus előadásai.* Gödöllő. 2006. április 10−12. (A MANYE Kongresszusok előadásai 3.) Pécs–Gödöllő: MANYE–Szent István Egyetem. 159–166.

Bańos, R., Bruti, S., Zanotti, S. 2013. Corpus linguistics and Audiovisual Translation: in search of an integrated approach. *Perspectives* Vol. 21. 483–490. https://doi.org/10.1080/0907676x.2013.831926

Bisiada, M. 2017. Universals of editing and translation. In: Hansen-Schirra, S., Czulo, O., Hofmann, S. (eds) *Empirical modelling of translation and interpreting*. Berlin: Language Science Press. 241–275. https://doi.org/10.5281/zenodo.1090972

Bowker, L., Pearson, J. 2002. *Working with specialized language. A practical guide to using corpora*. London: Routledge. https://doi.org/10.4324/9780203469255

Castagnoli, S., Niemants, N. 2018. Corpora worth creating: A pilot study on telephone interpreting. *InTRAlinea* Különkiadás. Elérhető: http://www.intralinea.org/specials/cbis

Dankó Sz. 2017. Alulreprezentált célnyelvi elemek a fordításban, avagy a „szokott" esete *Fordítástudomány* 19. évf. 1. szám 75–84.

Devaux, J. 2018. Technologies and role-space: How videoconference inter- preting affects the court interpreter's perception of her role. In: Fantinuoli, C. (ed.) *Interpreting and technology*. Berlin: Language Science Press. 91–117. https://doi.org/10.5281/zenodo.1493297

Dörnyei, Z. 2011. Research Methods in Applied Linguistics. Quantitative, Qualitative, and Mixed Methodologies. Oxford: Oxford University Press.

Dörnyei, Z., Csizér, K. 2012. How to Design and Analyze Surveys in Second Language Acquisition Research. In: Mackey, A., Gass, S. M. *Research Methods in Second Language Acquisition: A Practical Guide.* New Jersey: Wiley-Blackwell Publishing. 74–94. https://doi.org/10.1002/9781444347340.ch5

Fantinuoli, C. 2018. Interpreting and technology: The upcoming technological turn. In: Fantinuoli C. (ed.) *Interpreting and technology*. Berlin: Language Science Press. 1–12. http://doi.org/10.5281/zenodo.1493289

Holmes, J. S. 1975. The Name And Nature Of Translation Studies. In: Holmes, J. S. 1988. *Translated!: Papers on Literary Translation and Translation Studies*. Amsterdam: Rodopi. 66–80.

Jablonkai, R. 2009. „In the light of": A corpus-based analysis of lexical bundles in tow EU-related registers. *WoPaLP* Vol 3.

Karakanta, A., Dehdari, J., van Genabith, J. 2018. Neural machine translation for low-resource languages without parallel corpora. *Mach Tranlat* No. 32. 167–189. https://doi.org/10.1007/s10590-017-9203-5

Károly K. 2002. Az alkalmazott nyelvészeti kutatások néhány alapvető módszertani kérdéséről. *Alkalmazott Nyelvtudomány* 2. évf. 1. szám. 77–87.

Károly K. 2003. Korpusznyelvészet és fordításkutatás. *Fordítástudomány* 5. évf. 2. szám. 18–26.

Klaudy K. 2004. Fordítástudomány az ezredfordulón. *Publicationes Universitatis Miskolciensis* Vol. 9. 157 – 170. Elérhető: http://efolyoirat.oszk.hu/02100/02137/00001/pdf/EPA02137_ISSN_1219-543X_tomus_9_fas_1_2004_hun_eng_157-170.pdf

Klaudy K. 2005. Párhuzamos korpuszok felhasználása a fordításkutatásban. In: Lanstyák I. és Vančoné Kremmer I. (szerk.). *Nyelvészetről – változatosan. Segédkönyvek egyetemisták és a nyelvészet iránt érdeklődők számára*. Dunaszerdahely: Gramma Nyelvi Iroda. 153–185.

Laviosa, S. 1998a. The Corpus-based Approach: A New Paradigm in Translation Studies. In: *Meta: Translators' Journal* Vol. 43. No. 4. 474–479. http://doi.org/10.7202/003424ar

Laviosa, S. 1998b. Core Patterns of Lexical Use in a Comparable Corpus of English Narrative Prose. *Meta: Translators' Journal* Vol. 43. No. 4. 557–570. https://doi.org/10.7202/003425ar

Malaczkov Sz. 2020. A vonatkozó mellékmondat használata a feliratokban: TED-előadások magyar nyelvű összehasonlítható korpuszának vizsgálata. In: Robin E., Seidl-Péch O. (szerk.) *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordításkutatás Magyarországon.* Segédkönyvek a nyelvi közvetítésről I. Budapest: ELTE BTK Fordítástudományi Doktori Program, MANYE Fordítástudományi Szakosztály. https://doi.org/10.36252/Nyelvikozvsegedkonyv1.10

Mány D. 2020. Idegen szavak félautomatikus elemzése autentikus és fordított orvosi szövegekben: fordítási stratégiák angolról franciára és magyarra fordított betegtájékoztatók tükrében. In: Robin E., Seidl-Péch O. (szerk.) *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordításkutatás Magyarországon.* Segédkönyvek a nyelvi közvetítésről I. Budapest: ELTE BTK Fordítástudományi Doktori Program, MANYE Fordítástudományi Szakosztály. https://doi.org/10.36252/Nyelvikozvsegedkonyv1.6

McCarthy, M., O'Keeffe, A. 2010. Historical perspective: what are corpora and how have they evolved? In: McCarthy, M., O'Keeffe, A. (eds) *The Routledge Handbook of Corpus Linguistics.* New York: Routledge. 3–13. https://doi.org/10.4324/9780203856949.ch1

Nagy A. L. 2020. A szógyakoriság gépi vizsgálata műszaki szövegeken: terminusok és műszaki formulaszerű elemek. In: Robin E., Seidl-Péch O. (szerk.) *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordításkutatás Magyarországon.*

Segédkönyvek a nyelvi közvetítésről I. Budapest: ELTE BTK Fordítástudományi Doktori Program, MANYE Fordítástudományi Szakosztály. https://doi.org/10.36252/Nyelvikozvsegedkonyv1.8

Robin E. 2018. *Fordítási univerzálék és lektorálás.* Budapest: Eötvös József Kiadó. http://www.eltereader.hu/kiadvanyok/robin-edina-forditasi-univerzalek-es-lektoralas/

Robin, E., Götz, A., Pataky, É., Szegh H. 2017. Translation Studies and Corpus Linguistics: Introducing the Pannonia Corpus. In: *Acta Universitas Sapientiae Philologica* Vol. 9. No. 3. 99–116. https://doi.org/10.1515/ausp-2017-0032

Sass B., Váradi T., Pajzs J., Kiss M. 2011. *Magyar igei szerkezetek. A leggyakoribb vonzatok és szókapcsolatok szótára.* Budapest: Tinta Könyvkiadó.

Seidl-Péch O. 2011. *Fordított szövegek számítógépes összevetése*. In: Bocz Zs., Sárvári J. (szerk.) 2013. *Válogatott cikkek, tanulmányok, 2010–2013*. Budapest: BME GTK Idegennyelvi Központ. 369–386.

Seidl-Péch O. 2018. Melyek a (szak)fordító és a fordításkutató munkáját segítő legfontosabb nyelvi korpuszok? In: Robin E., Zachar V. (szerk.) *Fordítástudomány ma és holnap*. Budapest: L'Harmattan Kiadó. 175–191.

Seidl-Péch O. 2019. *Korpusznyelvészeti kutatások a fordítástudomány szolgálatában*. Elhangzott: Nyelvi Közvetítés a digitalizáció korában. Miskolc: Miskolci Egyetem, BTK. (2019. január 23.)

Seidl-Péch O.; Robin E. 2019. *Alkalmas-e a korpuszalapú módszer a terminuskezelési stratégiák kutatására?* In: Dróth, J. (szerk.) *Korpusz és kontrasztivitás a szakfordítás oktatásában és gyakorlatában.* Budapest: Károli Gáspár Református Egyetem – L'Harmattan Kiadó. 93–104.

Seresi M. 2016. *Távtolmácsolás és távoktatás a tolmácsképzésben*. Budapest: ELTE Eötvös Kiadó.

Vincze V. 2018. A Miskolc Jogi Korpusz nyelvi jellemzői. In: Szabó M., Vinnai E.: *A törvény szavai. Prudentia Iuris* Vol. 33. 9–36

Explicitáció

– a fordítás univerzális jellemzője?

Pápai Vilma

Széchenyi István Egyetem

Kivonat: A jelen tanulmány korpuszalapú vizsgálat alapján mutatja be az explicitációt, amelyre többnyire a fordítás univerzális jellemzőjeként hivatkozik a szakirodalom. A kutatás angol és magyar nyelvű szövegeket tartalmazó párhuzamos korpusz, valamint autentikus magyar szövegek összehasonlítható korpusza alapján végzett kettős elemzésre épül. Célja a szabályszerűségek feltárása mind a fordítási folyamatban végbemenő explicitáció, mind a fordítás eredményének explicitsége tekintetében. A tanulmány álláspontja, hogy szoros összefüggés van az explicitáció és az egyszerűsítés között, amely egy újabb potenciális fordítási univerzálénak tekinthető.

Kulcsszavak: explicitáció, explicitség, fordítási univerzálé, korpuszkutatás, egyszerűsítés

1. Bevezetés

Mivel minden szöveget meghatároz egy bizonyos cél, amelyért létrehozták, egy bizonyos kontextus, amelyben született, egy bizonyos hallgatóság, amelynek szól, a fordított szövegek szükségszerűen eltérnek a nem fordított szövegektől. Az egyik meghatározó különbség a szövegalkotás céljában rejlik. A szerző alapvető célja, hogy új, élő szöveget alkosson: „Mert a szerző mindig új mondatot csinál, kreál, olyat, amilyen addig még nem volt abban a nyelvben” (Esterházy 1996: 182); a fordító azonban egy olyan szöveget ad vissza, amelyet valaki más hozott létre. Más szóval a szöveg szerzője a szavak egyedi formájának létrehozására törekszik, hogy rögzítse és továbbítsa mondanivalóját, legyen

Pápai Vilma: Explicitáció – a fordítás univerzális jellemzője? (Ford.: Heltai Edit Fruzsina, Juhász Dóra, Kis Elizabet, Nagy Laura, Szász Barna) In: Robin E., Seidl-Péch O. (szerk.) 2020. *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordításkutatás Magyarországon.* Segédkönyvek a nyelvi közvetítésről I. Budapest: ELTE BTK Fordítástudományi Doktori Program, MANYE Fordítástudományi Szakosztály. DOI: https://doi.org/10.36252/Nyelvikozvsegedkonyv1.5

ez történet, kapcsolat vagy egy gondolat. Másrészről a fordító célja, hogy olyan szöveget fogalmazzon meg, amellyel valaki más mondanivalóját rögzíti és közvetíti – olyan gondolatot, amelyet elsőként olyan, a fordító és az olvasóközönség nyelvi formájától eltérő formában alkottak (és fogalmaztak!) meg. Baker szavaival élve: „A fordított szöveget rendszerint korlátozza egy más nyelven már készre formált szöveg" (Baker 1996: 177).

Az elmúlt évtizedre jellemző, hogy egyre nagyobb méreteket ölt a fordítás útján születő szöveg természetének, nyelvi sajátosságainak és a rá jellemző diskurzus típusának vizsgálata, különösen korpuszalapú módszerek segítségével.

2. Elméleti háttér

2.1. Explicitáció

Az explicitáció azon szövegsajátosságok egyike, amelyre a fordított szöveg sajátosságaként, univerzáléjaként hivatkozik a fordítástudományi szakirodalom. Számos tanulmány foglalkozik Blum-Kulka alábbi hipotézisével, amely

> […] azt állítja, hogy a FNY-i szövegről a CNY-i szövegre való áttérés folyamatában a kohézív explicitség növekedése figyelhető meg, függetlenül a két nyelv és a két szöveg rendszereinek különbségeire visszavezethető növekedéstől. (1986: 19)[1]

A fordítástudományban egészen mostanáig két fő megközelítést alkalmaztak a hipotézis igazolására, illetve megcáfolására. A közelmúltig a kutatások elsősorban a forrás- és a célnyelvi szöveg összehasonlításán alapultak. Következésképpen a kapott eredmények – ahogyan Toury (1995: 77) nevezi – „(ad hoc) nyelvpárok" összehasonlító elemzésére épülnek, mint például a holland–angol (Vanderauwera 1985), angol–francia és francia–angol (Blum-Kulka 1985, Séguinot 1988), héber–angol (Weissbrod 1992), an-

1 Klaudy Kinga fordítása (1999) Az explicitációs hipotézisről. *Fordítástudomány* 1. évf. 2. szám. 7. oldal

gol, francia, orosz, német–magyar (Klaudy 1993a, 1993b, 1996), angol–héber (Shlesinger 1995), valamint norvég–angol és angol–norvég (Øverås 1996) nyelv viszonylatában.

Végeredményben számos szövegsajátosságot sikerült azonosítani az elméleti és/vagy empirikus kutatás során. Az *1. táblázat* azokat a fő szövegjellemzőket mutatja be, amelyekre a fordított szövegeknek az autentikus szövegekkel szemben mutatott magasabb szintű explicitség formáiként tekinthetünk: terjedelem növekedése, nagyobb mértékű redundancia, erősebb kohéziós és logikai kapcsolatok, könnyebb olvashatóság, jellegzetes központozás és világosabb téma–réma kapcsolat. A táblázat továbbá megjeleníti az explicitációról mint stratégiáról formált nézeteket és álláspontokat is azt a fordításelméleti dilemmát illetően, vajon az explicitáció professzionális stratégia vagy a nyelvi közvetítés melléktermé-e.

Az egynyelvű összehasonlítható korpuszok használatának elterjedésével a fordított szövegek kutatásának teljesen új megközelítése tört utat magának; ezt a módszertani váltást „egynyelvű fordulatnak" is nevezhetjük. Baker (1995: 234) a következőképpen foglalja össze az összehasonlítható korpusz jelentőségét:

> Az összehasonlítható korpusz legfontosabb hozzájárulása a diszciplínához a fordított szövegekre jellemző nyelvi minták feltárása, függetlenül a forrás- vagy célnyelvtől.

A fordítástudományi – elsősorban a potenciális univerzálékra irányuló – vizsgálódások során egyre több kutató fogadja el a fordított szövegek elemzésének ezt az alternatív módszertani megközelítését (Laviosa-Braithwaite 1996; Kenny 1999; Olohan és Baker 2000), éppen ezért a kizárólag párhuzamos korpuszokra épülő szövegösszehasonlító megközelítés veszíteni látszik relevanciájából (lásd Laviosa 1998).

Olohan és Baker (2000) vezette be kutatásukban a nyelvi rendszer opcionális elemeinek használatában rejlő szabályszerűségek vizsgálatát. A Translational English Corpus (TEC) és a British National Corpus (BNC) adatbázisát vizsgálva felhívták a figyelmet

a *that* kötőszó függő beszéd bevezetésekor történő használatára angol nyelvű fordított szövegekben.

2.2. Definíciók és hipotézisek

Az explicitáció megvitatásához szükséges, hogy a fogalmat mind a fordítói folyamat, mind a fordítás eredménye szempontjából értelmezzük. A jelen kutatás célját tekintve a következő definíciót dolgoztam ki az explicitáció fogalmának meghatározására.

A folyamat szempontjából az explicitáció egy fordítási technika, amely a forrásnyelvi szövegtől való szerkezeti és tartalmi eltolódást jelenti. Ez egy technika a félreérthetőség feloldására, a forrásnyelvi szöveg kohéziójának fejlesztésére és fokozására, valamint a nyelvi és nyelven kívüli információk átadására. A fő cél, hogy a fordító tudatos vagy nem tudatos erőfeszítései megfeleljenek a célközönség elvárásainak. *Az eredmény tekintetében* az explicitáció egy szövegsajátosság, amely hozzájárul az explicitség egy magasabb szintjéhez, az autentikus szövegekhez viszonyítva. Megmutatkozhat olyan nyelvi sajátosságokban, amelyek gyakrabban fordulnak elő fordított, mint autentikus szövegekben, vagy hozzáadott nyelvi és nyelven kívüli információkban.

Ezek tudatában a következő hipotéziseket fogalmaztam meg: (1) az angolból magyarra való fordítás folyamatában a két nyelv közötti szerkezeti különbségek ellenére megjelennek az explicitációs stratégiák, (2) a fordított magyar nyelvi szövegek explicitebbek, mint az autentikus magyar szövegek, és (3) az explicitség mértéke magasabb a tudományos, mint a szépirodalmi szövegekben.

1. táblázat

Az explicitáció jellegének és formáinak szakirodalmi összegzése

	Természete						**Formája**			
	Tudatos stratégia	Fordítást kísérő jelenség	Hosszabb szöveg	Redundancia	Kohéziós elemek erősödése	Logikai kapcsolatok	Olvashatósági kapcsolatok	Központozás	Téma–réma	Opcionális
										that
Blum-Kulka 1986		+	+	+	+					
Séguinot 1988	+*	+	+**		+		+	+	+	
Klaudy 1993		+	+		+		+			
Baker 1993, 1995, 1996		+	+		+	+				
Shlesinger 1995		+			+	+				
Toury 1995							+			
Øverås 1998			+		+					
Ishikawa 1998	+				+					
Olohan és Baker 2000		+			+					+

* tudatos stratégia a szerkesztő részéről

** bizonyos nyelvek több szóval fejezik ki ugyanazt az üzenetet (francia az angollal összehasonlítva)

3. A kutatás bemutatása

3.1. A korpusz összeállítása, szerkezete és mérete

A vizsgálat alapjául szolgáló korpusz (a későbbiekben: ARRABONA korpusz) három, 1969 és 1999 között keletkezett szépirodalmi (SZ) és tudományos szövegekből (T) álló alkorpuszból tevődik össze.

(1) az eredetileg angol nyelven létrehozott szövegekből (EAK) álló alkorpusz 8 szövegből áll, melyeket brit, amerikai és kanadai szerzők írtak,

(2) a fent említett szövegek magyar fordításából álló alkorpusz (FMK) hivatásos fordítók által készített és neves kiadók által publikált szövegeket tartalmaz,

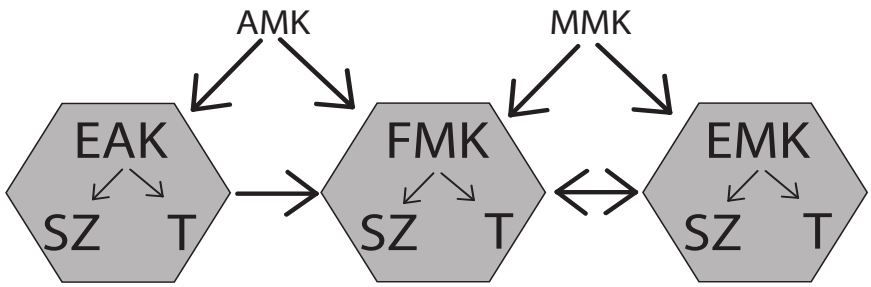
(3) az eredetileg magyar nyelven létrejött szövegekből (EMK) álló korpusz 8 összehasonlítható szöveget tartalmaz, amelyek ugyanekkor íródtak (a szöveglistát az első függelékben tekintheti meg).

Az alábbiakban az 1. ábra mutatja, hogy ezek az alkorpuszok egy párhuzamos (AMK) és egy egynyelvű összehasonlítható korpuszt (MMK) képeznek:

1. ábra

Az ARRABONA-korpusz szerkezete.

A párhuzamos korpusz (AMK) és az összehasonlítható korpusz (MMK)



A szépirodalmi és tudományos szövegek pontosan három évtizedet fognak át: 1969 és 1999 között keletkeztek. A 70-80-as évekre azért kellett visszanyúlni, hogy elemzésemben még tetten tudjam érni a klasszikusnak számító magyar könyvkiadói és fordítói gyakorlatot. Az eredetileg magyar nyelven létrehozott szövegek kiválasztásakor egy másik feltétel az volt, hogy rendelkezzenek angol nyelvű fordítással, ami megkönnyíti az elemzés későbbi kiterjesztését.

Az elemzésre szánt autentikus és fordított szövegek kiválasztásakor az általános cél és a fő elméleti megfontolás az volt, hogy a reprezentativitás érdekében minél változatosabb szövegek kerüljenek a korpuszba (Sinclair 1991: 13–36): földrajzi sajátosságok, (brit, amerikai és kanadai szerzők), nemek (férfi és női szerzők/fordítók) és társadalmi státusz tekintetében. Fordításvezérelt korpusz hiányában korlátokba ütköztem, különösen a tudományos szövegek kiválasztásakor. A tudományos szövegek azzal az elvárással kerültek be a korpuszba, hogy azok magasabb számú kohéziós kapcsolatról fognak tanúskodni, mint a szépirodalmi szövegek. A kohéziós elemek a leggyakrabban kutatott szövegsajátosságok közé tartoznak (lásd *1. táblázat*), és általuk bepillantást nyerhetünk az explicitáció természetébe.

A három alkorpusz a szövegek első 100 mondatát tartalmazza, így képviselik a szöveg egészét a szerző és a fordító stílusa, valamint a műfaji sajátosságok szempontjából. A kutatás alapjául szolgáló korpusz összességében 2 400 mondatot és körülbelül 45 000 szövegszót tartalmaz. A párhuzamos alkorpusz (AMK) összeállításában és az összehasonlítható alkorpusz (MMK) vizsgálatában a Wordsmith Tools (Scott 1998) számítógépes szövegelemző szoftvert alkalmaztam.

3.2. Vizsgálati módszerek

A korpusz két különböző elemzési módszer végrehajtásához készült. A párhuzamos korpusz (angol–magyar, AMK) vizsgálatával kísérletet tettem az explicitációs stratégiák feltárására, vagyis azon eltolódások típusainak azonosítására, amelyeket a fordítók az angol szövegek magyarra történő fordításakor alkalmaznak. A meghatározási folyamat során a

kiválasztási kritérium tágabb értelmezést tett lehetővé Blum-Kulka (1986) eredeti kategóriáinál. A listán ugyanis nemcsak a szövegkohéziós eltolódások szerepeltek, hanem olyan esetek is, amikor a fordító a szöveget nyelvi, illetve nyelven kívüli információkkal bővíti, valamint amikor a forrásnyelvi szöveg félreérthető vagy homályosabb részeit a fordító egyértelműbb módon fejezi ki a célnyelvi szövegben. Az irányadó elv az volt, hogy példákat találjak a forrásnyelvi szöveghez képest végrehajtott változtatásokra, és ezáltal feltárjam azokat a lépéseket, amelyek egy könnyebben érthető, jobban strukturált és szervezett, egyértelműbb szöveghez vezetnek.

Ez a manuális elemzés sokat elárul számunkra az angol–magyar nyelvi irányban történő *fordítás folyamatáról* is. Két nyelv összehasonlításakor olyan következtetésre juthatunk, amely a szóban forgó két nyelvre és az adott fordítási irányra korlátozódhat. Amennyiben a levont következtetéseket általánosítani kívánjuk, egy egynyelvű összehasonlítható korpusz hasznos eszközként szolgálhat, betekintést nyújtva magának a fordításnak és a fordítás eredményének a természetébe. Az elemzés másik módszere ezt a célt szolgálja. Az utóbbi eljárás gyökeresen különbözik az előzőtől: az első szakasszal ellentétben ugyanis, ahol az explicitációs stratégiákat kontrasztív szövegelemzéssel mutatom ki, a második elemzési módszer a „nagy mennyiségű" szövegek szövegjellemzőiben való megnyilvánulásként tekint az explicitségre. Ezen kívül a fordított, illetve az eredetileg is magyar nyelven keletkezett szövegekből álló összehasonlítható korpusz (magyar–magyar, MMK) vizsgálatát korpuszalapú módszerrel végeztem, gyakorisági adatok felhasználásával. A vizsgálat középpontjában az állt, hogy az angolról magyarra fordított szövegek explicitségi szintje magasabb-e, mint az autentikus magyar nyelvű szövegeké.

A két megközelítés összekapcsolódott a gyakorlatban: az első szakaszban meghatározott stratégiák közül ötöt a második szakaszban tovább vizsgáltam, és az összehasonlítható korpusz egészében teszteltem. A stratégiák kiválasztását a korpusz felépítése határozta meg. Mivel szófaji annotálás nem történt a szövegeken, csupán néhány stratégiát választottam ki a második szakaszba azok közül, amelyeket a *2.* táblázatban felsoroltam, és amelyek alkalmasak a gépi gyakorisági elemzésre. Ennek következtében a lexi-

ko-grammatikai kapcsolatok szintjét teljesen kizártam. A további elemzésre kiválasztott stratégiák a következők:

(1) írásjelbetoldás és írásjelváltás (kettőspont, pontosvessző, zárójelek);

(2) szófajváltók betoldása (*közötti* + melléknévképző, *belüli* + melléknévképző, *való*);

(3) kötőszavak betoldása (*hogy, aki, ami, amely, pedig, azonban*);

(4) kötőszavak és kataforikus referenciák betoldása (*az…, hogy, arr*..., hogy, ann*..., hogy* + mutatónévmások [három különböző esetvégződéssel]); valamint

(5) diskurzuspartikulák betoldása (*csak, még, is, például, így, tehát*).

Ezeket a nyelvi jelenségeket az MMK (*4.* táblázat) fordított és autentikus szövegeinek összevetése során az explicitség jellemzőinek tekintettem. Összefoglalva, ami az alkalmazott módszerek sorrendjét illeti, a részletes, kvalitatív szövegelemzésektől haladtam a korpuszalapú módszer általánosabb, teljes szövegállományra irányuló technikái felé.

4. Eredmények bemutatása

4.1. Az explicitációs stratégiák

Az angol nyelvű forrásszövegek és azok magyar fordításának elemzése eredményeképpen 16 típusú explicitációs stratégiát azonosítottam és kategorizáltam. A megállapítások azt sugallják, hogy az eltolódások a nyelv minden szintjén előfordulnak, a logikai-vizuális kapcsolatoktól a szövegen és nyelven kívüli elemek szintjéig (*2.* táblázat).

A kutatás ezen szakaszában azonosított műveletekkel az volt a cél, hogy lefedjék az explicitáció típusainak nagy részét, de a teljesség igénye nélkül; semmi esetre sem képviselik az explicitációs fordítási stratégia teljes palettáját.

2. táblázat

A párhuzamos korpuszban (AMK) észlelt explicitációs stratégiák összefoglalása

Szint	Eltolódás	Megjegyzés (oka/jellege)
1. logikai-vizuális kapcsolatok szintje	1. központozás: írásjelbetoldás és írásjelváltás 2. 1 m* → 2 m, 2 m →1 m 3. magyarázó kötőszavak: pl. *azaz*	tudatos stratégia és/vagy egyéni/közösségi stílus
2. lexiko-grammatikai kapcsolatok szintje	4. lexikai ismétlés 5. grammatikai párhuzamosságok 6. elliptikus szerkezetek kiegészítése 7. helyettesítések visszaállítása 8. angol névmás → magyar főnév	párhuzamos szerkezetek létrehozása
3. szókapcsolatok szintje	9. szófajváltók I.: *lévő, való*** 10. szófajváltók II.: *közötti, belüli*	a FNY és a CNY nyelvtipológiai különbségeiből eredő betoldások
4. mondatok szintje	11. kötőszavak betoldása 12. kataforikus referencia + kötőszó betoldása	a FNY és a CNY nyelvtipológiai különbségeiből eredő betoldások (pl. mondattömörítő szerkezetek feloldása), tudatos stratégia (pl. a FNY-i implicit viszonyok explicitté tétele)
5. szövegen és nyelven kívüli elemek szintje	13. lexikai magyarázat 14. diskurzusszervező elemek 15. szituatív betoldások 16. reáliák hozzáadott információval	tudatos stratégia, nyelvi/műfaji konvenciók

*m – mondat

** magyarázatért lásd: 4.1.3

Elegendőnek bizonyulnak azonban arra, hogy megteremtsék a második elemzés alapját, illetve különböző nyelvi szinteken bemutassák az eltolódások változatosságát. Az alábbiakban válogatott példák következnek a stratégiákra, világosan szemléltetve, hogy a feltárt stratégiák minden egyes típusának részletes vizsgálata és bemutatása túlmutatna a jelen tanulmány keretein.

4.1.1. Eltolódások a logikai-vizuális kapcsolatok szintjén: írásjelek

A szövegszerkezet logikai-vizuális kapcsolatai szintjén az írásjelek eltolódásai közé tartozik a) az írásjelbetoldás és b) az írásjelek helyettesítése egy erőteljesebb írásjellel. Az előbbihez az alábbiakban láthatunk példát.

3. táblázat

Az írásjelek gyakorisága a párhuzamos korpuszban (AMK)

Műfaj	**Szöveg**	**Kettőspont**	**Pontosvessző**	**Zárójel**	Összesen
		A – M	A – M	A – M	A – M
SZ	CA	11 – 14	14 – 4	1 – 1	26 – 19
SZ	UP	2 – 7	11 – 14	1 – 1	14 – 22
SZ	PA	0 – 4	1 – 9	3 – 2	4 – 15
SZ	ME	1 – 9	2 – 7	0 – 0	3 – 16
Összesen		14 – 34	28 – 34	5 – 4	47 – 72
T	DA	10 – 25	10 – 7	4 – 7	24 – 39
T	EY	3 – 5	3 – 3	35 – 33	41 – 41
T	HA	2 – 11	0 – 5	10 – 10	12 – 26
T	HI	6 – 9	4 – 4	13 – 13	23 – 26
Összesen		21 – 0	17 – 19	62 – 63	100 – 132

(1a) <DAenT, S 83> Paley here appreciates the difference between natural physical objects like stones, and designed and manufactured objects like watches.

(1b) <DAhuT, S 83> Paley itt azokat a különbségeket értékeli, amelyek a természetes fizikai objektumok (*mint a kő*) és a megtervezett és elkészített dolgok (*mint az óra*) között fennállnak.

Azon kívül, hogy a fordító a plusz megjegyzéseket szintén a *mint* appozícióval vezeti be, a zárójelek beillesztésével gördülékenyebbé teszi a mondatokat. A párhuzamos korpuszban (AMK) előforduló írásjelek gyakoriságáról szóló összefoglalás (*3. táblázat*) megmutatja ezen jellemző gyakoriságának az általános emelkedését, ezenfelül pedig rámutat a fordítói stílusok különbözőségeire azt illetően, hogy a fordítók milyen mértékben őrzik vagy változtatják meg a forrásnyelvi szövegben előforduló írásjeleket. A gondolatjeleket a párbeszédek jelölésének lényeges különbségei miatt kizártam az elemzésből. Az angol nyelv ugyanis az idézőjelet, míg a magyar nyelv inkább a gondolatjelet részesíti előnyben.

Az írásjelek erőteljesebb irányba történő eltolódását „egy tudat alatti stratégia részeként" értelmezhetjük, amelynek célja, „hogy érthetőbb megfogalmazással egyszerűbbé és könnyebbé tegye a feldolgozást" (Baker 1996: 182). Az is lehetséges, hogy a fordítók végső célja az üzenet mondanivalójának világosabb és összefüggőbb megfogalmazása, ezért ennek a stratégiának a végeredménye az egyszerűbb és könnyebben olvasható szöveg.

4.1.2 Eltolódások a lexiko-grammatikai kapcsolatok szintjén

Az egyik kohéziós kötés a helyettesítés. A második példában az eltolódás két kohéziós eszköz között megy végbe, vagyis a helyettesítést lexikai ismétlés váltja fel. Habár a helyettesítés „egy olyan kohéziós eszköz, amely egy korábban megnevezett dologra utal" (Halliday és Hasan 1976: 90), a fordító nem támaszkodott az anaforikus utalásra, és – valószínűleg stilisztikai okok miatt – inkább egy erősebb kohéziós kötést választott helyette.

(2a) <HAenT, S 56> As far as Kepler was concerned, elliptical orbits were merely an ad hoc hypothesis, and a rather *repugnant one* at that, because ellipses were clearly less perfect than circles.

(2b) <HAhuT, S 56> Kepler az ellipszispályákat alkalmi hipotézisnek tekintette, méghozzá fölötte *visszataszító hipotézisnek*, mivel az ellipszis nyilvánvalóan tökéletlenebb a körnél.

A helyettesítések visszaállítása – vagyis főneves szerkezetre való cseréje – azonban nem minősül kötelezően elvégzendő műveletnek, amikor angolról magyarra fordítunk, ahogyan a harmadik példában is látható:

(3a) <HAenT, S 47> Then two astronomers-the German, Johannes Kepler, and the Italian, Galileo Galilei – started publicly to support the Copernican theory, despite the fact that the orbits it predicted did not quite match the ones observed.

(3b) <HAhuT, S 47> Két csillagász: a német Johannes Kepler és az olasz Galileo Galilei nyilvánosan támogatni kezdte ezt a világképet, annak ellenére, hogy a Kopernikusz által megjósolt pályák nem minden esetben feleltek meg a megfigyelteknek.

A lexiko-grammatikai kapcsolatok szintjén elvégzett stratégiák elemzése Halliday és Hasan (1976) kohéziós eszközöket kategorizáló tipológiáján alapul, a párhuzamos grammatikai struktúrák azon típusaival kiegészítve, amelyek a párhuzamos korpuszban azonosított példák közül többet is lefednek.

A megállapítások arra mutatnak, hogy az angol forrásnyelvi szövegekben megtalálható kohéziós eszközök minden típusában előfordulnak eltolódások. Ezeket a kohéziós eszközöket ugyanazon a szinten másfajta kohéziós kötések helyettesítik a magyar célnyelvi szövegekben. A legtöbb esetben lexikai ismétléshez, következésképpen pedig redundanciához (Blum-Kulka 1986) vezetnek az eltolódások, amelyek többek között elliptikus szerkezetek kiegészítését, helyettesítések visszaállítását, valamint névmások lexikalizálódását foglalják magukba. Ám vajon miért döntenek a fordítók gyakran egy másik típusú kohéziós kötés mellett?

A lexikai ismétlés sokat vitatott nyelvi eszköz. Egyrészt a fordítók igyekeznek elkerülni a lexikai ismétléseket, és ez a tendencia tulajdonképpen ugyancsak egy lehetséges fordítási univerzálénak tekinthető (Baker 1998: 288). Másrészt viszont, ahogyan az adatokból is kiderül, a fordítók bőségesen alkalmaznak lexikai ismétléseket, hogy megteremtsék vagy megerősítsék a kohéziót a célnyelvi szövegekben. Mivel elsődleges céljuk, hogy érthető és világos célnyelvi mondatot hozzanak létre, alkalmazott stratégiáik felülírhatják az egyébként tiszteletben tartott fordítási normát, vagyis az ismétléskerülést. Ez a fordítási jelenség azonban megmagyarázható azzal a ténnyel, miszerint „a kohézió a nyelv rendszerének a része (…), és beépült magába a nyelvbe" (Halliday és Hasan 1976: 5).

4.1.3. Szintaktikai szintű eltolódások: szófajváltók

A jelzői bővítmények szerkezeti különbségéből azt a következtetést vonhatjuk le, hogy az angol jobbra, míg a magyar inkább balra bővít. A magyarra fordított szövegekben gyakran használunk olyan szavakat, amelyek a főnévi szerkezetben lehetővé teszik a balra történő bővítést. A melléknévi igenevek (a szemantikailag gyakran üres *való* vagy *lévő*), a félig üres melléknévi igenevek (*végzett*) vagy a névutómelléknevek (*közötti* vagy *belüli*) gyakran ezt a célt szolgálják. A negyedik példa egy szemantikailag üres melléknévi igenév használatát mutatja be:

(4a) <EYenT, S 8> In a fairly recent survey of academic psychologists in America, it was found that over three-quarters of them claimed to be cognitive psychologists!

(4b) <EYhuT, S 8> Egy amerikai egyetemi pszichológusok között végzett újabb felmérésben azt találták, hogy több mint háromnegyed részük kognitív pszichológusnak tekintette magát.

A helyhatározói, időhatározói és egyéb határozói jelentést hordozó magyar névutókat (például fák *között*) képzővel láthatjuk el, így névutómelléknévvé válnak, mint

például *között – közötti* (közötti – [között] + [i] – [névutó] + [melléknévképző]). „A névutómelléknevek a legfiatalabb csoport a magyar szófaji rendszerben, és a szintetikus nyelvi szerkezetre jellemzők, míg a névutó az analitikus nyelvi szerkezet tipikus példája” (Mátai 2002: 84).

(5a) <HIenT, S 2> The term Regional Economic Association (REA) defines collectively the various forms of economic integration among independent states.

(5b) <HIhuT, S 2> A regionális gazdasági társulás (REA) a független államok közötti gazdasági integráció különböző formáinak együttes definíciója.

4.1.4. Szintaktikai szintű eltolódások: kötőszavak

(6a) <HIenT, S 8> The dynamics of integration arise from increasing openness and political and economic interdependence among the participating countries.

(6b) <HIhuT, S 8> Az integráció dinamikája a részt vevő országok növekvő nyitottságának, illetve egymástól való kölcsönös gazdasági és politikai függésének eredménye.

A hatodik példában látható eltolódás két mellérendelő kötőszó között megy végbe. Az illetve kötőszó megváltoztatja a mellérendelt elemek felsorolásának tagolását: a fordító az *openness and political and economic interdependence* helyett a következő kapcsolatot valósította meg: *nyitottságának, illetve egymástól való kölcsönös gazdasági és politikai függésének eredménye*. A mellérendelő kötőszavak eltolódásának valószínűleg két oka lehetséges. Először is, hogy az *és* kötőszót többféleképpen értelmezhetjük:

> Szemantikailag a kapcsolatot egy kohéziós skálán helyezhetjük el... Az összes kötőszó közül az *és* kötőszó jelentése a legáltalánosabb – nevezhetnénk „általános célú kapcsolóelemnek", mivel csupán annyit jelent, hogy két gondolat között pozitív kapcsolat van, azt pedig már az olvasónak kell eldöntenie, mi is ez a pozitív kapcsolat. (Leech és Short 1989, idézi Øverås 1998: 576)

Másodszor, motivációként feltételezhetjük, hogy a magyar megfogalmazás x + (y + z) struktúráját valószínűleg könnyebb megérteni, mint az angol x + y + z struktúrát.

4.1.5. Szintaktikai szintű eltolódások: kötőszavak és kataforikus referenciák

A magyar vonatkozói mellékmondat és a hogy-os tagmondat egyik jellemzője a főmondatban bevezető névmással kifejezett kataforikus referencia. Ugyan nem képezi szükségszerűen részét a mondatnak, de „jelenlétével sokkal teljesebbnek érezzük a mondatot" (Hell 1980: 157). Az alábbiakban a hetedik példa egy vonatkozói mellékmondatot, míg a nyolcadik egy hogy-os alárendelő tagmondatot illusztrál:

(7a) <EYenT, S 80> If that is the case, then science may have practically no special features which elevate it above ancient myths or voodoo.

(7b) <EYhuT, S 80> Ha ez így van, akkor a tudománynak gyakorlatilag nincs semmi *olyan* jellemzője, *mely* az ősi mítoszok vagy mágiák fölé emelhetné.

(8a) <DAenT, S 19> The reader's reaction to this may be to ask, 'Yes, but are they really biological objects?

(8b) <DAhuT, S 19> Az olvasó reakciója valószínűleg *az* lesz, *hogy* megkérdezi: "Rendben, de vajon tényleg biológiai objektumok ezek?

4.1.6. Szövegszintű eltolódások: kultúraspecifikus elemek

Amennyiben a háttérismeret különbözik két nyelv, kultúra vagy kontextus között, a fordító mellékelhet egy olyan rövidebb vagy hosszabb magyarázó megjegyzést a fordításhoz, mint az alábbiakban látható *amerikai*, amely a kilencedik példamondatban a könyvkiadó nemzetségi hovatartozását jelöli, és a *Kiadó* szó, amely értelmezőként szolgál.

(9a) <PAenF, S 7> A dozen years ago, a senior man from Knopf recognized his former prison guard inside the well-pressed suit of a Heibon-sha executive, stood staring at him for a moment or two, then threw his champagne into the startled Japanese face.

(9b) <PAhuF, S 7> Tíz-tizenkét éve történt, hogy az *amerikai* Knopf egyik vezető munkatársa felismerte a Heibon-sha *Kiadó* valamelyik igazgatójában azt a hajdani őrt – pedig jól-szabott öltönyt viselt –, aki annyit gyötörte valamikor a hadifogolytáborban; egy darabig csak állt és némán nézte, aztán pezsgőjét a megdöbbent japán arcába löttyintette.

A párhuzamos korpuszban azonosított explicitációs stratégiák széles skálája betekintést nyújt a fordítói folyamatba, amely közben számos eltolódás megy végbe különböző tényezők és motivációk hatására: ezek közé tartozik a fordítók tudatos vagy nem tudatos szövegalkotási stratégiája, a fordítók vagy a nyelvi közösség sajátos stílusa, a műfaji hagyományok vagy a fordítói normák.

4.2. Eltolódások az explicitségben

A második elemzés célja az MMK fordított és autentikus szövegeinek összehasonlítása. A *4. táblázat* az összehasonlítható korpusz alapján azt mutatja be, hogy az explicitséget befolyásoló szövegjellemzők milyen gyakorisággal fordulnak elő, és hogyan oszlanak el. A legfontosabb összehasonlítások az MMK eredetileg magyar nyelven keletkezett (E) és

fordítás útján létrejött (F) magyar nyelvű szövegei között lelhetők fel. Az adatok szerint az esetek 80 százalékában – azaz 20 esetből 16 alkalommal – a fordításokban gyakoribbak ezek a szövegjellemzők, mint az autentikus szövegekben. A legtöbbször előforduló különbség a szófajváltók – mint a *közötti* és a *belüli* – esetében mutatkozik meg, mivel az autentikus szövegekben nem fordulnak elő, míg a fordított szövegekben 21 példát találtam rájuk.

Csupán négy eset mond ellent a hipotézisnek: a *való*, *amely*, *tehát* és *is* (*4. táblázat*). A legfeltűnőbb eredményeket a *való* üres melléknévi igenév és a *tehát* következtetésre utaló kötőszó használata mutatja. Az eredeti szövegekben 20 esetben fordul elő a *való*, míg a fordított szövegekben csak 8, a *tehát* esetében az arány 18:7. A *való* szócska, a *van* létige melléknévi igeneve, fontos szintaktikai funkciót tölt be: lehetővé teszi a balra történő jelzői bővítést. Mivel a magyar és az angol jelzői bővítmények között szerkezeti különbségek mutatkoznak (lásd *2. függelék*), azt feltételezhetnénk, hogy a *való* gyakrabban fordul elő az indoeurópai nyelvekről fordított szövegekben, mint az autentikus magyar szövegekben. Más szavakkal, nem várnánk az autentikus szövegek szerzőitől, hogy ezeket az eszközöket gyakrabban használják a fordítóknál, akik a forrásszöveg vagy maga a fordítási folyamat vagy mindkettő korlátai közé kényszerítve dolgoznak; ennek ellenére a szerzők is használják őket, hiszen egy kivételével az összesre találtam példát a tudományos szövegekben. Feltételezhetően tehát létezhetnek olyan normák, amelyek az említett eszközök használatát szabályozzák a szakszövegírók munkájában.

A korábbi szakirodalmi nézeteknek erősen ellentmond, hogy a *való* gyakrabban fordul elő a nem fordított szövegekben, mint a fordított szövegekben. Ez a szokatlan minta az *amely* vonatkozó névmásra, valamint az *is* diskurzus partikulára és kapcsoló bővítményre egyaránt vonatkozik, de természetesen további, nagyobb korpuszokon végzett elemzésre is szükség van, hogy ez a megállapítás igazolást nyerhessen.

Érthető tehát, hogy nem vagy nem kizárólag a fordítók használnak eddig a fordításokra jellemzőnek hitt nyelvi elemeket a tudományos szövegekben. Ezek a jelenségek a műszaki szövegek szerzőinek céljaival magyarázhatók, akik a lehető legtöbb információval szeretnék megtölteni a szövegeket, és – tudatosan, vagy sem – arra törekszenek, hogy

a szöveg minél érthetőbb legyen. Ugyanakkor még valószínűbbnek tűnik, hogy a már létező, fordított célnyelvi szövegek vannak rájuk hatással.

A két vizsgált műfajban megmutatkozó explicitséget illetően magasabb előfordulási gyakoriságot állapíthatunk meg a tudományos szövegekben, mint az irodalmi szövegekben. Habár az esetek csupán 65 százalékában (20 esetből 3 alkalommal) erősíthető meg a hipotézis, a műfajváltók és kötőszavak csoportja teljes mértékben alátámasztja a 3. hipotézist, míg a diskurzus partikulák csoportja ellentmond neki. A műfajokra jellemző explicitség részletesebb vizsgálatára van szükség, és erre a korpuszok minden bizonnyal alkalmasak.

Összefoglalásképpen azt a következtetést vonhatjuk le, hogy az MMK gyakorisági adatai bizonyítják azt a feltevést, miszerint a fordított magyar szövegek nagyobb mennyiségben tartalmaznak explicitséget, mint a nem fordított szövegek (2. hipotézis). Ez azt is jelentheti, hogy az explicitáció valószínűleg a fordított szövegek egy általános tulajdonsága, azaz az adatok sora Blum-Kulka (1986) hipotézisét támasztják alá.

4.3. Szótípus/szövegszó arány az összehasonlítható korpuszban

A szótípus/szövegszó arány a szöveg felszínén megmutatkozó lexikai változatosság mutatója. A *szövegszó* kifejezés a szövegben előforduló összes szóra utal, míg a *szótípus* kifejezés a szövegben szereplő eltérő szóalakokat jelenti. Minél magasabb a százalékos arány, annál változatosabb a szókincs (Baker 1995; Munday 1998). A szótípus/szövegszó arány nagyon érzékeny a szöveg hosszára, ezért az elemzések során a standardizált szótípus/szövegszó arányt használtam, mert bár az ARRABONA korpusz szövegei ugyanannyi mondatot tartalmaznak, terjedelmük meglehetősen eltérő.

4. táblázat

Az explicitség szövegjellemzőinek gyakorisága az összehasonlítható korpuszban (MMK)

Szint	Sorszám	Csoportkód	Szövegjellemző	Szépirodalom			Tudományos szövegek			Autentikus	Fordított
				EMK	FMK	Σ	EMK	FMK	Σ	E	F
1.	1	K[1*]1	kettőspont	21	34	55	56	50	106	77	84
	2	K2	pontosvessző	1	34	35	10	19	29	11	53
	3	K4	zárójel	21	4	25	30	63	93	51	67
3.	4	SZV*1	közötti, belüli	0	4	4	0	17	17	0	21
	5	SZV2	való	1	0	1	19	8	27	20	8
4.	6	KSZ*1	hogy	80	121	201	86	176	262	166	297
	7	KSZ2	aki*	12	29	41	8	12	20	20	41
	8	KSZ3	ami*	30	38	68	28	39	67	58	77
	9	KSZ4	amely*	7	13	20	63	48	111	70	61
	10	KSZ5	pedig	6	8	14	13	18	31	19	26
	11	KSZ6	azonban	5	7	12	14	20	34	19	27
4.	12	KU*1	az*…, hogy	5	26	31	28	54	82	33	80
	13	KU2	arr*…, hogy	2	4	6	11	21	32	13	25
	14	KU3	ann*…, hogy	2	2	4	2	9	11	4	11
5.	15	DP*1	csak	18	37	55	18	30	48	36	103
	16	DP2	még	24	30	54	10	23	33	34	53
	17	DP3	is	59	59	118	51	33	84	110	92
	18	DP4	például	1	1	2	10	19	29	11	20
	19	DP5	így	11	23	34	18	16	34	29	39
	20	DP6	tehát	5	0	5	13	7	20	18	7
	21	SZ/SZ*	szótípus/ szövegszó	65,73	61,75	átlag	60,84	54,54	átlag	63,29	58,15
	22					63,74			57,69		

Az 5. táblázatban megfigyelhető, hogy a számítógépes statisztikai elemzés eredményei két egyértelmű tendenciát mutatnak. Először is, az összehasonlítható korpuszban szereplő fordítások alacsonyabb szótípus/szövegszó arányt tartalmaznak, mint az autentikus szövegek (58,15–63,29). Ez pedig azt a konklúziót vonja maga után, hogy a fordított szövegek szókincse kevésbé változatos, mint az autentikus szövegeké.

Továbbá az is látható, hogy az összehasonlítható korpuszban megtalálható tudományos szövegek alacsonyabb szótípus/szövegszó arányt mutatnak, mint a szépirodalmi szövegek (57,69–63,74). Ez a jelenség pedig arra enged következtetni, hogy a tudományos szövegekben megmutatkozó szókincs kevésbé változatos, mint a szépirodalmi szövegeké.

5. táblázat

Szótípus/szövegszó arány az összehasonlítható korpuszban

	Alkorpusz	Autentikus magyar	Fordított magyar	*Átlag*
1.	Szépirodalom	65,73	61,75	63,74
2.	Tudományos szövegek	60,84	54,54	57,69
	Átlag	63,29	58,15	
	Különbség	4,89	7,21	

Ezen a ponton szeretnék a megvizsgált műfajokra is kitérni. Az autentikus szövegek szótípus/szövegszó aránya 5 százalékos különbséget mutat a két vizsgált műfaj között, míg ez a különbség 7 százalék a fordított szövegek esetében. Ezen paraméterek konvergenciája azt sugallhatja – ahogyan már a kutatás kezdetekor is felvetettük –, hogy a műszaki szövegek fordítói a forrásnyelvi szövegben szereplő információk minél pontosabb és tisztább közvetítése érdekében elkerülhetetlenül is gyakrabban alkalmaznak explicitációs stratégiákat, mint a szépirodalmi szövegek fordítói. Így az explicitációs stratégiák lexikai ismétlésekhez, következésképpen kevésbé változatos szókincshez vezetnek. Más szóval ez a sajátosság jól visszatükrözi azt a normát, amely a műfaj elvárásait jellemzi.

5. Konklúzió

A jelen tanulmányban leírt kutatás célja az volt, hogy próbára tegye az explicitációs hipotézist és megvizsgálja, hogy a fordításokban magasabb szintű explicitség fedezhető-e fel, mint az autentikus szövegekben.

Ha közelebbről szemügyre vesszük a két vizsgált nyelv strukturális különbségeit (az agglutináló magyar a jelentés kifejezésére kevesebb szót használ, mint az analitikus angol, például *I love you* → *Szeretlek*), az angolról magyarra való fordítások esetében inkább implicitációra (általánosabb fogalmazás, a forrásnyelvi szöveg nyelvi és nem nyelvi információinak kihagyása) számíthatunk, mintsem explicitációra. Habár a párhuzamos korpusz alapján létrehozott 16 explicitációs stratégiával az explicitáció az angol–magyar fordítási irányban erős tendenciának tűnik.

A második elemzés eredményei magasabb szintű explicitséget mutatnak az egész összehasonlítható korpuszban. A legtöbb eredmény alátámasztja azt a hipotézist, miszerint a fordítások explicitsége magasabb, mint az autentikus szövegeké. Továbbá, minden adatot egybevetve arra következtethetünk, hogy a harmadik hipotézist el kellene vetni. Az elemzés nem adott semmilyen választ a műfajok különbözőségére vonatkozó kérdésre. A bemutatott eredmények alapján nem állíthatjuk, hogy egyértelmű különbség van a szépirodalmi és tudományos szövegek között, ami a jelen tanulmányban megvizsgált nyelvi elemeket illeti.

A szótípus/szövegszó arány kérdésében azt látom, hogy a fordított szövegek esetében az explicitációs stratégiák következményeképp alacsonyabbak a százalékok. A logikai–vizuális kapcsolatok és a szövegen és nyelven kívüli elemek szintjeinek elmozdulásaitól függetlenül (lásd 2. táblázat) minden elmozdulás elkerülhetetlenül lexikai ismétlésekhez, és így a szókincs leegyszerűsítéséhez vezet. Például a kötőszavak, kataforikus referenciák + kötőszók betoldása és a szófajváltók használata közvetlenül növeli az ismételt elemek, vagyis a szövegszók számát, és így a szótípusok száma alacsonyabb lesz. Az elliptikus szerkezetek kiegészítése, a helyettesítések visszaállítása és a névmások főnévvel való helyettesítése is hozzájárul a szókincs alacsonyabb szintű változatosságához.

Ezért megállapítható, hogy a fordításban az explicitáció jelensége szorosan kapcsolódik az egyszerűsítéshez.

Összefoglalva a párhuzamos és összehasonlítható korpuszok elemzési eredményeit kijelenthetjük, hogy az 1969 és 1999 közötti időszakban olyan fordítási norma volt érvényben, amely szerint a fordítók a célszöveg szabályaihoz próbáltak igazodni, és így a célnyelvi olvasók igényeit kielégíteni. Összességében ez az explicitációs stratégiák végső funkciója.

1. függelék

Angol – magyar szövegek párhuzamos korpusza

Szépirodalmi szövegek

Kód	Író	Cím	Fordító	Cím
UP	Updike, John 1977	Marry me	Göncz Árpád 1981	Gyere hozzám feleségül
CA	Capote, Truman 1980	Music for Chameleons. Mojave	Osztovics Levente 1982	Mozart és a kaméleonok. Mojave
PA	Porter, Anna	The Book-fair Murders	Bart István 1998	Gyilkosság a könyvvásáron
ME	McEwan, Ian 1999	Amsterdam	Tandori Dezső 1999	Amszterdam

Tudományos szövegek

Kód	Író	Cím	Fordító	Cím
DA	Dawkins, Richard 1986	The blind watchmaker	Szentesi István (1. fejezet) 1994	A vak órásmester. Gondolatok a darwini evolúcióelméletről
HA	Hawking, Stephen 1988	A brief history of time: from the big bang to black holes	Molnár István 1989	Az idő rövid története
HI	Hitiris, Theo	European Community Economics	Roboz András 1995	Az Európai Unió gazdaságtana

EY	Eysenck, Michael W., Keane, Mark T. 1990	Cognitive Psychology. A Student's Handbook	Bocz András 1997	Kognitív pszichológia. Hallgatói kézikönyv

Összehasonlítható korpusz magyarul

Szépirodalmi szövegek

Kód	Író – fordító	Cím	Kód	Író	Cím
UP	Updike, J. 1977; Göncz Á. 1981	Gyere hozzám feleségül	KO	Konrád György 1969	A látogató. Budapest: Magvető Kiadó
CA	Capote, T. 1980; Osztovics L. 1982	Mozart és a kaméleonok. Mojave	MM	Mészöly Miklós 1984	Megbocsátás. Budapest: Szépirodalmi Kiadó
PA	Porter, A. 1997; Bart I. 1998	Gyilkosság a könyvvásáron	EP	Esterházy Péter 1990	Hrabal könyve. Budapest: Magvető Kiadó
ME	McEwan, I. 1999; Tandori D. 1999	Amszterdam	PO	Polcz Alaine 1991	Asszony a fronton. Budapest: Szépirodalmi Kiadó

Tudományos szövegek

Kód	Író – fordító	Cím	Kód	Író	Cím
DA	Dawkins, R. 1986; Szentesi I. 1994	A vak órásmester. Gondolatok a darwini evolúcióról	SI	Simonyi Károly 1978/1998	A fizika kultúrtörténete. Budapest: Akadémiai Kiadó
HA	Hawking, S. 1988; Molnár I. 1989	Az idő rövid története	BE	Bekker Zsuzsa 1978	Növekedési utak – dinamikus ágak. Budapest: Közgazdasági és J. K.
HI	Hitiris, Th. 1995; Roboz A. 1995	Az Európai Unió gazdaságtana	KJ	Kornai János 1983	Ellentmondások és dilemmák. Budapest: Magvető Kiadó
EZ	Eysenck, M. W., Keane, M. T. 1990; Bocz András 1997	Kognitív pszichológia. Hallgatói kézikönyv	KF	Kozma Ferenc 1992	A menedzser közgazdasági szemlélete. Budapest: Közgazdasági és J. K.

2. függelék

Magyar és angol főneves szerkezetek balra és jobbra történő bővítése

tágas	szoba			spacious room	
háromlábú	szék			three-legged chair	
esernyős	férfi			man	with an umbrella
nagyértékű	könyv			book	of great value
barnaöltönyös	férfi			man	in a/the brown suit
baloldalt lévő	ajtó			door	on the left
ablak mellett álló	lány			girl	(standing) by the window
művelésre alkalmas	föld			land	suitable for cultivation
egyezmény elérésére tett	erőfeszítések			efforts	(made) to reach an agreement
az az	ember,	aki éppen most érkezett		the man	who has just arrived

(Heltai és Pinczés 1993: 55)

Irodalom

Baker, M. 1993. Corpus Linguistics and Translation Studies: Implications and Applications. In: Baker, M., Francis, G., Tognini-Bonell, E. (eds) *Text and Technology: in Honour of John Sinclair*. Amsterdam/Philadelphia: John Benjamins. 233–250.

Baker, M. 1995. Corpora in Translation Studies: An Overview and Suggestions for Future Research. *Target* Vol. 7 No. 2. 223–243.

Baker, M. 1996. Corpus-based translation studies: The challenges that lie ahead. In: Somers, H. (ed) *Terminology, LSP and Translation: Studies in language engineering in honour of Juan C. Sager*. Amsterdam/Philadelphia: John Benjamins. 175–186.

Baker, M. 1998. *Routledge Encyclopedia of Translation Studies*. London/New York: Routledge.

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 Studies*. Tübingen: Gunter Narr Verlag. 17–35.

Esterházy P. 1996. *Egy kék haris*. Budapest: Magvető Könyvkiadó.

Halliday, M. A. K., Hasan, R. 1976. *Cohesion in English.* London/New York: Longman.

Hell, Gy. 1980. Object and subject *that* clauses in English and corresponding *hogy* clauses in Hungarian. In: Desző, L., Nemser, W. (eds) *Studies in English and Hungarian Contrastive Linguistics*. Budapest: Akadémiai Kiadó. 145–175.

Heltai P., Pinczés É. 1993. *Fordítás és szövegértés a nyelvvizsgán*. Budapest: Druck Kft.

Ishikawa, L. 1998. *Cognitive Explicitation in Simultaneous Interpreting*. Cambridge: RCEAL University.

Kenny, D. 1999. *Norms and Creativity: Lexis in Translated Text.* PhD Thesis. Manchester: UMIST. Centre for Translation Studies.

Klaudy, K. 1993a. Optional additions in translation. In: *Translation the vital Link, Proceedings of the XIII. FITWorld Congress* Vol. 2. London: ITI. 373–381.

Klaudy, K. 1993b. On explicitation hypothesis. In: Klaudy, K., Kohn, J. (eds) *Transferre necesse est… Current issues of Translation Theory. In honour of Radó on his 80th birthday*. Szombathely: Berzsenyi Dániel Főiskola. 69–77.

Klaudy, K. 1996. Back Translation as a Tool for Detecting Explicitation Strategies in Translation. In: Klaudy, K., Lambert, J., Sohár, A. (eds) *Translation Studies in Hungary*. Budapest: Scholastica. 99–114.

Laviosa-Braithwaite, S. 1996. *The English Comparable Corpus (EEC): A Resource and Methodology for the Empirical Study of Translation.* PhD Thesis. Manchester: UMIST. Centre for Translation Studies.

Laviosa, S. 1998. The corpus-based approach. *Meta* Vol. 43 No. 4. 473–659.

Leech, G. N., Short, M. H. 1981. *Style in Fiction. A Linguistic Approach to English Fictional prose*. London.

Mátai M. 2002. A névutók és a névutómelléknevek története. *Magyar Nyelvőr* 126. *évf*. 1. szám. 72–87.

Munday, J. 1998. A computer-assisted approach to the analysis of translation Shifts. *Meta* Vol. 43. No. 4. 542–556.

Olohan, M., Baker, M. 2000. Reporting *that* in Translated English. Evidence for Subconscious Processes of Explicitation? *Across Languages and Cultures* Vol. 1 No. 2. 141–158.

Øverås, L. 1998. In Search of the Third Code. An Investigation of Norms in Literary Translation. *Meta* Vol. 43 No. 4. 571–588.

Scott, M. 1998. *WordSmith Tools.* Oxford: Oxford University Press.

Séguinot, C. 1988. Pragmatics and the Explicitation Hypothesis. *TTR Traduction, Terminologie, Rédaction* Vol. 1. No. 2. 106–114.

Shlesinger, M. 1995. Shifts in Cohesion in Simultaneous Interpreting. *The Translator* Vol. 1. No. 2. 193–214.

Sinclair, J. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Toury, G. 1995. *Descriptive Translation Studies and beyond*. Amsterdam: John Benjamins.

Vanderauwera, R. 1985. *Dutch Novels Translated into English: The Transformation of a 'Minority' Literature*. Amsterdam: Rodopi.

Weissbrod, R. 1992. Explicitation in translations of prose-fiction from English to Hebrew as a function of norms. *Multilingua* Vol. 11. No. 2. 153–171.

(Footnotes)

1 * K – központozás, SZV – szófajváltók, KSZ – kötőszók, KU – kataforikus referenciák, DP – diskurzuspartikulák, SZ/SZ – szótípus/szövegszó

Az árnyékkal jelölt kapcsolatok elvetik a hipotéziseket.

Idegen szavak félautomatikus elemzése autentikus és fordított orvosi szövegekben: fordítási stratégiák angolról franciára és magyarra fordított betegtájékoztatók tükrében

Mány Dániel

manydaniel91@gmail.com

ELTE Nyelvtudományi Doktori Iskola

Fordítástudományi Doktori Program

***Cervical dystonia** (or **spasmodic torticollis**) is*

*the most common form of focal or localized **dystonia**.*

***Neurological** in origin, it is characterized by muscular **spasms**…*

(Ipsen weboldal 2018)

Kivonat: A jelen leíró jellegű, feltáró tanulmány középpontjában az idegen szavak jelensége áll, azon belül pedig ezen orvostudományi terminusok előfordulása és fordítása autentikus angol, valamint angolról magyarra és franciára fordított betegtájékoztatók párhuzamos korpuszában. Az idegen szavak gyakorisága nyelvenként eltérő, ezért a fordítónak szem előtt kell tartania a laikus célnyelvi olvasó feldolgozási erőfeszítését. A vizsgált korpuszban az angolra jellemző a legkevésbé az idegen szavak használata. Az angolhoz képest több idegen szó azonosítható a francia alkorpuszban, a legtöbb idegen szó a magyar alkorpuszban van. Mindez főként a vizsgált nyelvek szókészleti különbségeivel magyarázható. Az elemzés során összesen 5+1 fordítási stratégiát állapítottam meg, amelyek tekintetében a fordító törekszik az idegen szavak fordításakor megfelelni a laikus célnyelvi olvasó

Az Innovációs és Technológiai Minisztérium ÚNKP-20-4 kódszámú Új Nemzeti Kiválóság Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott szakmai támogatásával készült.

Mány Dániel: Idegen szavak félautomatikus elemzése autentikus és fordított orvosi szövegekben: fordítási stratégiák angolról franciára és magyarra fordított betegtájékoztatók tükrében. In: Robin E., Seidl-Péch O. (szerk.) 2020. *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordításkutatás Magyarországon.* Segédkönyvek a nyelvi közvetítésről I. Budapest: ELTE BTK Fordítástudományi Doktori Program, MANYE Fordítástudományi Szakosztály. DOI: https://doi.org/10.36252/Nyelvikozvsegedkonyv1.6

és az egészségügyi személyzet elvárásainak. Jelen tanulmány a 2018-as TransELTE Konferenciakötetben megjelenő cikkem bővített változata.

Kulcsszavak: betegtájékoztatók, fordítási stratégiák, idegen szavak, laikus célközönség, párhuzamos korpusz

1. Bevezetés

Jelen leíró jellegű, feltáró tanulmány középpontjában a fenti mottóban is fellelhető idegenítő stratégia áll, azon belül pedig ezen orvostudományi terminusok és/vagy szinonimáik előfordulása autentikus angol, valamint angolról magyarra és franciára fordított betegtájékoztatókban.

Kutatásomat az motiválta, hogy bár számos kutató (pl.: Bősze 2011, Heltai 2004, Jiménez-Crespo és Tercedor Sánchez 2017, stb.) foglalkozott az idegen szavak, a tudományos terminusok és a köznyelvi szavak nyelv- és fordítástudományi vonatkozásaival az általam vizsgált három nyelv mindegyikén, tudomásom szerint mindeddig nem került sor az angolról magyarra és franciára fordított szövegek átfogó, összevető vizsgálatára az idegen szavak gyakorisága, eloszlása és a rájuk jellemző fordítási stratégiák tükrében.

Ahogy azt a továbbiakban bemutatom, az idegen szavak előfordulásukat, gyakoriságukat és eredetüket tekintve nyelvközösségekként, kultúrákként, műfajokként, szövegtípusokként eltérők lehetnek, még az orvosi szaknyelven belül is. Különösen érdemesnek tartom az idegen szavak vizsgálatát forrásnyelvi szövegeken (FNYSZ), illetve azok összevetését célnyelvi szövegekkel (CNYSZ), ahol a fordító feladata pontosan a nyelvek és kultúrák közötti közvetítés. Korábbi vizsgálatok rávilágítanak, hogy az orvosi szakfordítás a fordítói piac egyik legjelentősebb ágazata (Montalt Resurrecció és González Davies 2007), s ennek köszönhetően a fordítástudomány kedvelt kutatási területe (Montalt Resurrecció 2011). Éppen ezért úgy gondolom, hogy kiemelkedően fontos a területen végzett leíró kutatások folyamatos szorgalmazása. Zethsen (2004: 139) rámutat, hogy bár hipotézisének alátámasztására további vizsgálatok szükségesek, a latin eredetű terminusok direkt átvitelre angolról dánra orvos–laikus kommunikációban jelentősen megnehezítheti az olvasó feldolgozási erőfeszítését, sőt, érthetetlenné is teheti a szöveget. Jelen tanulmány

igyekszik szűkíteni az említett kutatási űrt, annál is inkább, mert a korábban végzett, a jelen tanulmány szempontjából releváns vizsgálatok (Nisbeth Jensen és Zethsen 2012, Jiménez-Crespo és Tercedor Sánchez 2017) ellentmondásos eredményekről számolnak be. Tanulmányom eredményeit összehasonlítom az imént említett, korábbi vizsgálatok eredményeivel.

A betegtájékoztatók olyan orvosok által írt, laikusoknak szóló dokumentumok, amelyek jó lehetőséget adnak annak vizsgálatára, hogy az orvos–laikus közötti írott kommunikáció szövegeinél milyen fordítói stratégiák (pl.: explicitáció, implicitáció, orvostudományi terminus kerülése stb.) figyelhetők meg. A jelen tanulmány célja ezen stratégiák azonosítása és leírása. A stratégiák leírásának, tipizálásának célja gyakorló fordítók, fordítást oktató tanárok, nyelvtanulók segítése, valamint új információ gyűjtése az idegen eredetű terminusok fordítási stratégiáiról a már meglévő ismeretek, korábbi kutatások eredményeinek tükrében.

Fontos kutatási kérdés számomra, hogy a fordító hogyan jár el az idegen eredetű orvostudományi terminusok fordításakor. Legalább ilyen fontosnak tartom, hogy az idegen szavaknak más stilisztikai értéke van, mint a köznyelvi szavaknak. Ennek az orvosi szaknyelvben azért van kiemelkedő szerepe, mert az egészségügyi szakdolgozók a kommunikáció során elkerülési mechanizmusokat alkalmaznak. Amikor általuk kényesnek, nehéznek ítélt tényekről kell tájékoztatniuk a betegeket (Cselovszkyné Tarr 1999), eufemisztikus hatást kelthetnek, ha az általuk is ismert köznyelvi szó helyett tudományos terminust használnak. A vizsgálat további célja annak leírása, hogy az idegen eredetű, eufemisztikus hatást keltő tudományos terminusok hogyan jelentkeznek a forrásnyelvi és a célnyelvi szövegben.

2. Elméleti háttér

Mint minden nyelvben, sőt, mint minden szaknyelvben, az angol, francia és magyar orvosi szaknyelvben is vannak idegen szavak. A népek, népcsoportok folyamatosan érintkeznek egymással, és az új fogalmak, tárgyak megnevezésére idegen eredetű szavakat vesznek

és adnak át. A nem belső keletkezésű szóállományban megkülönböztetjük egymástól a jövevényszó vagy kölcsönszó és az idegen szó fogalmát. Az idegen szó és a jövevényszó közötti különbség abban rejlik, hogy míg az idegen szó idegenszerűségét a nyelvközösség legtöbb tagja érzékeli, a jövevényszó nem idegenszerű, az átvevő nyelv rendszerébe beilleszkedett, a nyelvközösség számára nem idegen hatású. Természetesen az idegen szavak idővel ugyanúgy jövevényszavakká válhatnak (Fazekas 2007): bár a *vírus, doktor, fisztula, diftéria* szavak mind latin eredetűek, véleményem szerint az első kettő nem hat idegenszerűen a nyelvközösség döntő többségének.

2.1. Idegen szavak, tudományos terminusok és köznyelvi szavak

Jelen tanulmány központi kérdése az általam vizsgált orvosi szaknyelvekben megjelenő idegen szavak, tehát a laikus olvasó számára feltehetőleg idegenszerűleg ható terminusok jelensége. Úgy gondolom, hogy a vizsgált nyelvekben a görög, latin és angol (LGA) eredetű idegen szavakat érdemes vizsgálni. Ennek oka, hogy az egyetemes orvostudomány az ókori görög, majd a görög–latinná alakult szakkifejezések tárán alapszik, és erre támaszkodik ma is (Bősze 2011: 370–371). A 20. század második felében az angol vált a nemzetközi közös nyelvvé. A tudományos ismeretek folyamatos bővülése nap mint nap új terminusokat, kifejezéseket szül. Az orvostudomány lingua francája ma egyértelműen az angol, amely ugyanazt a szerepet tölti be a tudományos kommunikációban, mint az ókorban a görög, a középkorban a latin (Varga 2014: 35). Az újonnan keletkezett terminusokat angolul nevezik el, amelyek beépülnek a nemzetek tudományos nyelveibe (Bősze 2011: 373), így a francia és magyar orvosi szaknyelvbe is. Az orvosi nyelv a többi tudományterület szaknyelvéhez hasonlóan folyamatosan új, angol eredetű terminusokkal bővül (Putz 2010: 370). A legkonzervatívabb orvosi területnek tartott tudományág, az anatómia szaknyelvében is felfedezhető az angol folyamatos térhódítása (Varga 2014: 41). Elemzésemet a latin és a görög (LG) terminusok vizsgálatán túl az angol eredetű terminusokra is

kiterjesztem, hiszen nem kaphatunk teljes képet az adott szövegek jellemzőiről az idegen szavak tükrében, ha az elemzés nem számol be az angol eredetű terminusokról.

Nem meglepő tehát, hogy gyakran találkozunk magyar orvosi szövegekben görög–latin (pl.: *krónikus, karcinóma, anatómia, hematoxilin*) és angol (pl.: *biopszia, skin-sparing mastectomia, graft*) terminusokkal. Érdekes jelenség, amikor az idegen szavak más kifejezések tagjaként fordulnak elő, tehát például az angol, a magyar és a latin szavak egy fogalmon belül keverednek (pl.: *laparoszkópos gastric band* műtét, in-*lay implantatio, diagnosztikus és staging laparoszkópia*). Ilyenfajta hibrid kölcsönzésről, egyesek által *részfordításnak* nevezett jelenségről akkor beszélhetünk, ha az átadó nyelv legalább egy morfémáját közvetlenül átváltjuk (Lanstyák 2006: 26).

Ha szembe állítjuk egymással az idegen eredetű és a nem idegen eredetű terminusokat, a tudományos terminusok és a köznyelvi szavak közötti különbségre gondolhatunk. Szinonimákként tekinthetünk a *pulmonológus* és a *tüdőgyógyász* szavakra, de ezek egyértelműen más regiszterhez tartoznak. A *pulmonológus* idegen eredetű szakszó, terminus, míg a *tüdőgyógyász* lehet terminusértékű, de nem idegenszerű a nyelvközösség legtöbb tagja számára.

Ezt azért fontos tisztázni, mert a terminusokat szokás megkülönböztetni a köznyelvi szavaktól, de fontos leszögezni, hogy nincs éles határvonal a két fogalom között. Annál is inkább, mert lényegében bármely köznyelvi szó terminussá válhat egy adott beszédközösség számára, vagy akár egy adott szövegben is. A különbség a tudományos és a hétköznapi megismerés különbségeire vezethető vissza, hiszen a tudományos megismerési és megnevezési folyamat részletesebb, túlmegy a mindennapos megismerésen, és a valóság olyan szegletét osztályozza, amelyet a köznyelv egyáltalán nem, vagy csak kevésbé részletesen. A hétköznapi megismerés célja a gyakorlati hasznosság, míg a tudomány magyarázatot keres, pontosságra törekszik. A fő különbség, hogy a terminológia más szempont szerint osztályozza a valóságot, mindig a megismerő funkció kerül előtérbe, a megismerés túlmegy a hétköznapi tapasztalat körén. A köznyelvi szavak esetében a referenciális jelentés implicit, meghatározhatatlan, jelentős az egyéni variabilitás, jelentésük függ a kontextustól, fontos szerepet játszik a szinonímia és a poliszémia. Ezzel szemben a ter-

minusok referenciális jelentése kidolgozott fogalomrendszerre vonatkozik, nem jellemzi egyéni variabilitás, adott szakterületen belül független a kontextustól, nem játszik jelentős szerepet a poliszémia, a szinonímia (Heltai 2004: 29).

Az orvosi szaknyelv terminusainak jelentős része idegen eredetű, de az orvos–beteg kommunikációban kiemelten fontos, hogy ezeknek a tudományos terminusoknak hatása van az olvasóra. Az eufemizmus neves kutatói (Tournier 1985, Demers 1991, Warren 1992) szerint például az idegen szavak eufemizáló jellegűek, megnehezítve ezzel az olvasó feldolgozási erőfeszítését. Fontos kérdés, hogy mennyire jellemző az angol forrásnyelvű szövegekre a tudományos terminusok szinonimákkal való helyettesítése, illetve, hogy milyen irányban változik ennek aránya a magyar és francia fordításokban (ha változik). A fordítók igyekeznek fordításaikat a célnyelvi olvasó igényeihez igazítani, ezért felmerül a kérdés, hogy ha ezek a terminusok eufemisztikus hatást keltenek, tehát elhomályosítják a jelölőt, hogy ne nevezzék meg a nemkívánatosnak ítélt referenst, akkor a fordító milyen tudatos vagy tudattalan stratégiákkal igyekszik a szóban forgó terminusokat megfeleltetni. A következő alfejezetben áttekintem az idegen szavak és a fordítás viszonyát.

2.2. Idegen szavak a fordításban

Ahogy azt az előző fejezetben bemutattam, az orvosi szaknyelv több diasztratikus változata ismert, de az érthetőség, a kommunikációs cél megvalósításának módja az orvos–laikus kommunikációban eltérő lehet az orvos–orvos kommunikációban tapasztaltakétól. Ez az eltérő regiszter eltérő terminológiát eredményez (Wright 2011: 246), amelynek egyik legjellemzőbb vonása a görög és latin terminusok eltérő eloszlása, valamint azok explicitációja, az orvostudományi terminus kerülése, körülírása, a regiszter „lesüllyesztése” (Meyer és Mackintosh 2000).

Ahogy azt később bemutatom, az orvostudományi terminusok kerülésének stratégiája nemcsak interlingvális, de intralingvális fordításokban is megfigyelhető és kutatásra érdemes téma. Különösen érdekes akkor, ha a kommunikáció nyelvek és kultúrák közötti közvetítéssel, interlingvális fordítással egészül ki, hiszen különböző nyelvekben és kul-

túrákban eltérő lehet az idegen szavak eloszlása, gyakorisága. Előfordul például, hogy a FNYSZ két terminust szinonimaként használ, a fordítónak pedig meg kell győződnie arról, hogy terminológiai szinonímiáról van szó (Heltai 2004: 39–41).

Lehetséges, hogy egyes orvosi LG terminusok az orvos–laikus kommunikációban sem okoznak jelentős feldolgozási erőfeszítést angolul vagy franciául, magyarul azonban túl magas regiszterre utalnának (pl.: *appendicitis 'vakbélgyulladás'*). LG terminusok gyakran fordulnak elő orvosi szövegekben, de nyelvekként eltérő lehet a terminusok gyakorisága, illetve az LG terminusok vernakuláris változatainak száma egyik vagy másik nyelven. Sokszor előfordul, hogy az angol és a francia LG terminust használ olyan helyen, ahol a magyar nyelv kerüli az LG terminus használatát a laikusoknak szánt szövegekben, hiszen a latin és görög elemek nem gyakoriak ebben a nyelvben. Ez főként azzal magyarázható, hogy a nemzeti orvosi nyelvek történetében lényeges különbségek észlelhetők a terminológiai jellemzőkben. A francia, angol *dermatologiste, dermatologist* helyes megfelelője ilyen esetekben a *bőrgyógyász*.

Újlatin nyelveken tehát alacsony regiszterűnek tűnhet egy-egy olyan terminus, amely jól megállja a helyét angol tudományos szövegekben (Montalt Resurrecció és González Davies 2007: 242). A fordításban ilyenkor felmerül az orvostudományi terminus kerülése (*determinologization*), az átalakítás (*reformulation*) és az explicitáció lehetősége (Jiménez-Crespo és Tercedor Sánchez 2017: 407). A következő alfejezetben áttekintem a jelen tanulmány szempontjából releváns, korábbi kutatások eredményeit.

2.3. Korábbi kutatások eredményei

Muñoz-Miquel (2012) intralingvális fordításon végzett a jelen tanulmány szempontjából releváns vizsgálatot. Korpuszát tíz darab angol nyelvű, laikusoknak szóló összefoglalóból állította össze. A tudományos összefoglalókat orvosok által orvosoknak írt cikkek alapján tömörítették, fogalmazták át. Az intralingvális fordítás során a szelekciós folyamat része, hogy eltávolítsuk a szöveg olvasója szempontjából irrelevánsnak vagy túl nehéznek tartott tartalmakat. Gyakori, hogy hozzáadunk releváns, fontos információkat a kulcsfo-

galmakhoz, hogy az olvasó számára a szöveg explicitebb, érthetőbb legyen, hiszen az olvasó és a szöveg írója feltehetőleg eltérő ismeretekkel rendelkeznek. Az elemzés szerint ez a stratégia leginkább a szövegek elején jellemző. A korpusz intralingvális fordításaiban a következőképpen jártak el az összefoglalások írói a tudományos terminusok esetén (Muñoz-Miquel 2012: 200–202):

– a tudományos terminus változatlan marad, de magyarázat vagy definíció van a terminus előtt vagy után (általánosságban a magyarázat zárójelben van, és megelőzi a terminust, de ez nem törvényszerű):

(1) More recently, researchers have suggested using tadalafil (a medicine often used to treat erectile dysfunction),

– a terminust idézőjelbe tett köznyelvi szinonima kíséri:

(2) *herpes zoster* (also called *"shingles"*),

– az eredeti szövegben nem szereplő egységeket vezetnek be az olvasó segítésének céljából (ez akár a *-like* szócska is lehet, pl.: *cortisone-like drug* known as dexamethasone)

(3) People commonly use nonsteroidal anti-inflammatory drugs (NSAIDs) to relieve pain. *Examples of NSAIDs include aspirin, etodolac (Lodine), ibuprofen (Advil or Motrin), and naproxen (Aleve).*

– a tudományos terminus kiesik, és egy kollokviálisabb „pszeudo-ekvivalens" helyettesíti:

(4a) myocardial infarction,

(4b) heart attack,

– perifrázis használata:

(5a) second-line treatment options,

(5b) options for other drugs they could take,

– változás a mondathosszban és a morfoszintaktikai szerkezetben:

(6a) Study personnel who reconstituted the vials and inoculated the participants,

(6b) Staff injecting the solution under the skin.

Az elemzés rámutat, hogy a legtöbb esetben megőrizzük a tudományos terminusokat, de gyakran zárójelben, idézőjelben szereplő magyarázatokkal, szinonimákkal konkretizáljuk őket. Időnként előfordul, hogy a tudományos terminust pszeudo-ekvivalensekkel, perifrázisokkal helyettesítjük. Az elvont, absztrakt fogalmakat átfogalmazzuk, konkretizáljuk. A redundanciát növelve az összetett fogalmakat ismételjük, szinonimákat használunk. A szerző fontos és hasznos felosztást ad az idegen szavak lehetséges intralingvális átalakítási eljárásairól (*reformulation procedures*), de a különböző eljárások gyakoriságát nem számszerűsíti. Ahogy a tanulmány végén pontosítja, a korpusz kis méretű, és szövegei egy forrásból származnak, ezért további vizsgálatok szükségesek az eredmények alátámasztására. A tanulmány rámutat, hogy a stratégiák megfigyelése és leírása fontos feladat a további kutatásokban (Muñoz-Miquel 2012: 204). Figyelemre méltó, hogy az LG terminusok fordítására még intralingvális fordításon belül is van kutatási igény. Bár a bemutatott vizsgálat nem biztosít összehasonlíthatóságot az eredmények számszerűsítésének hiánya miatt, megerősíti a témában folyó kutatások fontosságát, és érdekes, hasznos, a jelen tanulmányban is felhasználható felosztást javasol az LG terminusok fordítási stratégiáinak leírására.

Zethsen (2004: 126) az angolról dánra fordított, laikusoknak szánt orvosi fordítások feldolgozási erőfeszítését rendkívül nehéznek tartja, hiszen a célnyelvi olvasó szakismeretek hiányában nem tudja, vagy nehezen tudja értelmezni a latin terminológia direkt átvitelét a fordításokban. A dán nyelvközösség számára vagy érthetetlen, vagy túl szakszerű egy-egy latin eredetű terminus, de egy átlagos angol beszélő számára semmiféle problémát nem okoz annak feldolgozása. Éppen ezért, ha figyelembe vesszük a fordítás szkoposzát, akkor sok esetben a latin terminusok direkt átvitele nem eredményez ek-

vivalencia-viszonyt a forrásnyelv és a célnyelv latin eredetű terminusa között (Zethsen 2004: 138), mert az angol és a dán laikus olvasók ismeretei eltérők az idegen eredetű szavak tekintetében, az egyes idegen eredetű terminusok a két nyelvben más regiszterhez tartoznak. Ez azzal magyarázható, hogy az angol és a dán köznyelvben alapvetően eltér a latin és görög eredetű szavak gyakorisága. Ahogy a magyarról is elmondható, az angolhoz képest a dán is kevesebb LG terminust használ a mindennapi nyelvhasználatban, és ha az angolról dánra történő fordításban az LG terminusokat magyarázat vagy köznyelvi megfelelő betoldása nélkül átváltjuk, a fordított szöveg nehézsége drámaian növekszik.

Mindezt Nisbeth Jensen és Zethsen (2012) évekkel későbbi kutatása is alátámasztja, ahol az LG terminusok eloszlását angol és dán nyelvű szövegekben hasonlítják össze. Vizsgálatukban különösen fontos, hogy az LG terminusok fordítását tipizálják: két fő kategóriát és számos alkategóriát különböztetnek meg (dán nyelvismeret hiányában a példák ismertetésétől eltekintek):

– laikusokat segítő megoldások (*lay-friendly option*):

  - köznyelvi szó használata LG terminus helyett,
  - a sorrend változtatása: a dán terminus megelőzi az eredeti szövegben is szereplő LG terminust,
  - laikusok igényéhez igazodó magyarázat vagy terminus betoldása az LG terminus megtartásával,

– laikusokat nem segítő megoldások (*non-lay-friendly option*):

  - ha az adott LG terminusnak nincs köznyelvi megfelelője, a fordító átváltja a forrásnyelvi LG terminust, és nem fűz hozzá magyarázatot [pl.: *polycystic ovarian syndrome (PCOS)*],
  - a forrásnyelvi LG terminus és a köznyelvi szó vagy magyarázat fordítása változtatás nélkül,

- a fordító megtartja az LG terminust, pedig létezik köznyelvi dán megfelelő is,
- a fordító betold LG terminust, pedig létezik köznyelvi dán megfelelő is.

A kutatók azt vizsgálták, hogy az LG terminusok tekintetében milyen különbségek és hasonlóságok írhatók le attól függően, hogy egészségügyi szakemberek, vagy profi fordítók tollából származik a fordítás. Az eredmények azt mutatják, hogy az egészségügyi szakemberek több LG terminust használnak a fordításban, és az is gyakoribb, hogy nem fűznek hozzá magyarázatot. Akkor is hajlamosak LG terminust használni, ha van köznyelvi dán megfelelő. Jelen tanulmány nem tesz különbséget az egészségügyi szakemberek és a (szak)fordítók által fordított szövegek között, hiszen az elemzett szövegek mindegyike mindenki számára elérhető: az adott egészségügyi intézmény által publikált, a fordítók által lefordított és a megrendelő által elfogadott, betegek számára elérhető dokumentumok. Érdekesnek tartom azonban a laikusokat segítő és nem segítő felosztást, illetve a kutató által bemutatott stratégiákat.

Montalt Resurrecció és González Davies (2007: 251–253) szerint a tudományos terminus kerülésének az orvosi szaknyelvben négy alkategóriája különíthető el. A négy alkategória akár egyszerre is megvalósulhat:

– a célnyelven megmarad a tudományos terminus, de magyarázat követi

(7a) most *dyskinesias* are due to basal ganglia disorders…

(7b) most *dyskinesias* (impairment of voluntary movements resulting in fragmented or jerky motions) are due to basal ganglia disorders…

– a célnyelven megmarad a tudományos terminus, de csak zárójelben szerepel

(8a) Rigidity progresses, and *bradykinesia, hypokinesia*, and *akinesia* appear.

(8b) Rigidity progresses, and movements becomes slow (*bradykinesia*), decreased (*hypokinesia*), and difficult to initiate (*akinesia*).

– a tudományos terminust köznyelvi szóval helyettesítjük

(9a) Large *haemorrhages*, when located in the hemispheres, produces *hemiparesis*.

(9b) Intense bleeding, when located in the hemispheres, produces paralysis on one side of the body.

– a tudományos terminusokat teljesen kerüljük, és magyarázattal helyettesítjük

(10a) Most *dyskinesias* are due to basal ganglia disorders…

(10b) most impairment of voluntary movements is due to basal ganglia disorders…

Ezek a stratégiák könnyebben értelmezhetővé teszik a fordított szövegeket a laikus olvasó számára. Közülük az explicitáció, az a fordítási művelet, amelynek során a fordító nyíltabban, világosabban, esetleg több szóval fejez ki valamit a CNYSZ-ben, mint ahogy azt a FNYSZ szerzője tette (Klaudy 2001: 371). Ezen a ponton kell megjegyezni, hogy az explicitáció önmagában jellemző a fordított szövegekre, valamint az orvos és laikus közötti kommunikációra.

Jiménez-Crespo és Tercedor Sánchez (2017: 421) szerint a fordítási univerzálék újra tanulmányozása szakszövegen, sajátos műfajokon, regisztereken fontos feladat, hiszen az adott nyelvkombináció, regiszter és műfaj tükrében különbségeket állapíthatunk meg az eredmények alapján. A szerzők az összehasonlítható korpusznyelvészet módszertanával 40 millió szövegszavas korpuszt tanulmányoztak: angolról spanyolra fordított, illetve eredetileg spanyolul írt, betegeknek szóló internetes oldalakat vizsgáltak. A VariMed adatbázisból véletlenszerűen tizenhárom LG terminust választottak ki. A tanulmány rávilágít, hogy az LG terminusok átfogalmazási stratégiáinak (*reformulation strategies*) lehetősége szélesebb körű, mint ahogy az az adatbázisban szerepel, és vannak, amelyek

közül néhány kizárólag fordított szövegekben azonosítható. Ez lehet interferencia eredménye, vagy a lexikalizált forrásnyelvi elemek szó szerinti átváltása.

A WordSmith Tools 6.0 segítségével a korpuszon mindegyik LG terminus relatív gyakoriságát mérték. Az elemzés második részében azt vizsgálták, hogy a weboldalak írói mennyire tartják szem előtt a laikus olvasó igényeit (*lay-friendliness*): mennyire és milyen formában jellemző az LG terminusok átfogalmazása, a determinologizálás Montalt Resurrecció és González Davies (2007) korábban bemutatott terminuskerülési stratégiáinak tükrében. Az elemzés második részét manuálisan végezték a WordSmith Tools 6.0 konkordancia mezőinek vizsgálatával, az előfordulást normalizálták.

A vizsgálat eredményei kimutatták, hogy a két alkorpuszt összehasonlítva több LG terminus azonosítható az autentikus spanyol szövegekben, mint a fordított spanyol szövegekben. A fordított szövegben lévő LG terminusok alacsony hányada ellentmond az angol-dán nyelvpár esetében végzett kutatás eredményeinek (Nisbeth Jensen és Zethsen 2012), ahol az LG terminusok direkt átvitele érthetetlenné tette a dán szöveget a laikus célközönség számára. Ezzel szemben a laikusoknak szánt spanyol orvosi dokumentumokban gyakrabban szerepelnek LG terminusok, mint az autentikus angol szövegekben. Az eredmények szerint a fordításokban gyakrabban figyelhető meg átfogalmazás, explicitáció, mint az autentikus szövegekben, mert a FNYSZ szerzője determinologizációra törekszik, hogy könnyítse az angol nyelvű olvasó szövegértését, és a fordító is igyekszik a célközönség igényeit figyelembe véve az LG terminusokat explicitálni, átfogalmazni.

Elméleti síkon több szempontból is megközelíthető, hogy pontosan mi történik a fordítás folyamatában az idegen nyelvű szavakkal, kifejezésekkel. Ha a spanyol fordításokban kevesebb idegen eredetű terminus szerepel, mint az autentikus spanyol szövegben, akkor a fordító Baker (1996) terminusával élve normalizációra törekszik, a célnyelv jellemzőit igyekszik kiemelni. Kenny (1998: 516) ezt úgy közelíti meg, hogy a fordító igyekszik minél jobban adaptálni a CNYSZ-et a célközönség igényeihez, ezért a fordítás eredményeképpen keletkezett szöveget „megtisztítja" (*sanitisation*), hogy elfogadhatóbbá tegye azt a célnyelvi olvasó számára. Előfordul, hogy az eredeti szöveget összehasonlítva a fordítással úgy érezzük, hogy a fordítás eltompítja a FNY-i szöveg esetleges ridegségét,

negatív töltetét. A fordításban tehát semlegesebb elemek jelenhetnek meg a forrásnyelv negatív szemantikai prozódiájú elemeinek megfeleltetésekor, és bár nehéz rájönni, hogy mitől válik a CNYSZ mássá, a végleges hatás az, hogy a fordítás az eredeti „megtisztított" változata lesz. Kenny (1998) szerint érdemes lenne lexikai elemek szemantikai prozódiájának mérése FNYSZ-en és fordításokon, ehhez szükség lenne párhuzamos, illetve forrásnyelvi és célnyelvi referenciakorpuszra. A tanulmányában hozott példák nem reprezentatívak, de érdekes empirikus kutatási irányokat vetnek fel, és rámutatnak a párhuzamos és egynyelvű összehasonlítható korpuszok használatának fontosságára.

Az idegen szavak hatását megközelíthetjük az *eufemizmus* jelenségén keresztül is, ahogy arra korábban utaltam. Egy kommunikációs helyzet akármelyik résztvevője akármikor érezheti úgy, hogy mondandóját szépíteni, enyhíteni kell. Mindez függ egyrészt attól, hogy az adott korszak társadalmi, kulturális és nyelvi normái milyen fogalomkörök esetében sugallják, hogy a szavak élét tompítani kellene. Ugyanakkor egy egyén által eufemisztikusnak ítélt egységet nem feltétlenül tart annak egy másik is: az eufemizmus tanulmányozása, elemzése tehát sosem lehet teljesen mentes a szubjektivitástól, de szubjektív jellege ellenére is levonhatók általános következtetések, megfogalmazhatók objektív jellegzetességek. Az eufemizmus elhomályosítja a jelölőt, mert lehetővé válik, hogy az információ közlője ne nevezze meg a nemkívánatosnak ítélt referenst, de maga a referens valójában változatlan. Az eufemizmus során tehát nem a fogalom vagy a jelölt változik, hanem maga a nyelvi jel. Voltaképpen egy olyan új nyelvi jelről van szó, amely egy másik nyelvi jel elhomályosított változata. A jelölttel és a fogalommal azonban ugyanolyan kapcsolata van. Az eufemizmus tanulmányozásának kérdése tehát az, hogy milyen kapcsolat van az eufemisztikus nyelvi jel és a diszfemisztikus nyelvi jel között (Jamet 2010: 35). Mindezt azért tartom relevánsnak, mert az eufemizmus több kutatója (Tournier 1985, Demers 1991, Warren 1992) szerint az idegen szavaknak, terminusoknak eufemisztikus hatása van.

A kutatás célja az idegen szavak azonosítása, gyakoriságuk és a fordítási stratégiák kimutatása és leírása angolról magyarra és franciára fordított betegtájékoztatókban, a különbségek és hasonlóságok megfigyelése a különböző nyelvű szövegek között, a lehet-

séges eltolódások kimutatása a fordított szövegekben. Mindez rámutatna az orvosi szövegek fordításának mindeddig kevéssé kutatott jellegzetességeire, általánosíthatóságaira. A tanulságokat alkalmazni lehetne az orvosi szakfordítók munkájában, képzésében, illetve össze lehetne hasonlítani korábban végzett kutatások eredményeivel.

3. Módszerek

Két okból választottam elemzésem tárgyaként az orvosi szaknyelv egyik sajátos műfaját, a *betegtájékoztatót* az idegen szavak vizsgálatára. Az első, hogy a gyógyszerészeti, terápiás és orvosi kísérőiratok nyelvészeti elemzése a nyelvtudomány számára egyelőre feltáratlan terület (Illésné Kovács és Simigné Fenyő 2009: 141), annak ellenére, hogy napjainkban az egészségügyi felvilágosító irodalom a legolvasottabb szövegtípusok egyike (Tótfalusi 2008: 12). Fontosnak tartom, hogy a *betegtájékoztató* néven ismert műfajon kívül nem vonok be más műfajú orvosi dokumentumot a vizsgálatba, mert az eredmények műfajonként eltérők lehetnek (Jiménez-Crespo és Tercedor Sánchez 2017: 412). A műfaj angol nyelvű elnevezései: *patient information leaflet (PIL), package insert, information leaflet, patient package insert, consumer medicine information* (Montalt Resurrecció és González Davies 2007: 68).

3.1. A korpusz

Az orvosi szaknyelv egyik diasztratikus változatának tekintem az orvos–laikus kommunikációt. Az orvosi szaknyelv szakember és laikusok közötti nyelvhasználati rétegének jellemzője, hogy a kommunikáció résztvevői eltérő szakmai ismeretekkel rendelkeznek (Kurtán 2003: 33). A sikeres kommunikáció megvalósításához szükséges, hogy az orvosi szövegeket a szöveg szerzője és fordítója a laikus befogadó elvárásaihoz adaptálja. Az orvosi szakfordító fontos feladata, hogy az általa használt regiszter illeszkedjen a célközönség igényeihez (Jiménez-Crespo és Tercedor Sánchez 2017: 406), az autentikus és fordított szövegek regiszterbeli különbségének egyik legfontosabb alkotóeleme pedig az LG terminusok eloszlása, explicitációja vagy terminuskerülése, átfogalmazása (Meyer és

Mackintosh 2000). Éppen ezért a műfaj kiválasztására motiváló második ok, hogy ezek a közvetett, írott formájú, betegeknek és hozzátartozóiknak készített tájékoztatók jó lehetőséget adnak annak feltárására, hogy az idegen szavak milyen mennyiségben és formában (pl.: tudományos terminus kihagyása, körülírása, helyettesítése stb.) jelennek meg az eredetileg angolul írott betegtájékoztatókban és a magyar, francia fordításokban.

Az összegyűjtött szövegekre párhuzamos korpuszként (Baker 1993) tekintek, hiszen a korpusz angol nyelvű szövegeket és azok francia és magyar fordításait tartalmazza. Ezek a szövegek elérhetők a Pannónia Korpusz (Robin et al 2016) folyamatosan épülő orvosi alkorpuszában. A párhuzamos korpusz elemzése lehetővé teszi annak feltárását, hogy az eltolódások oka intralingvisztikai vagy interlingvisztikai explicitációban keresendő-e (Jiménez-Crespo és Tercedor Sánchez 2017: 420). A párhuzamos korpuszok hozzájárulnak, hogy az empirikus kutatások a leíró fordítástudományt szolgálják, és rávilágítanak arra, hogyan oldják meg a fordítók az egyes fordítási problémákat (Robin et al 2016), hogy milyen különbségek vannak az eredeti és a fordított szövegek között (Seidl-Péch 2018), tehát milyen tudatos vagy tudatalatti stratégiákat alkalmaz a fordító az LGA terminusok fordításakor. A párhuzamos korpusz rámutat a fordító műveleteire, stratégiáira, illetve azt is megmagyarázza, hogy miként befolyásolja a fordító megfogalmazását az eredeti FNYSZ.

Az elemzésben három alkorpuszt vizsgálok: autentikus angol, fordított magyar és fordított francia szövegeket. Mindegyik szöveg 2000 után született. A reprezentativitás szempontjából fontos, hogy a korpusz több szerző és fordító munkájából tevődik össze (Robin et al 2016).

Tematikájukat tekintve a szövegek egy része gyermeket vállaló szülőknek szól az esetleges genetikai betegségekről, illetve azok öröklődéséről, vizsgálatáról. A korpusz többi szövege a myeloma multiplex nevű rosszindulatú betegség kezeléséről, az őssejtterápiáról, szívelégtelenségről szól, illetve műtét előtti tájékoztató.

Az elemzéshez használt angol forrásnyelvű betegtájékoztatók, illetve azok francia és magyar fordításának terjedelmét, szószámát az *1. táblázat* foglalja össze. A kor-

pusz összesen 87 577 szövegszóból áll. Az angol forrásnyelvi alkorpusz mérete összesen 28 492 szövegszó, a francia fordítás összesen 33 301 szövegszóból áll, míg a magyar 25 784 szövegszóból. A fordított magyar és a fordított francia alkorpuszok közötti terjedelmi különbség azért nem befolyásolja a vizsgálat eredményeit, mert mindkét célnyelvi alkorpusz ugyanannak az angol forrásnyelvi alkorpusznak a fordítása.

1. táblázat

A párhuzamos korpusz szószáma

A betegtájékoztató angol címe	Az angol FNYSZ terjedelme (szószám)	A magyar CNYSZ terjedelme (szószám)	A francia CNYSZ terjedelme (szószám)
The Amniocentesis	1 366	1 149	1 360
Dominant Inheritance	925	959	1 171
Recessive Inheritance	978	896	1 212
Frequently Asked Questions	619	531	602
Stem Cell Therapies	3 340	2 797	4 171
Chromosome Translocations	1 401	1 249	1 537
What is a Genetic Test?	1 284	1 156	1 373
Multiple Myeloma, Cancer of the Bone Marrow	13 583	12 394	16 512
Carrier Testing	3 371	2 367	3 558
Chromosome Changes	1 625	22 86	1 805
Összesen	**28 492**	**25 784**	**33 301**

3.2. Kutatási kérdések

Ahogy azt korábban bemutattam, a jelen kutatás központi kérdése az idegen szavak előfordulása angolról magyarra és franciára fordított betegtájékoztatókban. Az idegen szavak eredetét, gyakoriságát, az előfordulásukra jellemző hasonlóságokat és különbségeket, a fordítási stratégiákat a bemutatott korpusz alapján a következő konkrét kutatási kérdések vizsgálatával elemzem:

(1) Milyen gyakorisággal fordulnak elő idegen szavak az autentikus angol és a fordított magyar és francia orvosi szövegekben?

(2) Milyen fordítási stratégiák azonosíthatók az idegen szavak fordításakor az adott nyelvek/nyelvpárok tükrében, hogyan alkalmazkodik a fordítói a laikusok célnyelvi olvasó igényeihez?

3.3. Hipotézis

A kutatási kérdések tekintetében a kutatás kvantitatív részére a következő hipotézis állítható fel:

> Az angol autentikus szövegekben kevesebb LG terminus van, mint a fordított magyar és francia szövegekben a vizsgált nyelvek szókészletének különbségei miatt.

3.4. Az elemzés menete

Az elemzéshez elengedhetetlen, hogy a pdf, html és word formátumban letölthető fájlokat UTF-8 kódolású txt formátumba konvertáljuk, mivel az internetről ingyenesen letölthető és használható AntConc nevű program csak txt fájl importálására alkalmas. Az elemzés következő lépésében az AntConc 3.5.0. segítségével lehívtam a gyakorisági listákat. A kinyert szavakat Excel formátumú dokumentumba másoltam, és kitöröltem az egyértelműen nem idegen eredetű szavakat, és kigyűjtöttem azokat, amelyek LGA terminusok lehetnek. Ennek a lépésnek köszönhetően kiszűrhetők a magas gyakoriságú névelők, kötőszók, elöljárószók és az egyértelműen köznyelvi szavak, amelyek kétséget kizáróan a mindennapos nyelvhasználathoz, a köznyelvhez tartoznak. Az elemzést lényegesen megkönnyíti, hogy kitörlöm azokat a szavakat, amelyek egyértelműen nem idegen eredetűek (pl.: *of, that, health, skin, family; az, Ön, egészség, családtag, vizsgálat; santé, mais, effet, corps*).

Az elemzés folytatásában elkülönítettem azokat a terminusokat, amelyek bizonyosan idegen eredetű egészségtudományi terminusok vagy részben idegen eredetű, hibrid terminusok részei (pl.: *izomdisztrófia, őssejttranszplantáció, sejtproliferáció*). Ennek köszönhetően kinyertem azokat a terminusokat, amelyek kétségkívül a jelen elemzés tárgyát képezik. Ezt a lépést a *2. táblázattal* szemléltetem.

2. táblázat

Egyértelműen idegen eredetű terminusok kiválasztása

A betegtájékoztató angol címe	**Az angol FNYSZ terjedelme (szószám)**	**A magyar CNYSZ terjedelme (szószám)**	**A francia CNYSZ terjedelme (szószám)**
The Amniocentesis	1 366	1 149	1 360
Dominant Inheritance	925	959	1 171
Recessive Inheritance	978	896	1 212
Frequently Asked Questions	619	531	602
Stem Cell Therapies	3 340	2 797	4 171
Chromosome Translocations	1 401	1 249	1 537
What is a Genetic Test?	1 284	1 156	1 373
Multiple Myeloma, Cancer of the Bone Marrow	13 583	12 394	16 512
Carrier Testing	3 371	2 367	3 558
Chromosome Changes	1 625	22 86	1 805
Összesen	28 492	25 784	33 301

A további szelektálási folyamatot két tényező nehezítette meg. Egyrészt ahogy azt korábban bemutattam, nem vonható éles határvonal az idegen szavak és a jövevényszavak között. Szubjektív és akár egyénenként változó lehet, hogy az adott szót vagy terminust mennyire érezzük idegenítőnek. Másrészt úgy gondolom, hogy a szövegek elemzésekor akarva-akaratlanul hajlamosak lehetünk arra, hogy idegen szónak érzékeljünk olyan szavakat is, amelyeket egyébként talán nem tekintenénk annak. Az elemzés során érzékenyebbé, vagy éppen kevésbé érzékennyé válhatunk elemzésünk tárgyára.

Úgy küszöböltem ki a szubjektív faktort, hogy segítségül hívtam az angolul, franciául és magyarul írt idegen szavak szótárát. Ahogy azt az elemzéshez használt magyar szótár szerzője kiemeli az előszóban, a szótárban nem szerepelnek a nyelvbe már teljesen beolvadt idegen elemek, ellenben külön hangsúlyt fektet a közhasználatba mindinkább átkerülő tudományos szakszókincsre (Bakos 2015: 4–5). Bár a szubjektív jelleg a szótárállományra is jellemző, az elemzés objektív, mivel a kérdéses terminusok meglétét ellenőrzöm az idegen szavak szótárában. Ha ezen kérdéses elemek idegenszerűsége az elterjedt köznyelvi használat miatt sem egyértelmű, de szerepelnek a szótárban (pl.: *abnormális, kromoszóma, diagnosztika*), akkor elfogadom ezt az álláspontot, és idegen eredetű terminusként tekintek rájuk, tehát bevonom őket az elemzésbe. Ha egy terminus esetleges idegen csengése ellenére nem szerepel a szótárban (pl.: *herpesz, bakteriális, monitorozás*), akkor azt nem vonom be az elemzésbe.

Az elemzés tehát rávilágít, hogy milyen gyakorisággal fordulnak elő LGA terminusok az elemzett korpusz orvosi szövegeiben. Fontos azonban, hogy ne csak a gyakoriságot vizsgájuk, mert az AntConc nem szótővenként adja ki a találatokat, ezért a toldalékolt terminusok lentebb lesznek a listában, mint a toldalékolatlan változatok. Az elemzés folytatásában lemmatizálással azonosítottam az idegen szavak körében a típusokat a tokenekhez képest. Az AntConc segítségével lehívott gyakorisági listát Excel táblázatba másoltam, és a már bemutatott módszer segítségével kiszűrtem az idegen eredetű terminusokat. Az LGA terminusok gyakoriságát az Excelben összeadtam, így megkaptam az egyes betegtájékoztatókra és az egyes nyelvekre jellemző LGA-gyakoriságot.

A folytatásban összehasonlítottam a LGA terminusok gyakoriságát az angolról magyarra és az angolról franciára fordított szövegekben, valamint az angol FNYSZ-ben. Így előzetes képet nyerünk arról, hogy a vizsgált nyelvek között milyen különbségek vannak az LGA terminusok gyakoriságának tükrében. A különbségek azonban fakadhatnak interferenciából is, ezért a fordított magyar alkorpusz eredményeit a kutatás folytatásában az autentikus magyar alkorpusz eredményeivel fogom összehasonlítani. Ennek köszönhetően adatot nyerhetünk arról, hogy gyakoribb-e az autentikus magyar szövegekben a LGA

terminusok használata, mint az angolról magyarra fordított szövegekben, és további fordítói stratégiákra következtethetünk az adott műfaj és az adott nyelvkombináció tükrében.

Ezen a ponton tértem át a vizsgálat kvalitatív részére. Az elemzés következő lépésében a MemoQ szövegpárhuzamosító funkcióját használtam, amely lehetővé teszi, hogy megvizsgáljam adott terminusok környezetét. Ez azért fontos, mert arra számítottam, hogy a szerző bizonyos esetekben például zárójelben szerepelteti az LGA terminust, de a magyar fordító esetenként köznyelvi szót választott a forrásnyelvi LGA terminus megfeleltetésére. A párhuzamosítás azért is szükséges, mert nem biztos, hogy az LGA terminusokat a fordító megtartja, és szükséges megvizsgálnunk, hogy az adott angol szöveg idegen eredetű terminusa hogyan jelenik meg a magyar és a francia fordításban. A párhuzamosítással tehát kigyűjtöm az LGA terminusokra jellemző fordítási stratégiákat, majd beillesztem őket az elméleti háttérben bemutatott, korábbi kutatásokban javasolt stratégiák különböző kategóriáiba, vagy új kategóriákat állapítottam meg.

Elemzésemben továbbá arra törekedtem, hogy hasonlóságokat vagy különbségeket figyeljek meg az angolról franciára és magyarra fordítás stratégiájának tükrében az adott korpusz alapján. A tanulmány következő fejezetében rátérek a kutatási kérdések megválaszolására, a hipotézisem igazolására vagy elvetésére.

4. Eredmények

A jelen kutatás központi kérdése az idegen szavak jelensége, gyakorisága angolról magyarra és franciára fordított betegtájékoztatókban. Az eredmények tekintetében az első kutatási kérdésre vonatkozó adatokról számolok be.

4.1. LGA terminusok az egyes alkorpuszokban

Az angol FNYSZ-ekben mindösszesen 1 374 idegen eredetű latin és görög terminust azonosítottam az *Oxford Dictionnary of Foreign Words and Phrases* (Speake 1997) szótár segítségével. Ez a szám az idegen szavak gyakorisági vizsgálatához felhasznált angol for-

rásnyelvi alkorpusz teljes szószámának (28 492 szövegszó) 4,82%-a. Az idegen szavak gyakorisági vizsgálatához használt fordított magyar nyelvű alkorpuszban (25 784 szövegszó) összesen 2 182 idegen eredetű terminust azonosítottam, ami az alkorpusz teljes szószámának 8,46%-a. A fordított francia alkorpusz (33 301 szövegszó) összesen 2 005 idegen szót tartalmaz, ez az alkorpusz teljes szószámának 6,02%-a. Megállapítható tehát, hogy a vizsgált korpusz alapján az angolra jellemző legkevésbé az idegen szavak használata. Az angolhoz képest több idegen szó azonosítható a francia alkorpuszban, és a legtöbb idegen szó a magyar alkorpuszban van. Ezt a tendenciát nem csak az egyes alkorpuszokat összehasonlítva figyelhetjük meg, hanem akkor is, ha az egyes betegtájékoztatókat egymással összehasonlítjuk. A számadatokat a *3. táblázat* foglalja össze.

3. táblázat

A párhuzamos korpusz LGA terminusainak száma

A betegtájékoztató angol címe	Idegen szavak az autentikus angol alkorpuszban (szószám)	Idegen szavak a fordított magyar alkorpuszban (szószám)	Idegen szavak a fordított francia alkorpuszban (szószám)
The Amniocentesis	5	19	12
Dominant Inheritance	11	28	21
Recessive Inheritance	10	45	32
Frequently Asked Questions	1	6	2
Stem Cell Therapies	34	99	81
Chromosome Translocations	56	131	106
What is a Genetic Test?	17	30	22
Multiple Myeloma, Cancer of the Bone Marrow	1 144	1 548	1 495
Carrier Testing	53	122	112
Chromosome Changes	43	154	122

Ennek tekintetében igazolódott az (1) hipotézis, amelyben azt feltételeztem, hogy a fordított magyar és a fordított francia betegtájékoztatókban több az LGA terminusok száma, mint a hasonló terjedelmű autentikus angol betegtájékoztatókban. Ez azzal magyaráz-

ható, hogy a fordító igyekszik ezeket a terminusokat a laikus célközönség igényeihez igazítani, de a betegedukációnak mégis fontos része, hogy a betegek megismerkedjenek a betegséget jellemző, alapvető tudományos terminusokkal, hogy hatékonyan tudjanak kommunikálni az őket ápoló egészségügyi szakszemélyzettel. Az eredmények oka továbbá abban rejlik, hogy a magyar köznyelven alapvetően kevesebb a latin és görög eredetű terminusok aránya, mint az angolban vagy a franciában, ezért a magyar nyelvű célközönség idegenítőnek érez olyan LGA terminusokat, amelyek az angolban a köznyelvi beszélő számára ismertek, nem okoznak megnövekedett feldolgozási erőfeszítést. Ugyancsak fontos megjegyezni, hogy az angol és a francia betegtájékoztatókban azonosított angol eredetű terminusok értelemszerűen nem hatnak idegenszerűen az angol FNYSZ-ben, de a francia és a magyar fordításokban igen. Felmerül továbbá a kérdés, hogy az elemzéshez használt szótárak között milyen különbségek lehetnek, hogy a különböző nyelvek idegen szavainak szótárai azonos rendszerező elvet használnak-e. Utóbbi módszertani kérdést úgy fogom megválaszolni, hogy a jövőben angol, francia és magyar anyanyelvi beszélőkkel végeztetek hatásvizsgálatot.

A számszerű eredményekből továbbá az a következtetés vonható le, hogy a fent felsorolt okok miatt a célnyelven megnövekedett idegen szavak száma miatt az adott nyelvközösség számára a magyar és a francia szöveg idegenszerűbben hat, mint az angol FNYSZ. Mindez felveti a kérdést, hogy a fordító milyen stratégiákkal igyekszik kielégíteni a laikus célnyelvi olvasó igényeit, hogy mit tesz, hogy csökkentse az idegen szavakkal találkozó olvasó feldolgozási erőfeszítését. Ennek megválaszolásához azt is fontos megállapítani, hogy milyen különbségek vannak az idegen szavak gyakoriságának tekintetében autentikus magyar és fordított magyar szövegek között. Erre egy másik kutatásban vállalkozom.

4.2. Az LGA terminusokra jellemző fordítási stratégiák

Felmerül tehát a kérdés, hogy milyen stratégiákkal törekszik a fordító kielégíteni a laikus célközönség igényeit? Milyen szövegkörnyezetben figyelhető meg, hogy a fordító megtartja a forrásnyelvben is fellelhető idegen szavakat?

A kérdés megválaszolásához szükséges áttérnünk a kutatás további kérdéseire, a betegtájékoztatók fordítási stratégiáira, bár a kvantifikált eredmények tükrében arra lehet következtetni, hogy bizonyos esetekben a fordító az idegen szavak mellé magyarázatokat told be. A fordítási stratégiák elemzéséhez vegyük példaként a *mielómáról* szóló angol betegtájékoztató leggyakrabban előforduló terminusát. A *myeloma* terminus (244), annak többesszámú alakja *myelomas* (1), és terminológiai szinonimája *MM* (4) a legtöbbször előforduló terminus az angol forrásnyelvű tájékoztatóban. A magyar fordítás elemzésénél gondosan ügyelnünk kell arra, hogy a találati listát ábécé sorrendbe állítsuk a gyakorisági lista lehívása után. Ez azért fontos, mert az AntConc nem szótővenként adja ki a találatokat, ezért a toldalékolt terminusok lentebb lesznek a listában, mint a toldalékolatlan változatok. Nem meglepő, hogy a *mielóma* terminus és annak toldalékolt vagy szóösszevonással keletkezett változatai adják a magyar célnyelvi korpusz leggyakrabban előforduló (269) idegen eredetű terminusát. Ebből az adatból látható, hogy a magyar fordítás ebben az esetben is többször szerepelteti az adott idegen szót, mint az angol FNYSZ. Itt feltehetőleg arról van szó, hogy ahol az angol időnként névmással utal vissza a betegség megnevezésére, vagy ahol a többszörösen összetett mondatokban nem szükséges a betegség nevét többször említeni, a magyar fordító a könnyebb érthetőséget szem előtt tartva a főnevet használja.

Ezen a ponton merül fel a kérdés, hogy a fordító milyen stratégiákkal igyekszik alkalmazkodni a célnyelvi olvasó, a laikus befogadó igényeihez. A *mielóma multiplex* a csontvelő daganatos megbetegedése, ahogy az a tájékoztatóból is kiderül. A fordító ebben az esetben nem laikusbarát megoldást (*non-lay-friendly option*) választott (Nisbeth Jensen és Zethsen 2012 terminusa), ha a meglévő köznyelvi szinonima helyett a tudományos terminust választotta, aminek több oka is van.

Először is a fordítónak nemcsak a célnyelvi olvasóhoz kell igazodnia, hanem meg kell találnia az arany középutat a befogadó, de az őket gondozó egészségügyi szakszemélyzet és a FNYSZ-et író tudósok elvárásai között. A *mielóma multiplex* a betegség hivatalos elnevezése a Betegségek Nemzetközi Osztályozásában (BNO). Nem várhatjuk az orvosi szaknyelvektől, hogy mellőzzék az alapjaiban LGA gyökereken nyugvó termin-

ológiájukat, és nem helyettesíthetünk mindent tudományos terminust köznyelvi szóval. A *leukémia* helyett sem használhatjuk orvosi szövegekben a közérthető *fehérvérűség* szót, és nem mondhatjuk egy betegnek, hogy menjen a *vesegyógyászat*ra, ha az adott osztály hivatalos neve *nefrológia*.

A másik ok a fordítás folyamatában rejlik. A tanulmányban azt vizsgálom, hogy hogyan változnak, vagy maradnak változatlanok az idegen eredetű terminusok a fordítás folyamatában. De vajon hol kezdődik a fordítás folyamata? Ott, amikor a magyar vagy francia fordító kézbe veszi az angol FNYSZ-et? Ebben az esetben nem lehet szó erről, hiszen a betegtájékoztatók műfaji sajátosságaiból fakad, hogy már a FNYSZ keletkezését is feltételezhetjük intralingvális fordítás eredményének. Ezt a feltételezést a következő, a korpuszomban szereplő betegtájékoztatóból vett idézettel támasztom alá:

> Tájékoztató a gyulladásos betegeknek. A kérdéseket nemzetközileg vizsgált, *az érintett betegek által felvetett problémák alapján állítottuk össze*. A válaszokban a The Crohn's and Colitis Foundation of America, az Europien Crohn and Colitis Organisation és a Magyar Gasztroenterológiai Társaság Colon Szekciójának állásfoglalásait összegeztük. (Lakatos et al. 2005: fedőlap)

Ahogy az idézetből is kitűnik, ebben az esetben három különböző szakmai szervezet képviselői a laikus betegekben felmerülő kérdésekre válaszolva dolgozták ki a betegtájékoztatót. Az orvosi szaknyelv szakember és laikusok közötti nyelvhasználati rétegének jellemzője, hogy a kommunikáció résztvevői eltérő szakmai ismeretekkel rendelkeznek (Kurtán 2003: 33), ezért feltételezhető, hogy az anyag kiadása előtti megbeszéléseken, konzultációkon a szakemberek kommunikációjára más terminushasználat volt jellemző, mint amilyen az a kiadott anyagban azonosítható. Éppen ezért gondolom, hogy a fordítás folyamata valójában ott kezdődik, amikor az egészségügyi szakemberek közérthető módon igyekeznek megfogalmazni valamit laikusok számára az anyanyelvükön. Mindennek nyomai az autentikus és a fordított szövegeken is azonosíthatók. Az autentikus angol lai-

kusoknak szóló szövegben tehát alapvetően kevesebb LG terminus van, mint egy autentikus angol, nem laikusoknak szóló klinikai kutatásban.

Fontos feladat a terminusok lexikai környezetének vizsgálata, kiemelt figyelmet fordítva a terminus első előfordulására, hiszen Muñoz-Miquel (2012: 200–202) kutatása rávilágít, hogy leginkább a szöveg elején igyekszünk az olvasó szempontjából túl nehéznek ítélt terminus magyarázatára, explicitálására, illetve releváns információk hozzáadására, amelyek szintén tartalmazhatnak idegen szavakat, terminusokat. Ugyancsak a szöveg elején azonosított terminusokra a legjellemzőbb intralingvális fordítási stratégia, amikor a tudományos terminus változatlan marad, de magyarázat vagy definíció van a terminus előtt vagy után.

Betegtájékoztatók vizsgálatakor számos esetben már maga a cím is rávilágít arra, hogy a tudományos terminusokat már a forrásnyelven is igyekeznek definiálni, magyarázatokkal kísérni. A következő példák az elemzés párhuzamos korpuszában elemzett tájékoztató címe és annak magyar és francia fordításai.

(11a) Multiple myloma, *cancer of the bone marrow*

(11b) Mielóma multiplex, *a csontvelő daganatos betegsége*

(11c) Myélome multiple, *cancer de la moelle osseuse*

Egyértelmű, hogy az orvos–orvos kommunikációban a betegség köznyelvi megfelelőjére nincsen szükség. Ha az angol címre intralingvális fordításként tekintünk, akkor a magyarázatra (*cancer of the bone marrow*) a szerző által végrehajtott tudatos betoldásként, explicitációként tekinthetünk. A magyar és a francia fordításokban megfigyelhető a betegség magyarázatának átváltása (*a csontvelő daganatos betegsége, cancer de la moelle osseuse*), a fordító nem vett el és nem adott hozzá a forrásnyelv címéhez. Azt gondolom azonban, hogy a fordító részéről éppen olyan tudatosság figyelhető meg, mint amilyen a szerző részéről, amikor megadja a tudományos terminusok közérthető magyarázatát. Az

orvos–laikus kommunikáció az orvosi szaknyelv egyik diasztratikus változata, és természetes, hogy az orvosi szaknyelvben (ellentétben például a szépirodalmi szövegekkel vagy a sajtószövegekkel) jelentős számú tudományos terminus azonosítható. Nem ritka, hogy a betegtájékoztatók végén külön fejezetet szánnak az adott témában releváns fogalmak tisztázására, meghatározására, ahogy az a korpusz több betegtájékoztatójában is látható (pl.: *Multiple myeloma, cancer of the bone marrow; What is a genetic test?*). A terminusok magyarázó jellegű, laikusok számára érthető tisztázása során ugyancsak felmerülhetnek idegen szavak, növelve ezzel a CNYSZ idegen eredetű terminusainak számát. Ebben a fordítási stratégiában az idegen szavak száma nem nő a fordításokban, de nem is csökken. Megjegyzendő azonban, hogy már ennél a stratégiánál is felmerül, hogy egy betegség tudományos megnevezése is lehet idegenítő a magyar olvasó számára úgy, hogy az angol laikus olvasó az angol terminust ismeri.

Úgy gondolom, hogy a betegtájékoztatók nyelvészeti elemzéséről szóló tanulmányokból kiemelt idézetek rámutatnak arra, hogy a laikusoknak szánt orvosi dokumentumokban fellelhető idegen szavak használata természetes, magától értetődő.

> A szituáció sajátosságai következtében ez egyirányú kommunikáció, azaz információközlés, melynek során nincs mód a visszacsatolásra, a dekódolás nem kontrollálható. Ezért a szövegszerkesztés során kell olyan elemeket […] alkalmazni, amelyek *megkönnyítik a laikusok számára íródott, de a szakszöveg elemeit is tartalmazó információ megértését,* lehetővé teszik, hogy a beteg valóban megértse az instrukciókat és azoknak megfelelően járjon el. (Illésné Kovács és Simigné Fenyő 2008: 168).

Szó sincs tehát arról, hogy az idegen szavakat mellőzni kellene laikusnak szánt orvosi iratok szerkesztésekor vagy fordításakor. A lényeg, hogy betegtájékoztatók esetén „a kommunikáció sikere […] a mondanivaló közérthetőségének, *a szakkifejezések világos használatának, ill. magyarázatának, az egyértelmű, pontos fogalmazásnak a függvénye*" (Illésné Kovács és Simigné Fenyő 2008: 168), és nem a szakkifejezések elhagyásának.

Úgy gondolom, hogy a megfelelő betegedukáció része, hogy a beteg képes betegségéről, az őt érintő orvosi beavatkozásokról beszélni olyan terminusok használatával, amelyeket a mindennapokban talán mellőz. Ennek fényében állapítható meg az (11) példában is fellelhető fordítási stratégia, amely a magyar és a francia fordításokra is éppúgy jellemző. Ha tehát elfogadjuk a feltételezést, hogy az angol autentikus szövegek voltaképpen intralingvális fordítás eredményképpen keletkezett szövegek, akkor érdemes az idegen szavakra vonatkozó fordítási stratégiát is ebből a szempontból megközelíteni. Ebben a gondolatban a betegtájékoztatókban azonosítható első fordítási stratégia elnevezése példákkal:

I. A tudományos terminus és a rá vonatkozó magyarázat vagy definíció is változatlan marad

(12) *Myeloma* is literally an "*oma*," or *tumor*, involving the "*myelo*," or *blood-producing cells in the bone marrow*. The cells that are affected are *plasma cells (a type of white blood cell)*, which are our *antibody-producing (immunoglobulin-producing) cells*.

(12a) *A mielóma* a szó szoros értelmében egy „*oma*", azaz *daganat*, ami a „*myelo*"-t, vagyis a *csontvelőben található vértermelő sejtek*et érinti. Az érintett sejtek a *plazmasejtek (a fehérvérsejtek egyik típusa)*, amelyek *az emberi szervezet antitest-termelői (immunglobulin termelők)*.

(12b) Le *myélome* est littéralement un « *ome* » ou *tumeur*, impliquant le « *myélo* », ou *les cellules produisant le sang dans la moelle osseuse*. Les cellules affectées sont les *plasmocytes (une catégorie de leucocytes)*, qui sont *les cellules produisant nos anticorps* (ou *immunoglobuline*).

A (12) angol nyelvű példában és annak magyar és francia fordításán jól látszik, hogy mind az autentikus, mind a fordított szövegek redundánsak, magyarázó jellegűek. Úgy gondolom, hogy ezen felüli információk betoldása, magyarázata valójában a fel-

dolgozást nehezítő, túlzott redundanciát eredményezne. A szövegek magyarázó jellegét nemcsak a példamondatokban dőlt betűvel szedett, idegen eredetű terminusok és azok köznyelvi megfelelője érezteti, hanem az egyes magyarázó szerkezetek és kötőszók (*literally, a szó szoros értelmében, littéralement; or, azaz, ou; involving, impliquant; or, vagyis, ou*) megléte, a vonatkozói alárendelésre használt vonatkozói névmások (*that, which; ami, amelyek; qui*), valamint az idegen eredetű, tudományos terminusok után zárójelben betoldott köznyelvi megfelelők (*a type of white blood cell, a fehérvérsejtek egyik típusa, une catégorie de leucocytes)*), vagy a köznyelvi terminusok után zárójelben betoldott tudományos terminusok (*immunoglobulin-producing, immunglobulin termelők, immunoglobuline*).

Ebben a fordítói stratégiában az idegen szavak száma nem feltétlenül nő a fordításokban, bár bizonyos esetekben a magyar szöveg az elemzéshez használt szótárak alapján idegenítőnek érez olyan szakszavakat (pl.: *antitest*), amelyeket az angol szótár (pl.: *antibody*) nem érez annak. A betegtájékoztatók magyarázó jellegűek, de a magyarázatban is lehetnek olyan szakkifejezések, amelyek a célnyelvi olvasó számára idegenszerűek.

Az *1.* és a *2. ábra* rávilágít, hogy gyakorlatilag semmilyen tudományos, idegen eredetű morféma vagy terminus nem marad köznyelvi megfelelő, definíció vagy magyarázat nélkül. Az ábrák bal oldali részén tudományos terminusok vannak, míg a jobb oldali rész inkább a köznyelvhez sorolható. Érdekes, hogy önmagában a bal oldali vagy önmagában a jobb oldali elemek is elengedők lennének az információ átadására, hiszen voltaképpen szinonimákról beszélünk. A szerzők tehát többletinformációt vezetnek be, a szövegeket redundánssá teszik. Közérthetőségre, de mindemellett a tudományos terminusok ismertetésére, magyarázatára is törekednek

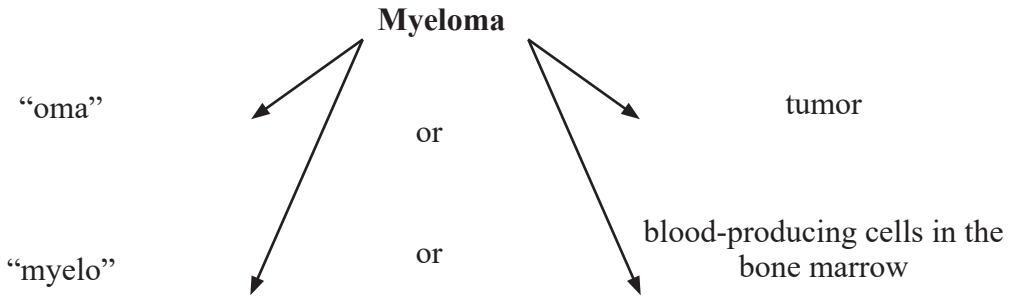
Ugyanez a tendencia figyelhető meg a *2. ábrán*, ahol a szerzők nem kötőszavakkal, hanem zárójelekbe tett köznyelvi megfelelőkkel vagy tudományos terminussal adnak meg szinonimákat.

A következő azonosított fordítási stratégia lényege, hogy a fordító feltehetőleg túl idegenszerűnek érzi az adott terminust, ezért mellékes információként, zárójelben

meghagyja azt, de alapvetően a köznyelvi megfelelőt részesíti előnyben. Ahogy arra Muñoz-Miquel (2012: 200–202) korábban rámutatott, ez a stratégia leginkább a szövegek elején jellemző. A következő példában rögtön az egyik betegtájékoztató címében találkozunk ezzel a jelenséggel. A jelen stratégiát a korpuszból vett példákkal szemléltetem.

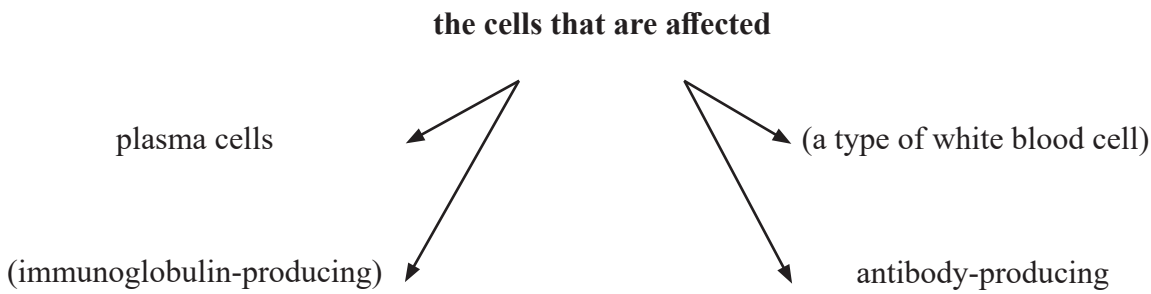
1. ábra

A myeloma értelmezése



2. ábra

A the cells that are affected értelmezése



II. Köznyelvi szó betoldása és az idegen eredetű terminus szerepeltetése zárójelben

(13) *amniocentesis*

(13a) magzatvízvizsgálat (*amniocentesis*)

Ez a fordítási stratégia törekszik arra, hogy a beteget megismertesse az adott beavatkozás tudományos megfelelőjével, de a magyar fordítás további részében az *amniocentesis* idegen eredetű terminus egyszer sem szerepel. Természetesen fontos, hogy ebben az esetben a fordítónak lehetősége nyílik arra, hogy a köznyelvi szót válassza, amit nem tehetne meg akkor, ha például egy betegség hivatalos elnevezéséről lenne szó (pl.: *leukémia* vs *fehérvérűség*), ahogy arra korábban utaltam. A szóban forgó, angol betegtájékoztatóban és annak francia fordításában a kérdéses idegen eredetű terminus (*amniocentesis, amniocèntese*) 28 alkalommal szerepel. Ez arra utal, hogy a francia fordító nem érzi szükségét, hogy köznyelvi szóval helyettesítse a tudományos terminust, ami azzal magyarázható, hogy angolban és franciában már a köznyelvben is lényegesen magasabb az idegen eredetű szavak száma, mint a magyarban.

Ezzel szemben a magyar fordításban összesen 1 alkalommal fordul elő az *amniocentesis* terminus, ahogy a mielómáról szóló tájékoztatóban is 1-szer fordul elő az *anemia* angol terminus közvetlenül átváltott alakja (*anémia)*, 1-szer a *vérszegénység (anémia)* alak, és 20 alkalommal a *vérszegénység* köznyelvi szó. A magyar fordító ezen kívül a következő köznyelvi szavakat használja az *amniocentesis* terminus magyar megfeleltetésére, összesen 27 alkalommal (a zárójelben pontosítom a szavak előfordulását): *magzatvízvizsgálat* (17), *vizsgálat* (9), *beavatkozás* (1). Rendszerint megfigyelhető, hogy a fordítás használ terminológiai szinonimát ott, ahol az angol egy terminust használ a szöveg egészén (ismétléskerülési hipotézis):

– *mutation (EN)*

- *kromoszóma elváltozás (mutáció) (HU), mutáció (HU), genetikai változás (HU)*;
- *modification (mutation) (FR), altération (mutation) (FR), mutation (FR)*,

– *foetus* (EN)

- *magzat (HU), születendő gyermek (HU), gyermek (HU);*
- *fœtus (FR), bébé (FR), enfant (FR)*.

Angolról magyarra fordításokban szintén megfigyelhető, hogy az idegen eredetű terminusok és a köznyelvi szavak a szövegben folyamatosan váltják egymást, hogy a fordító igyekszik szinonimákat használni ott, ahol a FNYSZ nem teszi. A fordított szövegekben használt szinonimák között is lehetnek azonban LGA terminusok, és a forrásnyelvi tudományos terminus is minden esetben szerepel a fordításban. Ennek értelmében bár a fordító törekszik köznyelvi szavak használatára, az idegen szavak száma a fordításokban feltehetőleg nem lesz jelentősen kevesebb. Sőt, a fordító valójában a zárójelben való pontosítást is többször végig viszi a szöveg egészén, ahogy azt a következő példák szemléltetik:

– translocation (6) →	transzlokáció (áthelyeződés) (a tájékoztató legelején szerepel) (1), transzlokáció (4), átrendeződés (1);
– chromosome deletion (6) →	kromoszóma törlődés (deléció) (a tájékoztató legelején), kromoszóma deléciók (tőrlődések) (2), deléció (2), kihagyás (1);
– duplication (3) →	duplikációk (kettőződés) (a tájékoztató legelején), megkettőződés (duplikáció), kihagyás (1);
– insertion (3) →	hozzáadás (inszerció) (a tájékoztató legelején), inszerció (hozzáadás), inszerció;
– inversion (4) →	megfordulás (inverzió) (a tájékoztató legelején), inverzió (megfordulás), inverzió (1).

Úgy gondolom, hogy a jelen példák is rámutatnak a korábban felvetett feltételezésre,amely szerint a fordító (vagy a szöveg szerzője) feltett szándéka, hogy az olvasó megismerje az adott betegség, beavatkozás szakterminológiájának egy részét. Az utóbbi felsorolás két utolsó példájában is látható, hogy a fordító először a tudományos termi-

nust szerepelteti mellékinformációként, zárójelbe téve. A szöveg további részén azonban a köznyelvi szó szerepel a tudományos terminus után tett zárójelben, míg végül a köznyelvi szót a fordító elhagyja, és csak a tudományos terminust szerepelteti.

Mielőtt rátérek a következő stratégiára, a II. stratégiára az előzőktől némileg különböző példát hozok. Ebben az esetben a fordító nemcsak az angol zárójelben szereplő mozaikszót és annak az angol szövegben is szereplő feloldását tartja meg, hanem mindezt zárójelbe téve megadja a fogalom magyar vagy francia tudományos megfelelőjét. Ilyen esetekkel akkor találkozhatunk, ha az angol mozaikszó elterjedt, használatos a magyar és a francia klinikumban is, de ismerete feltehetőleg a szakemberekre korlátozódik. Ezekben az esetekben a FNYSZ idegen szavainak száma voltaképpen kétszeresére nő, hiszen az összetett angol terminus mellett a célnyelvi tudományos megfelelő is szerepel.

(14) The very earliest stage is called *Monoclonal Gammopathy of Undetermined Significance (MGUS)*.

(14a) A legkorábbi stádiumot bizonytalan jelentőségű monoklonális gammopátiának *(MGUS, monoclonal gammopathy of undetermined significance)* nevezik.

(14b) Le stade le plus précoce est appelé gammapathie monoclonale de signification indéterminée *(MGUS, Monoclonal Gammopathy of Undetermined Significance)*.

Előfordul, hogy az adott kontextus lehetőséget ad arra, hogy a fordító kihagyja az idegen eredetű terminust.

Ezen a ponton érkezünk el a következő fordítói stratégiához:

III. Idegen eredetű terminus kihagyás

(15) Making a decision about having an amniocentesis test during pregnancy can be difficult. It is important to remember that you do not have to take the *amniocentesis test* if you do not want to.

(15a) Nem könnyű döntést hozni arról, hogy a terhesség során Ön szeretné-e a vizsgálatot. Fontos, hogy tudja, hogy *ez* nem kötelező.

(16) How long will it take to get the results of the *amniocentesis*?

(16a) Mennyi időbe telik, amíg kézhez kapom az eredményeket?

(17) Monitored by tracking monoclonal protein in serum using Serum Protein Electrophoresis *(SPEP)* (IgG) and/or quantitative immunoglobulin (QIG) measurement (IgA/D/E).

(17a) Monoklonális fehérje ellenőrzése *szérum elektroforézissel* (IgG) vagy/és immunglobulin (QIG) méréssel (IgA/D/E).

Ez a stratégia csak akkor figyelhető meg, ha a szóban forgó terminus elhagyható. Erre egyrészt akkor van lehetőség, ha például a fenti példában (15a) bemutatott mutató névmás (*ez*) alkalmas arra, hogy anaforikus funkciójának köszönhetően visszautaljon antécédensére (*a vizsgálat*). Lehetséges a megoldás akkor is, ha a kontextus egyértelművé teszi, hogy pontosan miről van szó, és a fordító nem terheli felesleges információkkal az olvasót (16a). Lehetséges a stratégia továbbá akkor, ha az angol szövegben többletinformáció van, mint ahogy az a (17) példában (*SPEP*) is megfigyelhető. Ezekben az esetekben a fordító feltehetőleg túlzott redundanciának véli a terminológiai szinonimák használatát, és egyetlen terminussal váltja át az angolban két terminussal pontosított fogalmat.

Efféle túlzott forrásnyelvi redundanciáról beszélhetünk akkor, amikor az angol szövegben a tudományos terminust magyarázat követi, de valójában a köznyelvi szinonima is megállná a helyét, vagy a terminus voltaképpen nem szorul magyarázatra. Érdekes, hogy a következő példában a magyarázat a magyarban és a franciában is elmarad, de ettől

függetlenül magát a tudományos terminust a francia megtartja, a magyar pedig köznyelvi szóval helyettesíti. Ez a példa ugyancsak rámutat, hogy az egyes fordítási stratégiák nem kizárólagosak, hiszen ebben az esetben is beszélhetünk az idegen szóra vonatkozó zárójeles megjegyzés kihagyásáról, a francia esetén magyarázat nélküli átváltásról (pl. 18b), a magyar esetén pedig a tudományos terminus köznyelvi szóval való helyettesítéséről (pl. 18a). Az újlatin ajkú francia olvasó számára valószínűleg nem igényel magyarázatot a *prognosis* terminus, a magyar olvasót figyelembe véve pedig feltehetőleg célravezetőbb a köznyelvi szó választása. A számadatok alapján azonban ez a fordítói stratégia ritkábban alkalmazott, hiszen a CNYSZ-ben több LGA terminus azonosítható, mint a FNYSZ-ben.

(18) The *prognosis (from the Greek words that mean "knowing ahead")* […] is better when treatment is started early and bone disease or other complications can be prevented.

(18a) A […] betegek *kilátásai* jobbak, ha a betegség kezelése korán elkezdődik, megelőzve a csontbetegségek és egyéb szövődmények kialakulását.

(18b) Le pronostic pour les patients […] est meilleur lorsque le *traitement* commence tôt et la maladie osseuse ou d'autres complications peuvent être évitées.

Újabb stratégiát fedezhetünk fel, amikor a fordító ugyancsak köznyelvi szóval feleltet meg az idegen eredetű terminust, de nem érzi szükségét, hogy megadja a tudományos terminust is. Ezek az esetek egyrészt általában a kevesebbszer előforduló terminusokra jellemzőbbek, amelyek ismerete nem feltétlenül fontos a beteg számára, hiszen az adott betegséghez vagy beavatkozáshoz közvetlenül nem kapcsolódnak. Másrészt akkor azonosíthatjuk ezt a stratégiát, ha a köznyelvi megfelelő az egészségügyi szakszemélyzet által is tökéletesen elfogadott, a fogalmat jól lefedő elnevezés. Ez az eljárás is elsősorban a magyar fordításokban figyelhető meg. Az angol és a francia korpuszban például nem szerepel köznyelvi megfelelője a *placenta, rectum, transmission, cornea, pleura, peritoneum*

terminusoknak, a magyarban pedig maga a tudományos terminus nem szerepel, csak a köznyelvi megfelelője (*méhlepény, végbél, átadás, szaruhártya, mellhártya, hashártya*). Anatómiai megnevezéseknél itt rendszerint metonimikus kapcsolat (szinechdoché) van a két fogalom között, hiszen a rendkívül precíz anatómiai terminus helyett általánosabb, „nagyobb területet jelölő" megfelelőt választunk. Érdekes, hogy ez a stratégia akkor is azonosítható, ha az angol szerepelteti a tudományos terminust és a köznyelvi megfelelőt is, de a magyar köznyelvi megfelelőt és közérthető definíciót használ [pl.: *cervix (entrance to womb)* → méhnyak, a méh bejárata]. A stratégia nem csökkenti a CNYSZ idegen szavainak számát, hiszen ahogy fentebb írtam, alapvetően a kevesebbszer előforduló, az adott betegség szempontjából nem kulcsfontosságú fogalmak megnevezéséről van szó. A következő azonosítható fordítói stratégia ennek értelmében:

IV. Az idegen eredetű terminus megfeleltetése köznyelvi szóval

(19) You have had another type of test that is done during pregnancy (such as an ultrasound, *nuchal translucency scan* or blood test). It has shown that there is an increased risk that your baby has a genetic condition.

(19a) A terhességi szűrővizsgálatok során (ultrahang, *nyaki bőrredő vizsgálata*, vérvizsgálat) az Ön eredménye arra utal, hogy magzatának fokozott az esélye egy esetleges genetikai betegségre.

(19b) Vous avez eu un autre type d'examen pendant votre grossesse (échographie, *mesure de la clarté* nucale ou test sanguin). Cet examen montre qu'il existe un risque que votre bébé soit atteint d'une maladie génétique.

A (19) példamondat érdekessége, hogy háromszorosan összetett terminust mutat be (*nuchal translucency scan*), amelyek mindegyike önállóan is terminusértékű. A magyar szöveg mindhárom terminust köznyelvi szóval váltja fel, míg a francia kettőt köznyelvi szóval feleltet meg (*mesure, clarté*), egy alkalommal pedig megtartja a tudományos terminust (*nucale*). Ezt azért tartom érdekesnek és fontosnak, mert a példa rávilágít, hogy

összetett terminusok esetén a fordítói stratégiák váltakozhatnak, használatuk nem kizárólagos. Többtagú terminusok fordításánál adott esetben több fordítási stratégia is azonosítható. Különösen érdekesnek tartom ezt az eljárást akkor, amikor a fordító a tudományos terminus megtartása mellett olyan köznyelvi szó betoldását választja, amelynek metaforikus stilisztikai értéke van. A következő példa erre a stratégiára mutat rá.

(20) The cells activate osteoclast cells, which destroy bone, and block *osteoblast cells*, which normally repair damaged bone.

(20a) A sejtek aktiválják az *oszteoklaszt (csontfaló) sejteket*, melyek elpusztítják a csontokat és blokkolják a csontképző sejteket, melyek normál esetben helyrehozzák a csontkárosodást.

V. A tudományos terminus átváltása magyarázat vagy definíció nélkül

(21) Common conditions inherited in this way include *fragile X, Duchenne muscular dystrophy and haemophilia.*

(21a) Az ily módon örökölt betegségek pl. a *fragilis X szindróma*, a *Duchenne izomdisztrófia és a hemofilia*.

(21b) Des exemples de maladies liées à l'X sont le syndrome de *l'X fragile*, la *dystrophie musculaire de Duchenne et l'hémophilie*.

(22) What is autosomal recessive inheritance?

(22a) Mi az *autoszomális recesszív* öröklődés?

(22b) Qu'est-ce que *l'hérédité autosomique récessive*?

Megesik, hogy a fordító valami miatt az idegen eredetű tudományos terminust átváltja a fordításban, és nem is fűz hozzá definíciót vagy magyarázatot. Ez a stratégia

hasonlít az első azonosított stratégiára. A különbség az, hogy az ötödik stratégia esetében az LGA terminusok lexikai környezetében sem a FNYSZ, sem a CNYSZ nem használ az LGA terminusokra vonatkozó magyarázatokat, egyik nyelv sem told be köznyelvi szavakat, szinonimákat az LGA terminusok előtt vagy után.

Ezekben az esetekben véleményem szerint az készteti a fordítót a magyarázat, definíció betoldásának elhagyására, hogy egyrészt az adott betegségek a betegtájékoztató további részében nem kapnak kiemelt hangsúlyt, nem fontosok azoknak, akik nem ezekben a betegségekben szenvednek. Másrészt feltételezhető, hogy bár az *izomdisztrófia* második tagja idegen eredetű terminus, a hibrid első tagja (*izom*) köznyelvi szó, a *hemofília* pedig talán többek számára ismeretes betegség. A (22) példában pedig feltehetőleg azért marad el a magyarázat, mert a példa valójában egy alfejezet címét jelöli. Ahogy korábban rámutattam, az idegen szavak elemzésekor fontos elemzési lépés az idegen szavak kontextusa, hiszen a jelen esetben sem beszélhetünk a tudományos terminus magyarázatának elmaradásáról, mert külön alfejezetet szenteltek arra, hogy definiálják az autoszomális recesszív öröklődést. Erre csak akkor deríthetünk fényt, ha nemcsak az idegen szavak számát, hanem kontextusát is tanulmányozzuk. A stratégia használatával tovább nő a CNYSZ idegen szavainak száma. Ugyanez jellemző a következő fordítási stratégiára is.

A következő részben olyan példákat mutatok be, amelyekre jellemző, hogy a fordító tudományos terminust használ, de a tudományos terminus eltér attól, amelyet a FNYSZ szerzője használt. Ez a fordítási stratégia a korpuszban angolról franciára fordításokban figyelhető meg.

VI. A tudományos terminus megfeleltetése más tudományos terminussal

(23) Some examples of genetic conditions include *Down's syndrome, cystic fibrosis* and *muscular dystrophy*.

(23a) Quelques exemples de maladies génétiques sont la *trisomie 21, la mucoviscidose* ou les *myopathies*.

A magyar fordításban akkor figyelhetjük meg ezt a stratégiát, ha az angol FNYSZ mozaikszót használ, a magyar fordító pedig ezt feloldani igyekszik (pl.: 24–24a). Fel is oldja, de minden mozaikszó minden betűjét talán lehetetlen köznyelvi szóval helyettesíteni. Ezért sorolom ezt az eljárást abba a kategóriába, amikor bár részlegesen, de a magyar fordító más tudományos terminust használt, mint az angol, vélhetőleg azért, mert az adott mozaikszó a célnyelvi laikus számára a legkevésbé sem ismert.

(24) Both *SPEP* and *UPEP* are negative (no monoclonal spike in serum or urine).

(24a) Mind a *szérum*, mind pedig *a fehérje elktroforézis vizsgálat* is negatív (nincs monoklonális tüske sem a szérumban, sem a vizeletben).

Ez a fordítási stratégia rávilágít a terminológiai szinonimák fontosságára. Az orvosi szaknyelv kötöttsége ellenére ismer terminológiai szinonimákat, és egy orvostudományi fogalmat nemcsak több köznyelvi szó, de több tudományos terminus is jelölhet. Érdekes, hogy az angol FNYSZ-ben (23) szereplő eponimát (*Down's syndrome*) a francia fordító a betegséget okozó 21. kromoszómapárra (*trisomie 21*) való átváltással felelteti meg a betegséget először leíró orvosról elnevezett tulajdonneves változat helyett. Úgy gondolom, hogy érdekes kutatásokat lehetne végezni abban a tekintetben, hogy az adott nyelvek és kultúrák mennyire törekszenek az eponimák vagy egyéb kulturális elemek használatára vagy elhagyására. A Down-kór története különösen érdekes, hiszen a betegséget diagnosztizáló orvos korában a szindrómát *mongolizmusnak* nevezték. E pejoratív jelölés ma már természetesen minden kontextusban kerülendő. Érdekesnek tartom, hogy a szaknyelvi kötöttség ellenére az orvosi szövegnél is fontos kérdéseket vethetnek fel a kulturális elemek, ahogy arra az imént utaltam. A korpuszomban észrevehető, hogy a fordító esetenként teljes bekezdéseket hagy ki, ha úgy érzi, hogy az adott szövegrész a fordítás nyelvének kultúrájában nem releváns. Itt azonban kevésbé beszélhetünk az idegen szavak kihagyásának fordítói stratégiájáról, sokkal inkább a szöveg pragmatikai adaptációjáról van szó. Az alábbi angol szövegrészlet fordítása mindenesetre elmaradt mind a magyar, mind a francia fordításban.

> There is an increased risk of having a child with a particular genetic condition because of your ethnic background. Examples of this include sickle cell in people of Afro-Caribbean descent, beta-thalassaemia in people of Mediterranean descent, cystic fibrosis in people of Western European descent and Tay Sachs in people of Ashkenazi Jewish descent. (Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2007: 4)

Az elemzés második kutatási kérdése arra irányult, hogy milyen fordítási stratégiákat alkalmaz a fordító az adott korpuszban az idegen szavak megfeleltetésekor, hogy hogyan igyekszik kielégíteni a laikus célnyelvi olvasó igényeit. Elemzésemben összesen hat fordítási stratégiát állapítottam meg: a tudományos terminus és a rá vonatkozó magyarázat vagy definíció is változatlan marad (I.), köznyelvi szó betoldása és az idegen eredetű terminus szerepeltetése zárójelben (II.), LGA kihagyása (III.), idegen eredetű terminus megfeleltetése köznyelvi szóval (IV.), a tudományos terminus átváltása magyarázat vagy definíció nélkül (V.), a tudományos terminus megfeleltetése más tudományos terminussal (VI.).

5. Összefoglalás és kitekintés

Az elemzés eredményei rámutatnak, hogy az adott szövegek vizsgálata alapján az autentikus angol szövegekre kevésbé jellemző az idegen eredetű terminusok használata, mint a fordított francia alkorpuszra. A legtöbb idegen szót a fordított magyar alkorpuszban azonosítottam. Ennek oka a nyelvek szókészleti különbségeiben keresendő, hiszen az angol köznyelvben alapvetően több LG terminus használatos, mint a magyarban, tehát a magyar beszélő idegenítőnek érzékelhet olyan terminusokat, amelyek az angolban a laikusok számára nem okoznak jelentős feldolgozási erőfeszítést. Fontos megjegyezni, hogy a magyar és a francia orvosi nyelvbe szüntelenül beáramló angol szakkifejezések növelik az idegen szavak számát, az angolban pedig értelemszerűen nem hatnak idegenszerűen az angolul elnevezett szakkifejezések. A betegedukáció része, hogy a betegek megismerkednek a be-

tegségüket megnevező, illetve az arra alkalmazott kezelések legfontosabb terminusaival, hogy hatékonyan tudjanak kommunikálni az őket gondozó szakszemélyzettel. A fordítók ezért is törekedhetnek az idegen szavak használatára. Arra következtethetünk azonban, hogy a magyar és a francia célnyelvi olvasónak a felsorolt okok miatt nagyobb feldolgozási erőfeszítésébe telik a betegtájékoztatók megértése.

Felmerül továbbá a kérdés, hogy milyen eredményeket kapnánk, ha az elemzéshez nem az adott nyelveken írt idegen szavak szótárát használnánk, hanem hatásvizsgálatot végeznénk angol, francia és magyar anyanyelvi beszélők körében. Lehetséges, hogy maguk a szótárak sem teljesen azonos rendszerező elvek mentén gyűjtötték össze az idegen szavakat. A kutatás folytatásában vállalkozom a hatásvizsgálat elvégzésére.

Az elemzés a kutatási irány első eredményeit mutatja be, amelyet a jövőben más adatokkal célszerű kiegészíteni. A kutatás folytatásában tervezem elkülöníteni egymástól az idegen szavakat eredetük szerint. Ennek fényében arról is információt kaphatunk, hogy a vizsgált nyelveken milyen arányban fordulnak elő görög–latin, angol, esetleg más eredetű idegen szavak. Tervezem továbbá a lexikai és a grammatikai szavak elkülönítését, hiszen a vizsgált nyelvekre más nyelvtani szerkezetek jellemzők, így további különbségek várhatók az eredmények tekintetében.

A kutatás folytatásában a párhuzamos korpuszelemzést összehasonlítható korpuszelemzéssel tervezem kiegészíteni. Az összehasonlítható korpusz rávilágít arra a kutatás szempontjából fontos kérdésre, hogy bizonyos szavak, terminusok gyakrabban fordulnak-e elő az azonos nyelvű fordított alkorpuszban, mint a nem fordított alkorpuszban, hogy az idegen szavak gyakorisága a két alkorpuszban eltérő-e, hogy azonosítható-e eltolódás.

A kutatásban a betegtájékoztatókra intralingvális fordítás eredményeképpen tekintek, ahol a FNYSZ szerzői az anyanyelvükön igyekeznek a laikus olvasó igényeihez igazítani a betegtájékoztatók szövegét. A fordító az elemzett korpusz alapján arra törekszik, hogy az intralingvális fordítás során keletkező betoldásokat, explicitációkat, magyarázatokat és definíciókat megtartja. Ez azért is fontos, mert az orvosi szaknyelven akkor

is jelentős az idegen szavak ismerete, ha nem orvos–orvos kommunikációról van szó, hiszen hozzátartozik a betegedukációhoz, hogy ne csak az orvos alkalmazkodjon a laikus nyelvhasználóhoz, hanem a beteg is megértse, ismerje a betegségével kapcsolatos alapvető fogalmakat, terminusokat.

A fordítási stratégiákat illetően összesen hat eljárást azonosítottam az általam vizsgált párhuzamos alkorpuszokban. A magyar és a francia fordító bizonyos esetekben megtartja a tudományos terminust és a rá vonatkozó magyarázatot vagy definíciót (I.), mert a FNYSZ eléggé explicit ahhoz, hogy a laikus befogadó igényeit kielégítse, és a további explicitáció túlzott redundanciát eredményezne. A következő stratégia (II.) lényege, hogy a fordító szerepelteti a forrásnyelven is meglévő idegen eredetű terminust, de zárójelbe teszi, és köznyelvi szót használ a megfeleltetéskor. Ez a stratégia leginkább betegtájékoztatók elején figyelhető meg, az idegen eredetű terminus első előfordulásakor, illetve akkor, ha az angol terminus elterjedt, használatos a magyar és francia klinikumban is, de ismerete feltehetőleg a szakemberekre korlátozódik. A stratégia a magyarra fordításra jellemző, mert a francia fordító nem érzi szükségét, hogy köznyelvi szóval helyettesítse a tudományos terminust, ami azzal magyarázható, hogy angolban és franciában már a köznyelvben is lényegesen magasabb az idegen eredetű szavak száma, mint a magyarban, ahogy arra fentebb utaltam. Lehetőség van az LGA terminus kihagyására (III.), ha a fordító más nyelvi elemmel helyettesíti azt (pl.: mutató névmás); ha a kontextus egyértelművé teszi, hogy pontosan miről van szó, ezért az idegen szó elhagyható; vagy ha a tudományos terminus, vagy annak esetleges további terminológiai szinonimákkal való együttes jelenléte túlzott redundanciát okozna. Újabb stratégiát állapíthatunk meg, amikor a fordító köznyelvi szóval felelteti meg az idegen eredetű terminust (IV.). Ezekben az esetekben általában egyrészt a ritkábban előforduló terminusok érintettek, amelyek ismerete nem feltétlenül fontos a beteg számára, hiszen az adott betegséghez vagy beavatkozáshoz közvetlenül nem kapcsolódnak. Másrészt akkor használható ez a stratégia, ha az adott köznyelvi elnevezés az egészségügyi szakszemélyzet számára is tökéletesen elfogadott, a kérdéses fogalmat kellőképpen lefedi. Ez az eljárás is elsősorban a magyar fordításokban figyelhető meg a korábban bemutatott okok miatt. A fordító a tudományos terminust átválthatja magyarázat vagy definíció nélkül (V.), ha az adott betegségek a fordító szerint a laikus célközönség

számára is nagy valószínűséggel ismertek, vagy ha a betegtájékoztató további részében nem kapnak kiemelt hangsúlyt, nem fontosok azok számára, akik érintettek az adott betegségben. Az idegen eredetű tudományos terminus megfeleltetése más idegen eredetű tudományos terminussal (VI.) angolról franciára való fordításban figyelhető meg, és felhívja a figyelmet az orvosi szaknyelven is meglévő terminológiai szinonimák ismeretének fontosságára.

A stratégiák tekintetében megállapítható, hogy a hat fordítási stratégia közül csupán kettő (III.) és (IV.) esetében figyelhető meg a FNYSZ idegen eredetű terminusának kihagyása vagy helyettesítése köznyelvi szóval. Ezekre a stratégiákra azonban csak akkor kerül sor, ha a tudományos terminus elhagyható (pl.: névmás használata a célnyelven), vagy a kevesebbszer előforduló terminusokról van szó, amelyek ismerete nem feltétlenül fontos a beteg számára, hiszen az adott betegséghez vagy beavatkozáshoz közvetlenül nem kapcsolódik. A másik négy stratégia növeli a CNYSZ idegein szavainak számát.

A kutatást a továbbiakban érdemes lenne elvégezni újabb kiadású idegen szavak szótár segítségével. Ezek az eredmények azért is lennének érdekesek, mert rávilágítanának, hogyan változik a vizsgálat szempontjából releváns nyelvekben az idegen szavak gyakorisága. Érdemesnek tartanám a vizsgálatot kiterjeszteni további műfajokra: internetes betegtájékoztatókra, beleegyező nyilatkozatokra stb. Fontos lenne bevonni az elemzésbe az orvos–orvos diasztratikus változat műfajait, mint például a zárójelentés. A két egymástól eltérő műfajból lehetne következtetni azokra a terminushasználati különbségekre és fordítási stratégiákra, amelyek az orvos–orvos és az orvosi–laikus kommunikáció közötti eltérésekre jellemzők. Ugyancsak izgalmas eredmények születhetnének olyan vizsgálatokból, amelyek az intralingvális fordítást is bevonják az elemzésbe, hiszen a betegtájékoztatók is intralingvális fordítás eredményei. Efféle vizsgálattal általános képet nyerhetnénk arról, milyen gyakorisággal használnak LGA terminusokat az orvos–orvos diasztratikus változat részvevői anyanyelvükön, és hogy hogyan és milyen formában változik ez a gyakoriság az intra- és interlingvális fordítás folyamatában.

Irodalom

Baker, M. 1993. Corpus Linguistics and Translation Studies – Implications and Applications. In Baker, M. et al (eds.) *Text and Technology. In Honour of John Sinclair*. Amsterdam/Philadelpia: Jonh Benjamins. 233–250. https://doi.org/10.1075/z.64.15bak

Baker, M. 1996. Corpus-based translation studies: The challenges that lie ahead. In: Somers, H. (ed) *Terminology, LSP and Translation. Studies in language engineering in honour of Juan C. Sager*. Amsterdam/Philadelphia: John Benjamins. 175–186. https://doi.org/10.1075/btl.18.17bak

Bakos F. 2013. *Idegen szavak és kifejezések szótára*. Budapest: Akadémiai Kiadó.

Bősze P. 2011. Latin, magyar, angol? A magyar orvosi szaknyelvről. *Debreceni Szemle*. Vol. 19. No. 4. 369–374.

Cselovszkyné Tarr K. 1999. *Tabu és eufemizmus az egészségügy nyelvében*. Pécs: Pécsi Tudományegyetem.

Demers, G. 1991. L'euphémisme en anglais et en français. *Langues et Linguistique*. No. 17. 17–37.

Fazekas E. 2007. *Bevezetés a magyar nyelvtörténetbe*. Kolozsvár: Egyetemi Műhely Kiadó.

Heltai P. 2004. Terminus és köznyelvi szó. In: Dróth J. (ed.) *Szaknyelv és szakfordítás*. Gödöllő: Szent István Egyetem. 25–46.

Illésné Kovács M., Simigné Fenyő S. 2008. A metainformációs nyelvi elemek funkciója a közvetett kommunikáció egyik fajtájában, betegtájékoztatók szövegében. *Alkalmazott Nyelvészeti Közlemények*. Miskolc: Miskolci Egyetem. 161–171.

Illésné Kovács M., Simigné Fenyő S. 2009. Gyógyszerészeti kísérőiratok kommunikációs szempontú elemzése Hymes szociokulturális modellje alapján. In: Gecső T., Sárdi

Cs. (eds.) *A kommunikáció nyelvészeti aspektusai*. Budapest: Tinta Tankönyvkiadó. Székesfehérvár: Kodolányi János Főiskola. 141–148.

Jamet, D. 2010. Historique et procédés linguistiques de l'euphémisme. In: Jamet, D. et Jobert, M. (eds) *Empreintes de l'euphémisme. Tours et détours.* Paris: L'Harmattan. 30–49.

Jiménez-Crespo, M. Á., Tercedor Sánchez, M. 2017. Lexical variation, register and explicitation in medical translation. A comparable corpus study of medical terminology in US websites translated into Spanish. *Translation and Interpreting Studies*. Vol. 12. No. 3. 405–426. https://doi.org/10.1075/tis.12.3.03jim

Kenny, D. 1998. Creatures of Habit? What Translators Usually Do with Words. *Meta*. Vol. 43. No. 4. 515–523. https://doi.org/10.7202/003302ar

Klaudy K. 2001. Az aszimmetria hipotézis. In: Bartha M. (szerk.). *A X. Magyar Alkalmazott Nyelvészeti Konferencia előadásai*. Székesfehérvár: KJF. 371–378.

Kurtán Zs. 2003. *Szakmai nyelvhasználat*. Budapest: Nemzeti Tankönyvkiadó.

Lanstyák I. 2006. A kölcsönszavak rendszerezéséről. In: *Nyelvből nyelvbe. Tanulmányok a szó* kölcsönzésről, kódváltásról és fordításról. Pozsony: Kalligram Könyvkiadó. 15–56.

Meyer, I., Mackintosh, K. 2000. When terms move into our everyday lives: an overview of determinologization. *Terminology*. Vol. 6. No. 1. 111–138. https://doi.org/10.1075/term.6.1.07mey

Mihály A. 2006. *Anatomia Essentialis. I. kötet. Egységes kötet orvostanhallgatók számára*. Szeged: Szegedi Tudományegyetem, Általános Orvostudományi Kar, Anatómiai, Szövet- és Fejlődéstani Intézet.

Montalt Resurrecció, V. 2011. Medical translation. In: Chapelle, C. A. (ed) *Encyclopedia of Applied Linguistics*. Hoboken, NJ: Wiley. 3649–3653.

Montalt Resurrecció, V., González Davies, M. 2007. *Medical Translation Step by Step. Translation Practices Explained*. Manchester: St. Jerome.

Muñoz-Miquel, A. 2012. From the original article to the summary for patients: Reformulation procedures in intralingual translation. *Linguistica Antverpiensia*. No. 11. 187–206.

Nisbeth Jensen, M., Zethsen, K. K. 2012. Patient information leaflets: Trained translators and pharmacists-cum-translators – a comparison. *Linguistica Antverpiensia New Series. Themes in Translation Studies*. No. 11. 31–49.

Putz O. 2012. Idegen szavak napjaink orvosi szaknyelvében. In: Horváthné Molnár K., Sciacovelli, A. D. (eds.) *Az alkalmazott nyelvészet regionális és globális szerepe. A XXI. Magyar Alkalmazott Nyelvészeti Kongresszus előadásai.* Budapest-Szombathely-Sopron: MANYE-NYME. 365–371.

Robin E., Dankó Sz., Götz A., Nagy A. L., Pataky É., Szegh H., Török G., Zolczer P. 2016. Fordítástudomány és korpuszkutatás: bemutatkozik a Pannónia Korpusz. *Fordítástudomány*. 18. évf. 2. szám. 5–26.

Seidl-Péch O. 2018. Melyek a (szak)fordítő és a fordításkutató munkáját segítő legfontosabb nyelvi korpuszok? In: Robin E. – Zachar V. (szerk.): *Fordítástudomány ma és holnap*. Budapest: L'Harmattan Kiadó. 175–191.

Speake, J. 1997. *The Oxford Dictionary of Foreign Words and Phrases*. Oxford: Oxford Univerity Press.

Tótfalusi I. 2008. *Idegenszó-tár. Idegen szavak értelmező és etimológiai szótára*. Budapest: Tinta Könyvkiadó.

Tournier, J. 1985. *Introduction descriptive à la lexicogénétique de l'anglais contemporain*. Paris-Genève: Champion-Slatkine.

Varga É. K. 2014. Anatómiai nevek régen és ma. Latin helyett angol? *Modern Nyelvoktatás*. Vol. 20. No. 1–2. 35–42.

Warren, B. 1992. What euphemisms tell us about the interpretation of words. *Studia Linguistica*. Vol. 46. No. 2. 128–172. https://doi.org/10.1111/j.1467-9582.1992.tb00833.x

Wright, S. E. 2011. Scientific, technical and medical translation. In: Malmkjaer, K., Windle, K. (eds.) *The Oxford Handbook of Translation Studies*. Oxford: Oxford University Press. 243–261.

Zethsen, K. K. 2004. Latin-based terms: True or false friends? *Target*. Vol. 16. No. 1. 125–142.

Internetes hivatkozások

AntConc 3.5.0. Elérhető: https://www.laurenceanthony.net/software/antconc/

Betegségek Nemzetközi Osztályozása (BNO). Elérhető: www.webbeteg.hu/keresok/bno/13403

MemoQ. Elérhető: https://www.memoq.com/

VariMed adatbázis. Elérhető: varimed.ugr.es

WordSmith Tools 6.0. Elérhető: https://lexically.net/wordsmith/version6/

Források

Az autentikus angol korpusz

G.M. Durie, Brian. 2012/2013. *Multiple Myeloma, Cancer of the Bone Marrow. Patient Handbook*. North Hollywood: International Myeloma Foundation. http://www.amen.org.il/wp-content/uploads/2012/02/Patient_Handbook_2013.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2009. *Carrier Testing.* Leuven: EuroGentest. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/english/carrier_testing.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2007. *Chromosome Changes.* Leuven: EuroGentest. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/english/chromosome_changes.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2007. *Chromosome Translocations.* Leuven: EuroGentest.http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/english/chromosome_translocations.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2007. *Dominant inheritance.* Leuven: EuroGentest. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/english/Dominant_Inheritance.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2007. *Frequently Asked Questions about Genetic Testing.* Leuven: EuroGentest. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/english/FAQ.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2007. *Recessive inheritance.* Leuven: EuroGentest. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/english/recessive_inheritance.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2007. *The Amniocentesis Test: Information for Patients and Families.* Leuven: EuroGentest. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/english/amniocentesis.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2007. *What is a Genetic Test?* Leuven: EuroGentest. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/english/genetic_test.pdf

International Society for Stem Cell Research. 2008. *Patient Handbook on Stem Cell Therapies*. Skokie: International Society for Stem Cell Research.http://www.isscr.org/docs/default-source/patient-handbook/isscrpatienthandbook.pdf

A fordított francia korpusz

G.M. Durie, Brian. 2015. *Myélome multiple, Cancer de la moelle osseuse*. *Guide du patient*. North Hollywood: International Myeloma Foundation. https://www.myeloma.org/sites/default/files/images/publications/International/PDF/french/guide_du_patient.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2007. *Amniocèntese : information pour les malades et leurs familles.* Leuven: Orphanet. Fordította: Orphanet. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/french/amniocentesis.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2009. *Anomalies chromosomiques: Information pour les malades et leurs familles.* Leuven: Orphanet. Fordította: Orphanet. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/french/chromosome_changes.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2007. *Hérédité dominante : Information pour les malades et leurs familles.* Leuven: Orphanet. Fordította: Orphanet. http://

www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/french/Dominant_Inheritance.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2007. *Hérédité récessive: Information pour les malades et leurs familles.* Leuven: Orphanet. Fordította: Orphanet. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/french/recessive_inheritance.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2009. *Test génétique de porteur: Information pour les patients et les familles.* Leuven: Orphanet. Fordította: Orphanet. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/french/carrier_testing.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2007. *Translocations chromosomiques: Information pour les malades et leurs familles.* Leuven: Orphanet. Fordította: Orphanet. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/french/chromosome_translocations.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2007. *Qu'est-ce qu'un test génétique?* Leuven: Orphanet. Fordította: Orphanet. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/french/genetic_test.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2007. *Questions fréquemment posées sur le test génétique: Information pour les malades et leurs familles.* Leuven: Orphanet. Fordította: Orphanet. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/french/FAQ.pdf

International Society for Stem Cell Research. 2008. *Guide à l'intention des patients sur les thérapies à base de cellules souches*. Ottawa: Réseau de cellules souches. For-

dította: Réseau de cellules souches. http://www.isscr.org/docs/default-source/patient-handbook/isscr_patientprimerhndbk_french_fnl.pdf

A fordított magyar korpusz

G.M. Durie, Brian. 2016. *Mielóma multiplex, a csontvelő daganatos betegsége. Betegek kézikönyve.* North Hollywood: International Myeloma Foundation. https://www.myeloma.org/sites/default/files/images/publications/International/PDF/hungarian/phb_2016_hu.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2009. *A domináns öröklődés*. Pécs: Orvosgenetikai Intézet, Pécsi Tudományegyetem. Fordította: Dr. Komlósi Katalin. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/hungarian/Dominant_Inheritance.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2009. *A magzatvízvizsgálat (amniocentesis).* Pécs: Orvosgenetikai Intézet, Pécsi Tudományegyetem. Fordította: Dr. Komlósi Katalin. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/hungarian/amniocentesis.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2009. *Gyakran feltett kérdések a genetikai vizsgálattal kapcsolatban*. Pécs: Orvosgenetikai Intézet, Pécsi Tudományegyetem. Fordította: Dr. Komlósi Katalin. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/hungarian/FAQ.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2009. *Hordozóság szűrés.* Budapest: Országos Környezetegészségügyi Intézet, Molekuláris Genetikai és Diagnosztikai Osz-

tály. Fordította: Dr. Karcagi Veronika PhD. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/hungarian/carrier_testing.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2007. *Kromoszóma rendellenességek.* Pécs: Orvosgenetikai Intézet, Pécsi Tudományegyetem. Fordította: Dr. Komlósi Katalin. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/hungarian/chromosome_changes.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2009. *Kromoszóma transzlokációk*. Pécs: Orvosgenetikai Intézet, Pécsi Tudományegyetem. Fordította: Dr. Komlósi Katalin. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/hungarian/chromosome_translocations.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2007. *Mi az a genetikai teszt?* Pécs: Orvosgenetikai Intézet, Pécsi Tudományegyetem. Fordította: Dr. Komlósi Katalin. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/hungarian/genetic_test.pdf

Guy's and St Thomas' Hospital, the Royal College of Obstetricians and Gynaecologists, London IDEAS Genetic Knowledge. 2009. *Recesszív öröklődés*. Pécs: Orvosgenetikai Intézet, Pécsi Tudományegyetem. Fordította: Dr. Komlósi Katalin. http://www.eurogentest.org/fileadmin/templates/eugt/leaflets/pdf/hungarian/recessive_inheritance.pdf

International Society for Stem Cell Research. 2008. *Az Őssejtkutatások Nemzetközi Szervezetének betegtájékoztatója az őssejtterápiás lehetőségekről*. Fordította: [origo] Egészség. http://www.angyalszarnyak.hu/files/cikk73_ossejt_handbook.pdf

Internetes források

http://www.webbeteg.hu/keresok/bno/13403

https://www.ipsennordic.com/en/therapeutic-areas/neurosciences/cervical-dystonia/

Szógyakoriság korpuszalapú vizsgálata: autentikus és célnyelvi magyar szórakoztató irodalmi szövegek összevetése

Dankó Szilvia

szilvia.danko@yahoo.com

ELTE Nyelvtudományi Doktori Iskola

Fordítástudományi Doktori Program

Kivonat: Minden nyelvben található olyan egyedi nyelvi elem, amelynek nem létezik ekvivalense valamely másik nyelven, ezért kihívás elé állítja a fordítót. Chesterman (2007) javaslata szerint a kutatásoknak az „egyedi" nyelvi elemeket (Tirkkonen-Condit 2002) nem a priori kellene meghatározni, hanem a fordított és a nem fordított szövegek korpuszalapú elemzésének eredményeképpen létrejövő gyakorisági minta alapján megkapni őket – a posteriori. Amely szavak előfordulási gyakorisága a legnagyobb különbséget mutatja a fordított korpuszban, azok lesznek az alulreprezentált, lehetséges „egyedi" nyelvi elemek, és lesznek felülreprezentált elemek. Ezen elemek eloszlására, nyelvi jellemzőire és a fordítási folyamatban játszott szerepükre kell magyarázatot keresni. E módszert követve készítettem el kutatásomat a *Pannónia Korpusz*ból (Robin et al. 2016) válogatott szórakoztató irodalmi szövegeken. A szógyakorisági vizsgálat eredménye több lehetséges „egyedi" nyelvi elemet is magában foglal, amelyeket további célzott elemzésekkel érdemes vizsgálni.

Kulcsszavak: „egyedi" nyelvi elem, előfordulási gyakoriság, korpuszalapú elemzés, (forrás)nyelvi inger, Pannónia Korpusz

Dankó Szilvia: Szógyakoriság korpuszalapú vizsgálata: autentikus és célnyelvi magyar szórakoztató irodalmi szövegek összevetése. In: Robin E., Seidl-Péch O. (szerk.) 2020. *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordításkutatás Magyarországon.* Segédkönyvek a nyelvi közvetítésről I. Budapest: ELTE BTK Fordítástudományi Doktori Program, MANYE Fordítástudományi Szakosztály. DOI: https://doi.org/10.36252/Nyelvikozvsegedkonyv1.7

1. Bevezetés

A fordított szövegek sajátos nyelvezetének feltárása a fordítástudomány egyik fontos feladata. Számos kutatás vizsgálta a fordítási univerzálékat (Baker 2003) és a fordított szövegek sajátosságait (Klaudy 1999a). Az elektronikus korpuszok és szövegtárak létrehozásával lehetővé vált nagy mennyiségű szöveg számítógépes kutatása, statisztikai összevetése és elemzése többféle szempontból (vö. Seidl-Péch 2018). Ide tartozik az úgynevezett „egyedi" nyelvi elemek vizsgálata is, amelyet nyelvenként és nyelvpáronként számos kutató vizsgált. Kutatási célom volt, hogy angol – magyar irányban találjak ilyen nyelvi elemeket.

A finn nyelv „egyedi" nyelvi elemeiről először Tirkkonen-Condit (2004, 2005) értekezett, miután bizonyossá vált, hogy bizonyos, a finnre jellemző nyelvi elemek előfordulása ritkább a fordított szövegekben. Sari Eskola (2004) összehasonlító kutatása az interferencia hatásának tudja be, amikor a fordításokban a célnyelvre jellemző egyedi nyelvi elemek alulreprezentáltak, míg a kézenfekvő, közvetlen ekvivalenssel bíró (mintegy stimulust kiváltó) elemek felülreprezentáltak. Kutatásában a fordított finn nyelvű szövegek több nem tipikus ismétlődést tartalmaztak, mint az eredetileg finn nyelven írt szövegek. Laviosa (1998, 2000) egynyelvű, összehasonlítható korpuszán végzett kutatási eredményei arra mutatnak rá, hogy a gyakori és kevésbé gyakori kifejezések aránya a fordításokban magasabb, a gyakori szavak pedig sűrűbben ismétlődnek. A fordításokban a célnyelvre jellemző „egyedi" nyelvi elemek alulreprezentáltak, míg a közvetlen ekvivalenssel bíró, nyelvi ingert kiváltó elemek felülreprezentáltak. Frankenberg-Garcia (2008) korpuszalapú kutatásában a fordított szövegekben előforduló alul- illetve felülreprezentált lexikai elemeket kereste eredeti és fordított portugál nyelvű szövegrészekben. Az „egyedi" nyelvi elemeken túl számos olyan kifejezés is az alulreprezentált csoportba került, amelyeknek létezik ekvivalense a forrásnyelvben. Ilyen esetben érdemes volna vizsgálni annak az okát, hogy a fordítók és lektorok miért nem használják ezeket a szavakat, a nyelvi norma játszik szerepet ebben vagy a nyelvi babona okán igyekeznek ezeket mellőzni? Götz (2017) a diskurzusjelölők közül a *vajon* jellemzőit, szerepét és eloszlását vizsgálta a fordított szövegekben. Még forrásnyelvi inger esetén is az esetek közel felében

kihagyták a fordítók a *vajon*-t, és nem került bele a fordításba, vagyis alulreprezentált maradt, így egyike a lehetséges „egyedi" nyelvi elemeknek.

2. A felülről indított Chesterman-féle módszer

Korábbi kutatásomban egy bizonyos „egyedi" nyelvi elemet („szokott") előre kiválasztottam, és megvizsgáltam, hogy a gyakorisági előfordulása alátámasztja-e az egyedi elemek (Tirkkonen-Condit 2002) hipotézist, amely szerint az „egyedi" nyelvi elemek fordított szövegekben ritkábban fordulnak elő. Autentikus és fordított szórakoztató irodalmi szövegek és szakszövegek együttes vizsgálata alapján az összes fordított szövegben mért előfordulási érték fele annyi volt, mint a magyar szerzők szövegalkotása során mért érték. Chesterman (2007) cikkében a hipotézist vizsgáló kutatások módszertanának gyengeségeiről ír. Véleménye szerint Tirkkonen-Condit (2004, 2005) kutatásaiban az adott nyelvpárok kontrasztív nyelvészeti elemzésével kiválasztott „egyedi" nyelvi elemek képezik a gyakorisági vizsgálatok alapját. Tehát az „egyedi" nyelvi elemek a priori kerülnek meghatározásra, és nem a kutatás eredményeképpen választódnak ki. Nyilvánvaló, hogy a kutatónak tudnia kell, mi az, amit keres egy adott szövegben. Ennek az ellentmondásnak a feloldására Chesterman az alábbi módszertani folyamatot javasolja a további vizsgálatokhoz (2007: 12):

1) Állapítsuk meg kontrasztív korpuszelemzések alapján, melyek azok az elemek, amelyeknek gyakorisága jelentősen eltér fordított és nem fordított szövegekben.

2) Válasszuk szét ezeket az elemeket aszerint, hogy a fordításokban vagy a nem fordítás útján keletkezett szövegekben fordulnak elő gyakrabban.

3) Tegyük sorrendbe ezeket a nyelvi elemeket gyakoriságuk szerint. Összpontosítsunk azokra az elemekre, amelyeknek előfordulási gyakorisága a legnagyobb különbséget mutatja a két korpuszban (fordítások és eredeti szöve-

gek). Az első csoportba azok az elemek tartoznak majd, amelyek a vizsgált fordításokban felülreprezentáltak, míg a második csoportban azok az elemek kapnak helyet, amelyek alulreprezentáltak.

4) Végezetül vizsgáljuk meg a fenti eredmények lehetséges magyarázatait. Az alulreprezentált csoportban megtaláljuk azokat a kifejezéseket is, amelyeket Tirkkonen-Condit (2002) „egyedi" nyelvi elemeknek nevezett. Előfordulhatnak azonban olyan elemek is, amelyek ebben az értelemben nem számítanak „egyedinek".

Ez a módszertani megközelítés nem az „egyedi" nyelvi elemek fogalmából indul ki. A fent ismertetett folyamat szerint egy adott elem egyedisége a vizsgálat során kerül meghatározásra, a posteriori. Az „egyediség" ilyen értelemben mint lehetséges magyarázat kínálkozik bizonyos nyelvi elemek alulreprezentáltságára.

3. A kutatási hipotézis és módszer

Korábbi kutatások alapján feltételeztem, hogy vannak olyan nyelvi elemek, amelyek disztribúciója jelentősen eltér a fordított és az autentikus szövegekben. Hipotézisem szerint a fordításokban alulreprezentált nyelvi elemek között szerepelnek az úgynevezett célnyelvspecifikus „egyedi" elemek. A fenti módszert alkalmazva reményeim szerint feltárhatóak a magyar nyelv „egyedi" elemei angol – magyar viszonylatban. A kutatási kérdésem, hogy mely nyelvi elemek fordulnak elő kisebb gyakorisággal a fordításban, mint az autentikus szövegalkotásban? Melyek az alulreprezentált elemek, avagy a lehetséges „egyedi" nyelvi elemek. Amennyiben sikerül „egyedi" nyelvi elemet találni, a fordítók céltudatosabban használhatják azokat, és nagyobb figyelmet kaphatnak ezek az elemek a fordítóképzés során. További kérdés, hogy lesz-e olyan alulreprezentált elem, melynek létezik ekvivalense az autentikus szövegekben, vagyis mégsem tekinthető „egyedi" nyelvi elemnek?

A szógyakorisági listában szerepelni fognak gyakran ismétlődő elemek, avagy felülreprezentált elemek. Kérdésként merül fel, hogy milyen nyelvi elemek ezek? Hipotézisem szerint előfordulási gyakoriságukat befolyásolhatja a kiválasztott műfaj, a forrásnyelvi inger vagy esetleg az adott korszak. További kérdés, hogy lehet-e ezekre az elemekre vonatkozóan valamilyen következtetést levonni?

A kutatást Chesterman (2007) módszere szerint a Laurence Anthony-féle online korpuszelemző eszközzel a következő lépések szerint végeztem:

1) Szógyakorisági listát hívtam le mindkét szövegkorpuszon.
2) Alfabetikus sorrendben egymás mellé helyeztem a fordított és a nem fordított gyakorisági listákat.
3) Megnéztem, hol mutatkozik nagy eltérés a gyakoriságban az autentikus szövegekhez képest (mindkét irányban alulreprezentált, illetve felülreprezentált elemek).
4) Lehetséges magyarázatot kerestem az eredmények értelmezéséhez.

4. A kutatási korpusz

A vizsgált kortárs szórakoztató irodalmi szövegek a *Pannónia Korpusz*ból származnak. A párhuzamos alkorpuszból kiválasztott 10 angolról magyarra fordított regény 906 728 szövegszót tartalmaz. Az összehasonlítható alkorpuszból letöltött 10 magyar nyelven írt könyv 865 194 szövegszóból áll. A szövegek átlagos hossza 60 000 – 80 000 szó között van, a könyveket öt különböző könyvkiadó adta ki. Összesen 7 magyar szerző és 9 angol szerző (9 fordító által fordított), túlnyomó részben női szerzők és fordítók műveinek az elemzése történt meg. Azért esett a szórakoztató irodalmi műfajra a választás, mert a szövegalkotási normát tekintve ez áll legközelebb a művelt köznyelvhez, továbbá nem köti sem szépirodalmi konvenció, sem szaknyelvi terminológia. Tartalmaz mind írásbeli, mind szóbeli nyelvhasználatot, tágabb és kreatívabb mozgásteret nyújt a fordítóknak, mint más

műfajok, szövegtípusok. A nyelvi megformálás és a kreativitás szempontjából ezek a legváltozatosabb, legérdekesebb szövegek. Ilyen módon kiváló alapul szolgálnak a fordítás általános sajátosságait vizsgáló empirikus kutatásokhoz.

5. Kutatási eredmények

A szövegeket digitális (*txt*) formátumban a Laurence Anthony-féle *AntConc* online korpuszelemző eszközzel, annotálás nélkül dolgoztam fel a korábbiakban részletesen ismertetett Chesterman-féle módszerrel. A program lefuttatta az összes szövegszó (kb. 1,8 millió szövegszó, fele-fele arányban autentikus magyar és fordított magyar korpusz) szógyakorisági listáját. Ez mindkét korpusz esetében kb. 95 000 szóból álló gyakorisági listát eredményezett a leggyakrabban előforduló szavakkal kezdve sorrendben a legkevésbé gyakori szavakig. Ezt a listát rövidebbé kellett tenni, hogy lexikailag feldolgozható legyen. Mindkét korpuszban kb. 55 000 olyan szó volt, mely egyetlen egyszer fordult elő, ami az összes szövegszó alig 6 százaléka. Az összes szövegszó 90%-át tették ki azok a szavak, amelyek a korpuszokban kevesebb, mint tízszer fordultak elő, azaz egyenletes eloszlást feltételezve minden elemzett műben mindössze egyszer. A maradék 10%-ból mindössze 1% azoknak a szavaknak a száma, amelyek legalább 100 alkalommal fordultak elő. A gyakorisági lista elején szereplő 300 szó eredményezte az előfordulások majd felét (kb. 419 000 előfordulás). Az elemzést ezen a mennyiségen volt ideális elvégezni, ezért a listában itt húztam meg a határt. Ezután a tulajdonneveket kivettem a listából, mert azok bár egyediek, de nem „egyedi" nyelvi elemek. Az autentikus szövegekben szereplő szavakat és a fordított szövegekben előforduló szavakat alfabetikus sorrendben egymás mellé tettem, majd összevetettem őket. Mindkét listában voltak olyan szavak, amelyek a másik listából hiányoztak, mert ott nem kerültek be az első 300 szóba. Így ezeknél a szavaknál jelentős volt a disztribúciós eltérés. Az eredmény teljesebbé tételéért – hogy legyen ismét legalább 300 szó –, ezeket a szavakat bent hagytam mindkét listában, és a hiányzó adatokat beemeltem a másik korpusz gyakorisági listájából. Az *1. táblázatban* látható szavak gazdagítják az alulreprezentált szavak listáját, mert habár a fordításokban ritkán szerepeltek, az autentikus szövegekben nagyon gyakori volt az előfordulásuk.

Fordításokban ritkábban szereplő szavak	Fordított szövegekben az előfordulás száma	Autentikus szövegekben az előfordulás száma
próbáltam	91	514
néztem	80	499
kezdtem	75	328
hiába	93	327

1. táblázat

Az alulreprezentált szavak listájának kiegészítése

A másik irányban a felülreprezentált szavak kiegészítése hosszabb listát eredményezett, ahogy a *2. táblázatban* látható. A számokból jól látható, hogy ezek a szavak a fordításokban gyakran előfordulnak, de az autentikus szövegekben nem – ott az előfordulásuk jóval 300 alatt van. Ez a táblázat azért lett hosszabb, mert a fordított szövegkorpuszban sok tulajdonnév szerepelt a gyakorisági lista elején. A lenti szavak a személynevek helyett kerültek be az összehasonlítási listába.

2. táblázat

A felülreprezentált szavak listájának kiegészítése

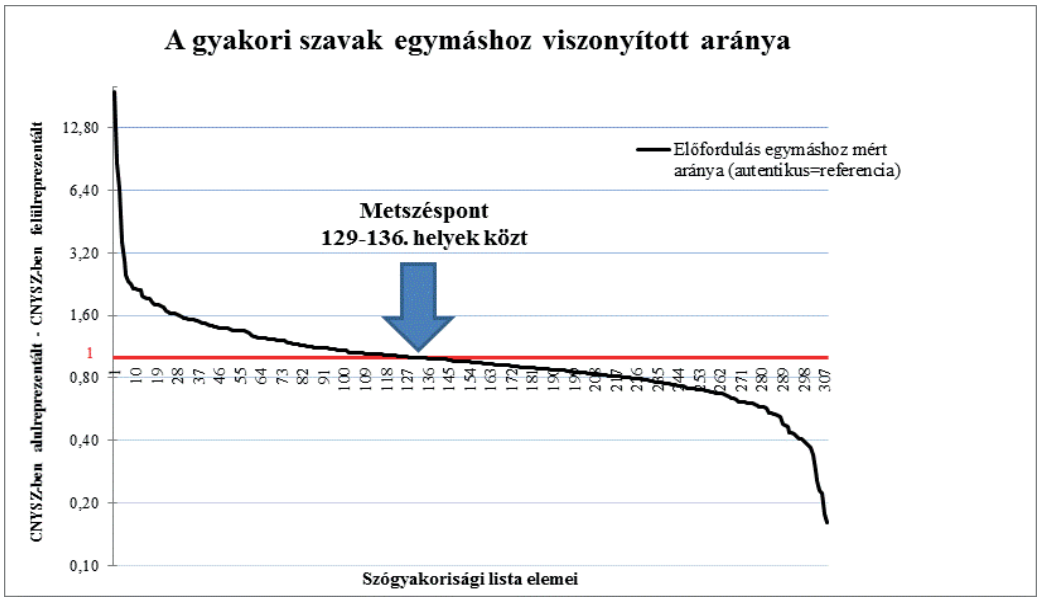
Fordításokban gyakran ismétlődő szavak	Fordított szövegekben az előfordulás száma	Autentikus szövegekben az előfordulás száma
felelte	726	246
szeme	585	251
amely	584	106
akár	514	260
tette	507	235
gondolta	482	133
vajon	481	261
szemét (*valakinek a szemét*)	434	197
apám	349	54
fölött	325	154
keresztül	316	126
látott	312	138
válaszolta	306	16

Ellenőrzés céljából és a kettős kódolás kiváltására lefuttattam a korpusz elemzését a *Wordsmith Tools 6.0* szövegelemző szoftverrel is. Összevetettem a szavak sorrendjét és az előfordulási értékeket minden listában, hogy meggyőződjek a kapott eredmény helyességéről. Mindkét szoftver szógyakorisági listája egyező előfordulási listát eredményezett. Az elemzésnél az autentikus magyar szövegek szógyakorisági listáját vettem alapul minden szónál. Ez a lista lett a referenciapont. A fordított szövegekben mért előfordulási számot viszonyítottam ehhez a listához, hogy megkapjam az arányt a kettő között. Amennyiben egy adott szó előfordulási gyakorisága közel azonos volt mindkét korpuszban, akkor az egyezés az egyes értéket adta. A kapott arányszámokból egy grafikont készítettem, mely az 1. ábrán látható. A metszéspont az egynél meghúzott referenciaértéket a 129–136 közötti helyen szereplő szavaknál található. Ez a nyolc szó ugyanannyiszor fordult elő mindkét korpuszban. Amikor egy szó gyakrabban fordult elő a fordított szövegekben, mint az autentikus szövegekben, akkor az arányszám egyesnél nagyobb lett. Összesen 128 szó fordult elő többször a fordításban, mint az autentikus szövegekben. Ezek a felülreprezentált, gyakori szavak. Amikor egy szó kevesebbszer fordult elő fordított szövegben, mint az autentikusban, akkor az eltérés mértéke egy nulla és egy közötti szám. Mivel a lista véges volt – 307 szó –, a kapott arányszám nem ment 0,16 alá. Ezen alulreprezentált szavak listája adta a maradék 171 szót.

A vizsgálatban szereplő szavak megoszlása alul- illetve felülreprezentáltság tekintetében látszólag szimmetrikus (lásd *1. ábra*). Az elemzés eredménye alapján mindkét korpuszban hasonló a szógyakorisági tendencia, az alul- és felülreprezentált szavak megoszlási aránya majdnem azonos az autentikus és a célnyelvi szövegprodukció esetén a vizsgált szövegtípusra, vagyis a szórakoztató irodalomra vonatkozóan. Az *1. ábrán* a vastag fekete görbe íve érzékelteti az arányszámok eloszlását. A görbe közepe laposan, hosszan elnyúlik, ezen szavaknál nem volt nagy eltérés a szógyakoriságban, mindkét korpuszban hasonló gyakorisággal szerepeltek. Ezeket nem tekinthetjük sem felül-, sem alulreprezentált szavaknak. A görbe mindkét vége rövid szakaszon meredek, a fekete vonal balról induló ereszkedő szakaszán találjuk a lehetséges felülreprezentált szavakat, a jobbra lefelé görbülő szakaszát – az egyes érték alatt – a lehetséges alulreprezentált szavak csoportja teszi ki.

1. ábra

A gyakori szavak egymáshoz viszonyított aránya



Az *1. ábra* legmeredekebb szakaszainál tapasztalható a legnagyobb eltérés az arányszámban. Ezek a szavak tűntek izgalmasnak a kutatás szempontjából, ezeket kellett részletesen elemezni a fenti szimmetria tükrében. A görbe két meredek szakaszát kitevő szavak számát 100 szóban határoztam meg. Ahhoz, hogy egymással összevethető legyen a görbe két végén található 50-50 szó előfordulási arányszáma, a jobb oldalt kitevő alulreprezentált szavaknál meg kellett fordítani a viszonyítási alapot. Azaz a fordított szövegekben mért előfordulási értéket alapul véve kellett kiszámítani az arányt az autentikus szövegekben mért előfordulási értékekhez viszonyítva. Így ezeknél is egynél nagyobb lett az arányszám. Az elemzést bizonyos határok között kívántam tartani, ezért a listát tovább szűkítettem. Mindkét szógyakorisági lista elejéről kiválasztottam azokat a szavak, amelyeknél a szó előfordulása majdnem kétszerese a másik szövegben való előfordulásnak. Így 20-20 szó maradt a végső listában (*3. táblázat* és *4. táblázat*).

A célnyelvi szövegben sűrűn előforduló szavakat a *3. táblázatban* láthatjuk. A felülreprezentált szavakat a fordítók gyakrabban használják, mint a magyar szerzők eredeti

szövegalkotás során. Miért éppen ezeket a szavakat részesítik előnyben a fordítók, mi ennek az oka? Fontos megjegyeznem, hogy a kiegészítéssel beemelt szavak mind bekerültek ebbe a rövid listába, mert a két szövegkorpuszban mért előfordulási gyakoriságuk annyira eltért egymástól.

3. táblázat

A felülreprezentált szavak statisztikai megoszlása

Fordításokban gyakran ismétlődő szavak. A 300-as lista első 20 szava	**Fordított szövegekben az előfordulás száma**	**Autentikus szövegekben az előfordulás száma**	**Előfordulási arányszám felülreprezentált elemek**
válaszolta	306	16	19,13
amely	584	66	8,85
apám	349	54	6,46
gondolta	482	133	3,62
felelte	726	246	2,95
keresztül	316	126	2,51
szeme	585	251	2,33
látott	312	138	2,26
tette	507	235	2,16
fiú	926	430	2,15
szemét	450	212	2,12
fölött	325	154	2,11
akár	514	260	1,98
kezét	811	413	1,96
mondta	2348	1220	1,92
látta	644	336	1,92
vajon	481	261	1,84
kérdezte	1084	600	1,81
éppen	877	488	1,80
lány	1196	666	1,80

A szavakat nem egyenként, hanem logikailag összetartozó szóbokrokba, csoportokba rendeztem, hogy könnyebben találjak magyarázatot a gyakori használat okára. Az első helyen szereplő szó a *válaszolta* együtt vizsgálható a *felelte, mondta, gondolta* szinonimákkal. Ezek a szavak felcserélhetőek a nyelvi norma szempontjából. Ide sorolhatjuk még a *kérdezte* ellentétes jelentésű igét is. Klaudy kutatása (2001) az idéző igékről bizonyította a fordítás során végbemenő konkretizációt, a fordított szövegek gazdagítását. A fordító explicitál a forrásnyelvi szöveg monoton *said,* esetleg *asked* szavainak állandó ismétlése helyett. A magyar irodalmi nyelvhasználati norma érvényesüléseként értelmezhetjük ezen gyakori szavak variációit. Az autentikus szövegekben való hiányuk alapján feltételezem, hogy a magyar szerzők további szinonimákat alkalmaznak széles körűen, vagy teljesen kihagyják ezeket az igéket a párbeszédes részekben. Párhuzamos korpuszon végzett kontrasztív elemzés kimutathatná a forrásnyelvi szavak arányát: vajon mennyivel többször fordul elő az angol szövegekben a *said* és *asked,* mint a magyar fordított szövegekben a *mondta* és *kérdezte*?

A második helyen szereplő *amely* vonatkozó névmás magas előfordulása meglepett, mert felbukkanását az alulreprezentált szólistában vártam. Miért ennyire gyakori a fordításokban? Kikerestem az *ami* disztribúciós értékét, tekintve, hogy ez a vonatkozó névmás is a mellékmondatok fűzését teszi lehetővé, de mindkét korpuszban szinte azonos volt az előfordulási aránya. Az egyik lehetséges magyarázat a gyakori előfordulás indoklására az, hogy az angol forrásnyelvi szöveg több vonatkozói mellékmondatot tartalmaz, mint a magyar irodalmi szöveg, és ezeket le kell fordítani. Másik lehetséges magyarázat, hogy a grammatikai felbontások során keletkezik több vonatkozói mellékmondat a célnyelven, és ezeket kell összekapcsolni egy alkalmas kötőszóval. Elképzelhető az is, hogy lektorálás során nő meg ennyire a számuk a fordított szövegekben, amikor a lektorok a célnyelvi norma, nyelvhelyesség és stílus alapján, tartalmilag pontos, nyelvileg helyes, az eredeti szöveggel egyenértékű, (Robin 2015), az irodalmi nyelvezetnek megfelelő szöveg létrehozásához az *amely* vonatkozó névmás nagyobb arányú alkalmazásával kívánják javítani a szöveget.

A következő leggyakoribb szó a fordításokban az *apám* szó. Feltehetően konkretizáció során nőtt meg a száma a fordításokban. A *fiú* és *lány* szó hasonlóképp szükséges konkretizáló betoldások az angol hímnemű és nőnemű személyes névmások (*he/she*) kiváltására. Az eddigi kutatások a névmások fordításával kapcsolatban az explicitációt hangsúlyozták, tehát a főnévi szócsoporttal történő helyettesítést, illetve a nyelvtani nem kérdését helyezték vizsgálódásuk középpontjába. Az angol – magyar fordítási irányra az implicitáció jellemző: az esetek 70-80%-ában az angol névmásokat nem fordítjuk magyarra (Heltai és Juhász 2002). A névmások fordítása során szükségessé váló átváltásokat a fakultatív átváltások közé, vagy külön átmeneti kategóriába sorolhatjuk, amelyben a forrásnyelvi és a célnyelvi nyelvhasználat eltérései olyanok, hogy – bár nem kötelező az átváltás – az esetek túlnyomó részében mégis sor kerül rá.

További csoportot alkotnak a listában a *szeme, szemét* és *kezét,* szavak. Klaudy (1999b) az átváltási műveletek kapcsán beszél a kötelező és a fakultatív betoldásokról. A különböző testrészekkel végzett cselekvések esetén pl. int, megsimogat, megszorít, a magyar nem tartja szükségesnek a testrész megnevezését, az angol viszont igen. Forrásnyelvi inger hatására lexikai redundancia eredményeképp kerülhetnek a fordításba. A következő példa bizonyítja a felesleges betoldást:

(1) His eyes had passed over her, empty; then sharpening and momentarily confused. (Cashore 2012: 38)

(1a) A bíró *szeme* üresen bámult el mellette, majd amikor észrevette Keserkéket, zavartság jelent meg benne. (Cashore 2012: 30)

További kvalitatív vizsgálat során kellene elemezni, hogyan lehetne csökkenteni ezen szavak használatát. A diskurzusjelölő szavak *akár, vajon, éppen* használata összetett képet mutat. Götz (2016) összegzése szerint a diskurzusjelölők fordítását műfaji, nyelvi, fordítástípusra jellemző normák befolyásolhatják, ami megkérdőjelezi, hogy a fordított szövegben mért adatok kizárólag a forrásnyelvi elemek funkciójára reflektálnának. Saját

kutatásában azt találta, hogy a legtöbb esetben a párhuzamos korpusz célnyelvi vagy forrásnyelvi eleme (*after all, but, because, Ø*) nem feleltethető meg a *vajon*-nak. A kérdés az, hogy mit fordítanak, vagy nem fordítanak ilyenkor a fordítók? A fordítói döntések szerepe sem elhanyagolható a fordított szövegek diskurzusjelölőinek szempontjából, Klaudy (2012) magyarázata szerint a diskurzusjelölők változatosságának különbségeit a fordítók eltérő magyar nyelvi normakövetésére is vissza lehet vezetni, akik a célnyelvi befogadók feltételezett igényeinek szem előtt tartásaként toldanak be, vagy ritkábban hagynak ki különféle kötőelemeket. A betoldás megfelel az explicitációs tendenciának. Kötőelemeket a *4. táblázatban* is találunk, ezért elemzésükre ott ismét visszatérek.

A *keresztül* és a *fölött* névutók túl gyakori használata terjengőssé teheti a szöveget, ezért szerettem volna megvizsgálni, hogy a fordításokban található magas számuk milyen összefüggésben van az explicitációval. Összevetettem ezeknek a szavaknak az eloszlási statisztikáját mindkét szövegkorpuszon szövegenként. Az autentikus szövegekben a *keresztül* összesen 126 előfordulásából 39 szerepel az egyik szövegben. Ez az összes előfordulás egy harmadát teszi ki, míg a korpusznak csak az egy tizedét. A másik esetben hasonlóképpen, az autentikus szövegekben lévő 154 darab *fölött* előfordulásából 79 egyetlen műben szerepel. Ez az előfordulások fele. Kijelenthető, hogy a *keresztül* és a *fölött* névutót egy-egy magyar szerző használja előszeretettel, míg a magyar szerzők többsége nem. A fordított szövegekben az eloszlás a *fölött* esetében teljesen kiegyenlített volt mind a tíz műben, tehát nem tekinthető túlreprezentáltnak egyikben sem. A *keresztül* esetében az összes előfordulás egy negyede szintén egyetlen szöveghez köthető, bár ugyanennek a fordítónak a másik munkájában nem volt magas ez a szám. Kvalitatív vizsgálatra volna szükség ahhoz, hogy eldőljön, vajon indokolt-e ennyi *fölött* és *keresztül* névutó használata a fordításokban?

A következő lista a fordítás során ritkán előforduló szavakat tartalmazza. Az ilyen elemek nem aktiválódnak a fordító mentális lexikonjában, részben forrásnyelvi inger hiánya folytán, részben a fordítói kompetencia korlátai miatt. A *4. táblázatban* szereplő szavak a lehetséges alulreprezentált, azaz a fordított szövegekben ritkán szereplő szavak.

4. táblázat

Az alulreprezentált szavak statisztikai megoszlása

Fordításokban ritkábban szereplő szavak. A 300-as lista utolsó 22 szava	Fordított szövegekben az előfordulás száma	Autentikus szövegekben az előfordulás száma	Előfordulási arányszám alulreprezentált elemek
néztem	80	499	6,24
próbáltam	91	514	5,65
éreztem	260	1160	4,46
kezdtem	75	328	4,37
kérdeztem	102	400	3,92
hiába	93	327	3,52
akartam	281	829	2,95
tudtam	490	1321	2,70
vettem	121	318	2,63
rám	558	1431	2,56
eddig	164	409	2,49
láttam	386	955	2,47
magam	478	1173	2,45
hozzám	136	326	2,40
mosolygott	130	304	2,34
voltam	460	1053	2,29
hallottam	236	540	2,29
hiszen	310	667	2,15
velem	288	611	2,12
azonnal	286	596	2,08

Itt is érdemes csoportokra bontani az eredményt. Kezdjük az egyes számban szereplő igékkel a lista elején: *néztem/éreztem/próbáltam/kezdtem*. Nem találtam pontos nyelvészeti magyarázatot arra, hogy miért több az autentikus szórakoztató irodalomban az érzékelést kifejező ige, mint a fordított szövegekben. Arra gondoltam, hogy a műfaj

személyes hangvétele indokolhatja, amit a szerzők nyelvileg kifejezőbbnek vélnek, mint a fordítók, de ezek a szavak nem tekinthetők „egyedi" nyelvi elemnek.

A kötőszavak, diskurzusjelölők közül a *hiába* és *hiszen* fordítása jelent problémát a fordítók számára. Mindkettő lehet kötőelem, töltelékszó és interperszonális viszonyokra utaló diskurzusjelölő (pl. *Hiszen te is ismered őt! Hiába beszélek neki!).* A *hiszen* lexikai jelentése szűk, és mint a diskurzusjelölőknek általában, inkább procedurális jelentése van, azaz a megnyilatkozások feldolgozásához nyújt eligazítást, ill. relevancia-megszorításokat kódol, korlátozva a levonható releváns következtetéseket (Heltai és Juhász 2002). Gyuris (2008: 657) szerint a *hiszen* a közös ismeretekre utaló, az információ ismertségét jelző kontextusjelölők közé tartozik. Ha a diskurzusjelölőket kifejezetten a fordítás szempontjából vizsgáljuk, akkor a következő kérdések merülnek fel Heltai szerint (Heltai 2016: 10):

1) Tekintve a diskurzusjelölők szerteágazó, egymást gyakran átfedő, nehezen azonosítható és parafrazálható jelentéseit és funkcióit, képes-e a fordító (ha nem anyanyelve a forrásnyelv) pontosan azonosítani, mi a szerepe a diskurzusjelölőnek az adott mondatban (szövegrészben)?

2) Van-e az adott jelentésben és funkcióban a diskurzusjelölőnek egyértelmű megfelelője a célnyelven, illetve a célnyelvi megfelelők közül melyeknek a jelentései (funkciója) áll hozzá legközelebb az adott kontextusban?

3) Ha nincs jó célnyelvi megfelelő, mennyire releváns az adott diskurzusjelölő által hordozott információ? Ha igen, érdemes-e explicitáló megoldást alkalmazni? Ha nem, milyen veszteséget okoz a kihagyás?

Úgy tűnik, hogy mivel a propozicionális jelentés a diskurzusjelölő fordítása nélkül is átvihető, pontos célnyelvi megfelelő hiányában a kihagyás tűnik célszerűbbnek, vagy egyéb más diskurzusjelölőt használnak a fordítók. A hiányt részben az is indokolhatja, hogy az angol általában kevesebb diskurzusjelölőt használ (Heltai 2016). Mivel ezeknek

az elemeknek a fordítása és használata kihívást jelent, indokoltan tekinthetjük őket lehetséges „egyedi" elemeknek.

Külön csoportot képeznek a személyes névmás egyes szám első személyű ragozott formái: *rám, hozzám, velem* és a visszaható formájú *magam*. A korpuszvizsgálatok ezen a területen tapasztalt disztribúciós különbségei eredeti magyar és fordított magyar szövegek között nem feltétlenül az implicitációt bizonyítják, hanem az ingadozó magyar nyelvhasználatot. Lehetséges, hogy a szerkesztők, lektorok ügyelnek a redundanciára, és igyekeznek törölni a felesleges elemeket (Robin 2015). Továbbá Heltai (2016) megállapítása szerint a képzett fordítók által készített és lektorált fordítások a nyelvi megformálás szempontjából sokszor jobb minőségű szövegek, mint egyes eredeti művek, akár irodalmi, akár szakfordításra gondolunk. Mindemellett ezeket az elemeket érdemes tovább kutatni más szövegtípusokon is, és feltárni azokat az eseteket, amelyek bizonyíthatják egyediségüket.

Cappelle és Loock (2013) kutatása bizonyította a *van* létige disztribúciós különbségét az autentikus francia szövegalkotás és a fordítás között. Jelen listában egy múlt idejű alak a *voltam* szó szerepel. Fordított szövegekben jelentősen kevesebbszer fordul elő, mint az autentikus szövegekben. Ennek oka lehet fordítói stratégia vagy lektori döntés, hátterében az ismétléskerülés állhat. Mivel egyszerűen fordítható elem, nem tekinthető lehetséges „egyedi" nyelvi elemnek. Feltételezésem szerint a fordítás során explicitáció valósul meg, és alkalomadtán kicserélik egy érzékletesebb, színesebb igére.

Az *akartam/tudtam/láttam/hallottam* és *mosolygott* igékre vonatkozóan hasonló megállapítást tehetünk, mint a lista elején található egyes szám első személyben szereplő igékre vonatkozóan. Eloszlásuk lehet véletlen, vagy inkább műfajspecifikus. Mindegyiknek létezik kézen fekvő ekvivalense a forrásnyelven, ezért nem lehetnek lehetséges „egyedi" nyelvi elemek annak ellenére, hogy az alulreprezentált szavak csoportjába kerültek.

A Chesterman-féle módszer alkalmazásával létrejött szógyakorisági listák elemzése bizonyítja a hipotézist, amely szerint vannak alulreprezentált nyelvi elemek. A lista alapján azonosíthatóak voltak a felülreprezentált elemek is. A módszer alkalmas volt a szavak eloszlásának vizsgálatára, a disztribúciós különbségek feltárására. Az alulrepre-

zentált elemek részletesebb elemzése során találtam lehetséges „egyedi” nyelvi elemet a diskurzusjelölők, kötőszavak között, de bizonyításukhoz nagyobb szövegkorpusz elemzése szükséges. Nem sikerült bizonyítanom a korábbi kutatásomban szereplő „szokott” elem egyediségét főként azért, mert az elemzett korpusz egyetlen szövegtípusra korlátozódik.

6. Összegzés

Kutatásomban szórakoztató irodalmi korpuszon, 10-10 autentikus és fordított regényen végeztem szógyakorisági vizsgálatot. A kiválasztott szövegekben Chesterman módszerével feltártam a ritka és gyakori szavakat. Az alul-, illetve felülreprezentált nyelvi elemek eloszlása mindkét korpuszban hasonló tendenciát mutatott. Számos olyan nyelvi elem akadt – a kutatási hipotézissel egyezően –, amelynek disztribúciója jelentősen eltért az autentikus és a fordított szövegekben. A kutatási kérdésre, hogy milyen nyelvi elemek fordulnak elő nagyobb, illetve kisebb gyakorisággal a fordításban, mint az autentikus szövegalkotásban a válasz nem egyértelmű. Mindkét gyakorisági irányban találtam diskurzusjelölőket, ezek közül a *hiába*, és *hiszen* lehetnek „egyedi” nyelvi elemek. További kutatás szükséges a bizonyításhoz. A *szokott,* mint lehetséges „egyedi” nyelvi elem nem került be a vizsgált szógyakorisági listába. Korábbi kutatásom alapján tapasztaltam, hogy disztribúciója ebben a szövegtípusban nem mutat jelentős eltérést.

Összegzésként elmondható, hogy a Chesterman-féle módszer működik a magyar nyelv vonatkozásában is, érdemes és könnyű alkalmazni. A kutatás egyetlen szövegtípuson történt, közel 1,8 millió szövegszón, összehasonlítható autentikus és fordítási korpuszon. Kibővített korpuszon más szövegtípusok bevonásával feltétlenül érdemes megismételni, mert az „egyedi” nyelvi elemek keresése így hatékonyabb, mint a priori kiválasztott egy-egy elem esetén. További kutatásként érdemes volna összehasonlítani a fordított és nem fordított célnyelvi szövegeket kvalitatív módon, vagy a fordított szövegeket kontrasztívan elemezni kétnyelvű, párhuzamos fordítási korpuszon.

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: Benjamins. 233–250. https://doi.org/10.1075/z.64.15bak

Chesterman, A. 2007. What is a Unique Item? In: Gambier, Y., Shlesinger, M., Stolze, R. (eds) 2007. *Doubts and Directions in Translation Studies*. Amsterdam: Benjamins. 3–13. https://doi.org/10.1075/btl.72.04che

Cappelle, B., Loock, R. 2013. Is there interference of usage constraints? A frequency study of existential there is and its French equivalent il y a in translated vs. non-translated texts. *Target* Vol. 25. No. 2. 252–275. https://doi.org/10.1075/target.25.2.05cap

Eskola, S. 2004. Untypical frequencies in translated language: A corpus-based study on a literary corpus of translated and non-translated Finnish. In: Mauranen, A., Kujamaki, P. (eds) *Translation Universals: Do they exist?* Amsterdam: Benjamins. 83–99. https://doi.org/10.1075/btl.48.08esk

Frankenberg-Garcia, A. 2008. 'Suggesting rather special facts': a corpus-based study of distinctive lexical distributions in translated texts. *Corpora* Vol. 3. No. 2. 195–211. https://doi.org/10.3366/e1749503208000154

Götz A. 2016. A diskurzusjelölők fordításának kérdései. *Fordítástudomány* 18. évf. 1. szám. 5–20.

Götz, A. 2017. Translating doubt: the case of the Hungarian discourse marker vajon. In: Wachowski, W., Kövecses, Z., Borodo, M. (eds) *Micro-scale perspectives on cognition, translation and cross-cultural communication.* Peter Lang. 125–146.

Gyuris B. 2008. A diskurzus-partikulák formális vizsgálata felé. In: Kiefer F. (szerk.) *Strukturális magyar nyelvtan IV*. A lexikon szerkezete. Budapest: Akadémiai Kiadó. 639–682.

Heltai P., Juhász G. 2002. A névmások fordításának kérdései angol–magyar és magyar–angol fordításokban. *Fordítástudomány* 4. évf. 2. szám 25–38.

Heltai P. 2016. Előadás az ELTE Fordítás- és Tolmácstudományi Tanszékén 2016. február 23-án, megjelenés alatt

Károly K. 2003. Korpusznyelvészet és fordításkutatás. A korpuszalapú fordításkutatás néhány elméleti és módszertani kérdéséről és eredményeiről. *Fordítástudomány* 5. évf. 2. szám. 18–27.

Klaudy K. 1999a. *Bevezetés a fordítás gyakorlatába*. Budapest: Scholastica.

Klaudy K. 1999b. *Bevezetés a fordítás elméletébe.* Budapest Scholastica.

Klaudy K. 2001. Az aszimmetria hipotézis. In: Bartha M., Stephanides É. (szerk.) *A nyelv szerepe az információs társadalomban.* A X. MANYE Kongresszus előadásai. Székesfehérvár: KJF. 371–379.

Klaudy K. 2012. Szószintű műveletek szövegszintű hatása a fordításban. In: Bárdosi Vilmos (szerk.) *A szótól a szövegig.* Budapest: Tinta Könyvkiadó. 153–161.

Laviosa, S. 1998. Core Patterns of Lexical Use in a Comparable Corpus of English Narrative Prose. *Meta* Vol. 43. No. 4. 557–570. https://doi.org/10.7202/003425ar

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. https://doi.org/10.1556/acr.1.2000.2.2

Nagy A. L. 2017. A szógyakoriság gépi vizsgálata műszaki szövegeken. *Fordítástudomány* 19. évf. 2. szám. 40–57.

Robin E. 2013. Alulreprezentált „egyedi" nyelvi elemek a fordításban. Sonja Tirkkonen-Condit hipotézisének hatása a fordítástudományra. *Fordítástudomány* 15.évf. 1. szám. 92–102.

Robin E. 2014. Nyelvi babona a fordításokban. In: Benő, Fazakas, Zsemlyei Borbála (szerk.) *Többnyelvűség és kommunikáció Kelet-Közép-Európában: XXIV. Magyar Alkalmazott Nyelvészeti Kongresszus.* Kolozsvár: Erdélyi Múzeum-Egyesület 167–174.

Robin E. 2015. A fordító mint lektor. In: Horváth Ildikó (szerk.) *A modern fordító és tolmács.* Budapest: ELTE Eötvös Kiadó. 35–46.

Robin E., Dankó Sz.; Götz A., Nagy A. L., Pataky É., Szegh H., Török G., Zolczer P. 2016. Fordítástudomány és korpuszkutatás: bemutatkozik a Pannónia Korpusz. *Fordítástudomány* 18. évf. 2. szám. 5–26.

Seidl-Péch O. 2018. Melyek a (szak)fordító és a fordításkutató munkáját segítő legfontosabb nyelvi korpuszok? In: Robin E., Zachar V. (szerk.) *Fordítástudomány ma és holnap*. Budapest: L'Harmattan Kiadó. 175–191.

Tirkkonen-Condit, S. 2002. Translationese – a Myth or an Empirical Fact?: A Study into the Linguistic Identifiability of Translated Language. *Target* Vol. 14. No. 2. 207–20. https://doi.org/10.1075/target.14.2.02tir

Tirkkonen-Condit, S. 2004. Unique items – over- or under-represented in translated language? In: Mauranen, A., Kujamaki, P. (eds) *Translation Universals: Do they exist?* Amsterdam: Benjamins. 177–186. https://doi.org/10.1075/btl.48.14tir

Tirkkonen-Condit, S. 2005. Do Unique Items Make Themselves Scarce in Translated Finnish? In: Károly, K., Fóris, Á. (eds) 2005. *New Trends in Translation Studies.* Budapest: Akadémiai. 177–156.

Internetes hivatkozások

AntConc. Elérhető: www.laurenceanthony.net/software.html

Wordsmith 6.0. Elérhető: www.lexically.net/wordsmith/version6/

Források

Pannónia Korpusz párhuzamos korpusz – Fordított szórakoztató irodalom:

Baldacci, D. 2008. *Isteni igazság.* Budapest: Európa Kiadó. Ford.: Holbok Zoltán 2011.

Cashore, K. 2009. *Fire – Zsarát.* Szeged: Könyvmolyképző. Ford.: Balog Eszter, Szalai Virág 2011.

Cashore, K. 2012. *Keserkék.* Szeged: Könyvmolyképző. Ford.: Szalai Virág 2014.

Clare, C. 2007. *Csontváros.* Szeged: Könyvmolyképző. Ford.: Kamper Gergely 2009.

Diamand, E. 2009. *Kalózok nyomában.* Szeged: Könyvmolyképző. Ford.: Görgey Etelka 2009.

Grey, M. 2015. *A lány éjfélkor.* Szeged: Maxim Könyvkiadó. Ford.: Robin Edina 2016.

Harris, J. 2007. *Rúnajelek.* Budapest: Ulpius-Ház. Ford.: Szűr-Szabó Katalin 2009.

Snyder, V. M. 2013. *Méregtan.* Szeged: Könyvmolyképző. Ford.: Robin Edina 2014.

Stiefvater, M. 2009. *Shiver – Borzongás.* Szeged: Könyvmolyképző. Ford.: Gazdag Tímea 2012.

Westerfeld, S. 2005. *Csúfok* Szeged: Könyvmolyképző. Ford.: Bosnyák Viktória 2007.

Pannónia Korpusz összehasonlítható korpusz – Eredeti szórakoztató irodalom:

Benina 2010. *A Boszorka fénye.* Szeged: Könyvmolyképző.

Bosnyák, V. 2011. *Lovessence.* Szeged: Könyvmolyképző.

Böszörményi, Gy. 2009. *3... 2... 1...* Szeged: Könyvmolyképző.

Böszörményi, Gy. 2004. *Gergő és az álomfogók.* Budapest: Magyar Könyvklub.

Fejős, É.2009. *Eper reggelire.* Budapest: Ulpius-Ház.

Spirit B. 2010. *Árnyékvilág.* Szeged: Könyvmolyképző.

Spirit B. 2009. *A múlt árnyai.* Szeged: Könyvmolyképző.

Szurovecz, K. 2010. *A gyémántfiú.* Szeged: Könyvmolyképző.

Szurovecz, K. 2014. *A fény hagyatéka.* Szeged: Könyvmolyképző.

Zakály, V. 2012. *Szívritmuszavar.* Szeged: Könyvmolyképző.

A szógyakoriság gépi vizsgálata műszaki szövegeken: terminusok és műszaki formulaszerű elemek

Nagy Annamária Lilla

annamarialillanagy@gmail.com

ELTE Nyelvtudományi Doktori Iskola

Fordítástudományi Doktori Program

Kivonat: A jelen tanulmány egy korábbi korpuszalapú kutatás (Nagy 2017) eredményeiből kiindulva kísérel meg egy lépéssel közelebb kerülni a műszaki szövegek terminológiakezelési és szógyakorisági jellemzőinek feltárásához. Elsősorban arra a kérdésre keresi a választ, milyen hatást gyakorolnak a műszaki szövegek szógyakoriságára a szakterület terminusai és a műszaki szövegkörnyezet formulaszerű elemei. A témában eddig viszonylag kevés kutatás született, bár az említett nyelvi elemek fontos szerepet játszanak a szakszövegek összetételének és szövegjellemzőinek alakításában. A jelen tanulmány egy korpuszalapú gépi szógyakorisági vizsgálat eredményeit mutatja be. Az elemzések alapjául egy párhuzamos, angol–magyar fordítási korpusz és egy autentikus szövegeket tartalmazó összehasonlítható korpusz szolgált. A vizsgálati korpusz szakszövegei a Pannónia Korpusz (Robin et al. 2016) műszaki alkorpuszából lettek kiválasztva a jelen kutatás céljára. A kapott eredmények a műszaki szakszövegek szógyakoriságának sajátosságairól adnak részletesebb képet, és fontos lépéseket tesznek a szakszövegek újszerű perspektívából történő vizsgálatainak terén.

Kulcsszavak: formulaszerű elemek, műszaki szövegek, Pannónia Korpusz, szógyakoriság, terminusok

Nagy Annamária Lilla: A szógyakoriság gépi vizsgálata műszaki szövegeken: terminusok és műszaki formulaszerű elemek. In: Robin E., Seidl-Péch O. (szerk.) 2020. *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordításkutatás Magyarországon.* Segédkönyvek a nyelvi közvetítésről I. Budapest: ELTE BTK Fordítástudományi Doktori Program, MANYE Fordítástudományi Szakosztály. DOI: https://doi.org/10.36252/Nyelvikozvsegedkonyv1.8

1. Bevezetés

A korpuszalapú kutatások egyre népszerűbbé válnak a nyelvészeti, főként a szógyakoriságot célzó kutatások körében (Seidl-Péch 2018). A vizsgálódások többségét irodalmi szövegekből álló szövegállományon végzik, elsősorban azok összefüggő, nagy terjedelme miatt, azonban más szövegtípusokon, így például a műszaki szövegeken végzett vizsgálódások is új, érdekes eredményekre vezethetnek, gazdagítva a különböző szövegfajtákról szerzett ismereteinket.

A műszaki szövegek vizsgálatai általában azok fordítási problémáit, problematikáit helyezik középpontba (lásd Sturcz 2010), vagy a más nyelvre történő átültetésükhöz szükséges speciális fordítói kompetencia kérdéskörét vetik fel (lásd Fischer 2012). Ha azonban a már jól ismert kutatási területek ismereteinek bővítése helyett a műszaki szövegek szógyakoriságának jellemvonásait igyekszünk feltárni, érdekes, új eredményeket kapunk. Egy korábbi kutatásban (Nagy 2017) összevetettem a műszaki szövegek szógyakoriságát irodalmi szövegeken végzett, ugyancsak szógyakoriságot feltáró kutatási eredményekkel, és a kapott eredmények alapján megállapítottam, hogy a műszaki szövegek az irodalmi szövegekhez képest kisebb lexikai változatosságot mutatnak, ami többek között valószínűleg a terminusok gyakori használatának, és ezáltali ismétlődésének tudható be.

A kutatás eredményeinek elemzése során felmerült azonban egy további kutatási kérdés, mégpedig annak feltárása, vajon az ismétlődések során szerepet játszott-e a műszaki szövegek szógyakorisági jellemzőiben a szakszövegekre jellemző formulaszerű nyelvhasználat. A jelen tanulmányban elvégzett kutatás erre a kérdésre keresi a választ angol forrásnyelvi szövegek és magyarra fordított változatuk párhuzamos korpuszának vizsgálatával, továbbá az autentikus szövegek szógyakorisági jellemzőit kielemezve az is kiderül, milyen mértékű eltérést mutatnak a fordított és az autentikus magyar műszaki szövegek a formulaszerű nyelvhasználat terén.

2. A műszaki szövegek jellegzetességei

A műszaki szövegek egyre elterjedtebbé válnak, ahogyan világunk és az életünket körülvevő technika folyamatosan fejlődik. A műszaki szövegek esetében nemcsak azok írásakor, hanem a fordítási munka során is szükség van az idegen nyelv klasszikus értelemben vett megfelelő szintű ismeretére, amely az érintett műszaki szakterület szakkifejezéseit, továbbá az adott tudományágban való jártasságot foglalja magában.

A szakszövegek egy további jellemvonása, hogy a szövegek úgynevezett „puha" és „kemény" részekre oszlanak (Klaudy 2004: 17). „Puha" vagy más szóval „rugalmas" résznek azokat nevezzük, amelyek jelentésének átültetésekor a fordítónak több különböző lehetőség áll rendelkezésre, és ezekből kiválaszthatja a véleménye szerinti legmegfelelőbbet. A „puha" részek fordítása ellentétben a kemény részekével tehát a fordító egyéni döntése szerint alakul, tehát különböző fordítók munkája egymástól eltérhet, mégis minden változat elfogadható és helyes lehet. „Kemény" részeknek nevezi a szakirodalom a szöveg azon részleteit, ahol a fordítónak nincs választási lehetősége arra, hogy különböző megoldások közül választhasson, hanem csak egyetlen behelyettesíthető lexikai egység fogadható el helyes fordításként. Az úgynevezett „kemény" elemek közé sorolandók a szakszövegek terminusai is.

2.1. Fordítói műveletek

Mivel a nyelvek eltérő jelleget mutatnak, a fordítás során különböző műveleteket kell elvégezni annak érdekében, hogy a célnyelvi szöveg ugyanazt a szerepet tölthesse be az adott célnyelvi rendszerben, mint a forrásnyelvi szöveg a sajátjában. A fordítástudomány egyik úttörőnek számító tanulmányaként tartjuk számon Mona Baker 1993-as írását, amelyben a szerző az univerzálékat a fordítási viselkedés egyik alapelvének tekinti, és a fordítás eredményeként keletkező szövegek tipikus sajátosságaként jelöli meg.

Az egyik elsőként felfedezett és általánosan elfogadott fordítási univerzálé az explicitáció és az implicitáció műveleti párosa. Az explicitáció fogalmát Vinay és Darbelnet (1958) határozta meg először. A későbbiekben számos fordításkutató végzett különböző

elemzéseket és vizsgálatokat az explicitációs jelenség jellemzőinek és okainak feltárására, többek között Blum-Kulka (1986), Toury (1995) Olohan és Baker (2000), Klaudy és Károly (2005), Heltai és Juhász (2002), valamint Robin (2018). Explicitációnak azt a jelenséget nevezzük, amikor a fordító egyértelműbben fogalmaz meg valamit a célnyelvi szövegben, ami a forrásnyelvi szövegben csak a kontextusból kikövetkeztethető információ volt, de sem lexikai, sem grammatikai formában nem szerepel konkrétan a szövegben. Klaudy (1999) tipológiájában rávilágít arra, hogy az explicitáció nem csupán tudatos fordítói döntés eredménye lehet, hanem a nyelvek különbözőségéből fakadó elkerülhetetlen művelet is. Ennek megfelelően négy fő kategóriát azonosít: kötelező, fakultatív, pragmatikai és fordításspecifikus fordítói műveletek. Klaudy tipológiájára támaszkodva Robin (2018) az alábbi felosztást javasolja: szabálykövető, normakövető, valamint stratégiakövető műveletek.

Az explicitáció és az implicitáció témája egyúttal mindig felveti az ismétlés, illetve az ismétléskerülés témáját is, hiszen ugyancsak az univerzálék között emlegeti a szakirodalom (Baker 1993). Toury (1991) arra a következtetésre jutott, hogy a fordítók jelentősen csökkentik a lexikai ismétléseket a munkájuk során. Klaudy és Károly (2000) hivatásos fordítókon és fordító hallgatókon végzett vizsgálata azt az eredményt hozta, hogy a gyakorlott fordítók nem változtatnak jelentősen az ismétléskapcsolatokon, a kezdő fordítók azonban igen. Robin (2018) szintén arra a következtetésre jutott, hogy az explicitáció és az implicitáció használatával a fordítók kerülik az ismétléseket. A szakirodalom tehát gyakran rámutat a gyakorlott fordítók azon stratégiájára, hogy bizonyos esetekben csökkentik a lexikai ismétléseket, akad azonban olyan eset is, amikor a forrásnyelvi és a célnyelvi változatok ismétléseinek száma között nem jelentkezik kifejezetten nagy eltérés. Az ilyen megfigyelésekből arra következtethetünk, hogy bizonyos befolyásoló tényezők alakítják az ismétlést vagy adott esetben az ismétlések kerülését. Az egyik ilyen tényező a szövegek fajtája. Ennek klasszikus példája az irodalmi szövegek fordítása, ahol könnyen véghez vihető a lexikai ismétlések kerülése anélkül, hogy az negatív hatást gyakorolna a szöveg koherenciájára és felépítésére, egyúttal pedig a végeredmény is választékos célnyelvi szöveg lesz. Szakszövegek esetén azonban más helyzettel szembesül a fordító. A műszaki szövegek esetén az ismétlések kerülése negatív hatást gyakorolhat a szöveg

kohéziójára és koherenciájára, így a fordított szöveg minőségét is negatív irányban befolyásolhatja, hiszen az adott szövegben a terminusok szinonimákkal való behelyettesítése gyakran értelmezésbeli bizonytalanságot okoz, ezért akár fordítási hibának is minősülhet.

2.2. A terminusok és fordításuk

Az explicitáció és az implicitáció ugyan univerzális fordítási műveletnek tekinthető (Klaudy 1999), gyakoriságuk mégis eltérő az egyes nyelvpároknál, illetve azon belül a különböző szövegtípusoknál (Kamenická 2008; Szegh 2017). A fordítási műveletpáros műszaki szövegek fordításánál is jelentős szerepet játszik. A műszaki szövegek egyik, már-már alapérvényűnek tekintett jellemvonása a gyakori terminushasználat. Korántsem egyszerű azonban annak pontos meghatározása, mit tekinthetünk terminusnak. 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. Előfordul azonban az is, hogy egyes szavak vagy kifejezések terminusvolta sokkal nehezebben ismerhető fel, mert időközben a köznyelv részévé váltak. Ilyen esetekben a fordítónak jóval nehezebb dolga van, mint az azonnal terminusként azonosíthatók esetében, hiszen a terminusok fordításának egyik fő feltétele az, hogy a fordító lehetőleg ne helyettesítse be más kifejezésekkel őket, és ne a körülírást válassza fordítói megoldásként.

Heltai (1988) továbbá rámutat, hogy a magyar szaknyelvben a szinonimák használata csak korlátozott keretek között történik, sőt jellemzőbb, hogy a laikus közönség számára készült szövegek esetén ugyanúgy megmarad a szakterminusok használata, mint a szakértő közönség számára készült írásoknál és beszédeknél. Ezzel szemben az angol szaknyelv gyakran alkalmaz szinonimákat is terminusok esetében.

Hasonló gondolatmenetből kiindulva, Fischer (2010) doktori értekezésében rávilágít egy további, fordítókat érintő problémafelvetésre is, miszerint a fordítók igyekeznek ragaszkodni ahhoz a szabályszerűséghez, hogy terminusok esetében nem alkalmazhatnak

sem körülírást, sem szinonimákat, sem más egyéb, a konkrét célnyelvi terminus behelyettesítésétől eltérő fordítói műveletet. Mint ahogy azt a későbbiekben bemutatom, a kutatás alapjául szolgáló szövegek fordítói valóban bizonytalanok voltak olyan kifejezések terminusnak való jelölése esetén, amelyeknek létezik ismert rokonértelmű változata is.

Érdemes azonban szem előtt tartani, hogy az a tény, miszerint bizonyos terminusoknak létezik szinonimájuk, nem jelenti feltétlenül azt, hogy bármilyen szövegkörnyezetben szabadon felhasználhatók, sőt mi több, a rossz minőségű, széttöredezett célnyelvi szövegek egyik fő okaként tekinthetünk erre a jelenségre. A magyar anyanyelvű fordítóhallgatóknál gyakori hiba, hogy terminusok esetében is igyekeznek szinonimát vagy körülírást alkalmazni a szóismétlés elkerülése végett, mivel iskolai tanulmányaik során a diákokat kifejezetten erre ösztönzik. A szakszövegekben azonban a terminushasználatnak kohézióteremtő ereje van, így a felesleges vagy esetleg az adott környezetben helytelen szinonimák használata negatív hatást gyakorolhat a szöveg kohéziójára (Fischer 2012).

A jelen vizsgálódás szempontjából ezek az elgondolások kiemelkedő jelentőségűek, mivel az explicit és implicit megfogalmazás, a szinonimák alkalmazása, az ismétlések kerülése vagy megtartása egyaránt hatással van az autentikus és a fordított szövegek szógyakoriságára.

3. A formulaszerű nyelvhasználat

A formulaszerű elemek kategóriájába a hagyományos nyelvészet olyan több elemből álló kifejezéseket sorol, amelyeket a nyelvhasználó egy egységként tárol és hív elő a mentális lexikonból (Weinert 1995). A pszicholingvisztikai kutatások szerint a nyelvi egységeket redundánsan tároljuk a mentális lexikonban, így az összetett szavak és az összetételt alkotó szavak egyaránt megtalálhatók a beszélő saját mentális lexikonjában (Nattinger és DeCarrico 1992). Wray (2002: 9) szerint a formulaszerű szekvenciák előre gyártott elemek – legalábbis annak tűnnek –, amelyeket a memória egyben tárol és így is hív elő, tehát nem a nyelvtan segítségével generált kifejezések. Más elméletek, így például Howarth (1996) szerint bizonyos szókapcsolatok egységként tárolódnak, míg mások nem.

A formulaszerű nyelvhasználatra először Pawley és Syder 1983-as cikke hívta fel a figyelmet. Az anyanyelvi beszélők nyelvhasználatának sajátosságait feltáró tanulmány két úgynevezett rejtélyt állapított meg: Az egyik az „anyanyelvi folytonosság" (*native fluency*), más szóval az a jelenség, hogy anyanyelvünk használata során alapvetően nem okoz problémát a folytonos beszéd. A másik pedig az „anyanyelvi szelekció" (*native-like selection*), amelyen a szerzők azt értik, hogy az anyanyelvi beszélők nyelvhasználata valóban anyanyelvinek (idiomatikusnak) hangzik, és a többféle nyelvtanilag helyes megoldásból az anyanyelvi beszélő ösztönösen az anyanyelvinek ható kifejezést választja, a többi pedig helyességük ellenére is idegennek hat számára. Az *I want to marry you* mondat helyett például nem használják az *I wish to be wedded you* változatot, holott nyelvtanilag helyes volna. A két rejtély kulcsát a szerzőpáros a formulaszerű nyelvhasználatban véli felfedezni.

A formulaszerű elemeknek egyre fontosabb szerepet tulajdonít a nyelvelsajátítással foglalkozó szakirodalom is. Nattinger és DeCarrio (1992) könyve szerint az emberi nyelvhasználatban az egyes lexikai frázisok „összeszerelt egységként" állnak rendelkezésünkre, és annak ellenére, hogy az adott nyelv szabályrendszere szerint elemezhetők és generálhatók, mégis általában nem a szabályrendszer segítségével hozzuk létre őket, hanem emlékezetből. A formulaszerű nyelvi elemek osztályozása különbözőképpen lehetséges, attól függően, mely irányvonalat követjük. Összességében ide sorolják a frazémákat és az idiómákat, az állandósult szókapcsolatokat, a különböző lexikai frázisokat, legyen szó pragmatikai formuláról, kliséről, sztereotípiáról, helyzetmondatról, közmondásról vagy társalgási manőverről, illetve közéjük tartoznak a különböző memorizált szövegek is.

Wray (2002) a formulaszerű elemeket formulaszerű szekvenciáknak nevezi, amelyek egyik kritériuma a gyakoriság, másik pedig azok hosszúsága és összetettsége. Weinert (1995) ezzel szemben a döntő kritériumot abban látja, hogy a formulaszerű kifejezéseket egységként tároljuk a mentális lexikonban, és az adott beszédszituációban ilyen formában hívjuk elő, nem a nyelvtani szabályrendszer segítségével. Habár az egyes kutatók véleménye bizonyos témákban eltérő, a szakirodalom áttekintése azt mutatja, hogy a formulaszerű elemek egységként történő tárolása és előhívása tekintetében egy véleményen vannak.

Ezt a gondolatmenetet követve feltételezhetjük, hogy a formulaszerű elemek tárolása és előhívása nemcsak az anyanyelvi nyelvhasználat során, hanem a fordítás során is egységként történhet. Az adott szövegrészlet vagy kifejezés értelmezését követően az ekvivalenst a fordító teljes egységként raktározza el a mentális lexikonba, majd adott esetben komplett egészként hívja elő.

4. A kutatás bemutatása

A jelen kutatás célja annak feltárása, hogy a műszaki szövegekben előforduló terminus értékű szóismétlések vizsgálatában milyen és mekkora szerepet játszanak a formulaszerű kifejezések. Feltételezésem szerint a forrásnyelvi, a célnyelvi és az autentikus szókapcsolatokat tartalmazó gyakorisági listákban egyaránt jelentős számban fordulnak elő formulaszerű nyelvi elemek. Úgy vélem, továbbá, hogy a legtöbb, a vizsgálat szempontjából értékes találatot a kéttagú elemekben, a legkevesebbet pedig a négytagú, általam vizsgált elemekben találom. Ennek okát abban látom, hogy a kutatásban nem értékelhető elemtagok (névmások, névelők, kötőszavak stb.) megjelenésének lehetősége jóval nagyobb a négytagú elemekben.

4.1. A kutatási korpusz és módszerek

A korpuszalapú nyelvészeti kutatások elsősorban Mona Baker (1993) iskolateremtő cikkének hatására terjedtek el a fordítástudományban. Baker ebben az írásában rámutat arra, hogy a korpuszalapú elemzések segítségével akár hatalmas szövegmennyiség vizsgálatával mutathatók ki különböző nyelvi jelenségek fordított és autentikus szövegeken egyaránt. Nemzetközi szinten számtalan digitálisan tárolt korpusz létezik, így például a CTF (*Corpus of Translated Finnish*), a TEC (*The Translational English Corpus*) és a GEPCOLT (*German-English Parallel Corpus of Literary Texts*), azonban Magyarországon korábban nem létezett nagyobb méretű párhuzamos szövegeket tartalmazó fordítási szövegtár. A jelen tanulmány alapjául szolgáló fordítási korpuszt az Eötvös Loránd Tudományegyetem Nyelvtudományi Doktori Iskoláján belüli Fordítástudományi Doktori Program egyik pro-

jektjeként 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 magyar fordítását használtam fel.

Az elemzéshez 23 autentikus magyar nyelvű értekezést és 7-7 angol–magyar párhuzamos forrásnyelvi és célnyelvi szöveget vizsgáltam. A Pannónia Korpuszból vett szövegek teljes terjedelme megközelítőleg 500 000 szövegszó mindhárom esetben, így a vizsgálódás több mint 1 500 000 szövegszóra terjed ki. Habár az autentikus és a párhuzamos szövegek darabszáma között jelentős különbség van, az egyes alkorpuszok terjedelme így vált megközelítőleg egyenlővé. Az autentikus szövegeket abból a megfontolásból nem vontam össze, mert bár mind műszaki szakszöveg, az egyes szakterületek sajátos szaknyelvi lexikája – amely a kutatás középpontjában áll – minden bizonnyal elveszett vagy háttérbe szorult volna a 23 szöveg egy korpuszba való összevonásával. 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, ami az angol nyelv analitikus jellegével magyarázható. A párhuzamos szövegek hossza nem arányosan oszlik el, mivel az informatikai szövegek könyvformátumban jelentek meg, míg a használati utasítások értelemszerűen rövidebbek, azonban a legrövidebb is eléri a 20 000 szövegszót. Ezzel szemben az autentikus magyar szövegek hossza nagyjából megegyező, minden értekezés megközelítőleg 20 000 és 30 000 szövegszó közötti. Az autentikus szövegkorpuszt 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, ezért az elemzésül szolgáló korpuszban is találhatók ilyen típusú szövegek.

A korpusz szövegállományán az *Antconc* nevezetű számítógépes szoftverrel végeztem kvantitatív vizsgálatokat. A gépi elemzéssel kigyűjtöttem a korpusz egyes elemeinek egyenként 100 leggyakoribb két-, három- és négytagú szócsoportjait, majd a lista elemeiből kézi elemzéssel kiválogattam a többtagú terminusnak vagy műszaki környezet-

ben gyakori szófordulatnak számító elemeket és az adott műszaki környezetben gyakori szófordulatokat. A kigyűjtött elemeket ezután szétbontottam két csoportra: a többtagú terminusok és a műszaki környezetben gyakori szófordulatok csoportjára, a kapott adatokat pedig ugyancsak feltüntettem a táblázatokban. A vizsgálat megerősítése és az esetleges elemzési hibák elkerülése céljából felkértem kollégámat, dr. Ábrányi Henriettát az eredmények kontrollelemzésére.

Habár jogos felvetés volna, miért nem került be minden – tehát a terjedelmes listán hátrébb szereplő – többtagú terminus vagy gyakori műszaki szófordulat, azonban úgy vélem, hogy a leggyakoribb jelzővel ellátott listáknál feltétlenül fontos meghúzni egy határvonalat (Laviosa 1998). Erre főként azért van szükség, hogy ne vesszünk el a részletekben, és ne vonjunk le komolyabb következtetéseket úgy, hogy számos elem csupán egyszeri előfordulást tud felmutatni. Továbbá, a jelen kutatásnak alapot adó korábbi kutatásban szintén csak az úgynevezett *list head* (Laviosa 1998) került a vizsgálandó elemek közé, így ettől a mintától ezúttal sem szerettem volna eltérni. A cikk végén található *1. függelék* mutatja be a két tagból álló szókapcsolatok gyakorisági listáját, zöld szín jelzi a kéttagú terminusértékű kifejezéseket.

5. Eredmények bemutatása

Az egymás mellett leggyakrabban előforduló kifejezések listáit megvizsgálva a következő táblázatokban bemutatott értékek mutatkoznak a többtagú terminusok és műszaki környezetben gyakori szófordulatok előfordulásával kapcsolatban. A táblázatokban külön mutatom a két-, három- és négytagú elemekkel kapcsolatos eredményeket, mivel értelemszerűen ezek között jelentős eltérés tapasztalható már egyetlen szöveg vizsgálata esetén is.

Az eredmények bemutatását megelőzően előrebocsátanám, hogy a felsorolt listákból nem számoltam találatnak azokat a kéttagú elemeket, ahol az egyik elem ugyan terminus, a másik azonban csupán névelő, kötőszó, névelő, névutó vagy más, az adott kifejezésnek vagy terminusnak nem szerves részét képező elem (például *an application*). A három- és négytagú elemeknél már megengedőbbnek kellett lenni, mivel a nyelvek

sajátosságából fakadóan a szószerkezetek részeit képezik a felsorolt elemek, és nélkülük a szerkezet nyelvileg létre sem tudna jönni – együtt alkotnak formulaszerű kifejezéseket (például *date and time*). Ezekben az esetekben sem számoltam azonban találatnak azokat a szókapcsolatokat, ahol a háromtagú szerkezet valójában egy névelővel ellátott kéttagú terminus, illetve négytagú szerkezet esetén névelővel ellátott háromtagú terminus (például *the hub transport server*).

5.1 Eredeti angol szövegek

Az *1. táblázat* a párhuzamos korpusz szövegeinek angol eredeti változatánál mutatja a vizsgált kifejezések és szószerkezetek válogatás utáni számadatait. Az egyes táblázatoszlopok az angol forrásnyelvi szövegekben leggyakrabban előforduló terminusok és műszaki szövegkörnyezetre jellemző formulaszerű nyelvi elemek számát mutatja két-, három-, és négytagú szerkezetekben vizsgálva. A szövegek mellett zárójelben a szavak száma látható, míg a tagokat jelző oszlopok számai melletti zárójelekben látható adatok a terminusok és a műszaki környezetben gyakori szófordulatok eloszlását jelzik, ahol is mindig az első szám vonatkozik a terminusokra.

1. táblázat

Forrásnyelvi angol szövegek elemzési adatai

	2 tag	**3 tag**	**4 tag**
1. szöveg (63 297)	12 (11-1)	14 (13-1)	2 (2-0)
2. szöveg (191 205)	26 (23-3)	20 (20-0)	1 (1-0)
3. szöveg (80 181)	25 (23-2)	22 (19-3)	8 (8-0)
4. szöveg (98 252)	7 (7-0)	14 (10-4)	1 (1-0)
5. szöveg (72 201)	21 (21-0)	15 (14-1)	9 (8-1)
6. szöveg (21 231)	16 (15-1)	11 (10-1)	3 (2-1)
7. szöveg (55 972)	29 (29-0)	12 (11-1)	2 (2-0)
Összesen	**136 (129-7)**	**108 (97-11)**	**26 (24-2)**

A táblázat alsó sora az eredeti angol szövegeket tartalmazó alkorpusz összesített adatait mutatja. Érdekes megfigyelni, hogy bár a két tagból álló szerkezetek összesített darabszáma a legmagasabb, és az egyes szövegeket nézve is többségében ebben az oszlopban szerepelnek a nagy számok, az 1-es és a 4-es számú szöveg esetében a legmagasabb érték érdekes módon a háromtagú szerkezeteknél jelentkezik. A két említett szöveget nézve különösen a negyedik szöveg eredménye kiugró, ott ugyanis a háromtagú szerkezetekből kétszer annyi a találat, mint a 2 tagból állókból. Ennek magyarázata feltehetően abban rejlik, hogy az említett szövegben a találatok a műszaki szövegekben gyakori formulaszerű kifejezések esetében magasabb értéket mutatnak. Ezt támasztják alá a találatok eloszlását jelző adatok, hiszen míg a 4-es számú szöveg kéttagú elemeinél minden kiválasztható elem a többtagú terminus kategóriába sorolható, addig a háromtagú elemeknél a műszaki környezetben gyakran előforduló kifejezések tekintetében ez a legmagasabb érték (4) a forrásnyelvi szövegek táblázatában. A kéttagú szerkezeteknél a legmagasabb találat 29, a legalacsonyabb pedig 7, a három oszlopból a legnagyobb különbséget adva. A háromtagú szerkezetek esetében a legmagasabb érték 22, a legalacsonyabb pedig 11, így a kéttagú kifejezések esetében nagyobb változatosságnak lehetünk tanúi az alkorpuszban a szógyakoriságot illetően.

(1) Példák 2 tagból álló listaelemekre: *running windows, then click, white balance, memory card, mail address, public folder*

(2) Példák 3 tagból álló listaelemekre: *remove the battery, figure below shows, red eye reduction, hub transport server, built in flash, shutter release button*

A négy tagból álló szókapcsolatok vizsgálatánál várható volt, hogy kevesebb számban fordul elő a vizsgálat szempontjából jelentőséggel bíró szerkezet. Ennek fő oka nem más, mint az az egyszerű tény, hogy minél több tagból álló szerkezeteket szeretnénk vizsgálni, annál kisebb a matematikai esély arra, hogy a szavak adott sorrendben ismétlődnek. Egy másodlagos, de a jelen vizsgálat szempontjából mégis fontos megjegyzés ezen a téren az, hogy a négy vagy több elemből álló szókapcsolatra alapvetően sem sok példa

hozható fel az angol nyelvben: a négytagú szerkezeteket bemutató oszlopban a legalacsonyabb találat 1, a legmagasabb pedig 9 – ezek közül is a legtöbb terminus, és csak elvétve találkozunk terminusnak nem minősülő formulaszerű kifejezésekkel.

(3) Példák 4 tagból álló listaelemekre: *displayed in the viewfinder, server based domain controllers, turn the camera on, preset manual white balance, public folder management controller*

5.2 Célnyelvi magyar szövegek

A *2. táblázatban* a célnyelvi magyar szövegek vizsgálata során kapott eredmények láthatók. Az egyes oszlopok – ahogyan a forrásnyelvi szövegek eredményeit bemutató táblázatban is – a két-, három- és négytagú szerkezetek előfordulását mutatja be a vizsgált szövegekben, zárójelben feltüntetve a terminusok és a formulaszerű gyakori elemek számainak eloszlását.

2. táblázat

Célnyelvi magyar szövegek elemzési adatai

	2 tag	3 tag	4 tag
1. szöveg (58 099)	13 (12-1)	9 (9-0)	6 (6-0)
2. szöveg (161 922)	14 (12-2)	8 (7-1)	4 (4-1)
3. szöveg (80 139)	18 (16-2)	19 (16-3)	10 (7-3)
4. szöveg (68 094)	11 (10-1)	5 (3-2)	0 (0-0)
5. szöveg (67 899)	11 (11-0)	7 (6-1)	7 (5-2)
6. szöveg (19 348)	8 (6-2)	8 (8-0)	8 (6-2)
7. szöveg (50 562)	10 (9-1)	7 (6-1)	14 (12-2)
Átlag	85 (76-9)	63 (55-8)	49 (39-10)

A kéttagú szerkezetek számadatait bemutató oszlop mondható a legkiegyensúlyozottabbnak, hiszen a legkisebb érték 8, míg a legmagasabb 18. A szövegekben előforduló összes

szerkezet száma pedig 85, több a három- és négytagú szerkezetek számánál, azonban jóval alacsonyabb, mint a forrásnyelvi angol szövegek esetében. Az utóbbi jelenség hátterében feltehetően az angol és a magyar nyelv rendszerbeli különbségei állhatnak, elsősorban a helyesírási szabályok terén, ugyanakkor ennek a megállapításnak az igazolásához kvalitatív vizsgálatokra van szükség.

(4) Példák 2 tagból álló listaelemekre: *operációs rendszer, biztonsági mentés, ISO* érzékenység, jobb gombbal, majd kattintsunk, élő nézetben

A háromtagú szerkezetek esetében a legalacsonyabb érték 5 a legmagasabb pedig 19, nagyobb változatosságról tanúskodva, utóbbi számadat pedig a táblázat legmagasabb értéke. Az érintett táblázat 3. számú szövege bizonyult a leggazdagabbnak a terminusok és műszaki szövegekben gyakori formulaszerű nyelvi elemek tekintetében, hiszen a két-, három- és négytagú szerkezetek száma egyaránt legalább 10 – tükrözve a forrásnyelvi szöveg jellemzőit.

(5) Példák 3 tagból álló listaelemekre: *h*ázon *belüli szoftverek, teljes képes visszajátszás, automatikus ISO* érzékenység, *súgó és támogatás*

A négytagú szerkezeteket bemutató táblázatoszlop a fordított szövegek esetében jóval magasabb számot, az angol szövegek közel kétszeresét adja, arról tanúskodva, hogy a magyar nyelvben magasabb a négytagú terminusok és formulaszerű kifejezések száma. Azonban még így is ez a teljes célnyelvi szövegeket bemutató táblázat legalacsonyabb értékeket felvonultató oszlopa, és fontos megemlíteni, hogy az egyes szövegek eredményei között szerepel 0 is.

(6) Példák 4 tagból álló listaelemekre: *kattintsunk az ok gombra, kattintsunk a next gombra, súgó és támogatás központ*

A többtagú nyelvi elemek csoportbontásánál azt figyelhetjük meg, hogy a műszaki szövegekben gyakori szófordulatok tekintetében nagyjából hasonló értékeket mutat a célnyelvi szövegek táblázata, mint a forrásnyelvi szövegeké – ez alól a négytagú kifejezések képeznek kivételt, amelyek jelentősen nagyobb számban szerepelnek a célnyelvi magyar szövegekben. Kiemelendő azonban, hogy kevesebbszer fordulnak elő – még a táblázat alapvetően alacsony értékei ellenére is –, hogy a gyakori szófordulatok darabszáma nulla, míg az angol szövegek esetében ennek az ellenkezőjét láthattuk. A magyar fordításokban tehát a terminusokhoz képest nagyobb százalékban fordulnak elő formulaszerű kifejezések a gyakori szókapcsolatok között, mint angol szövegekben.

5.2 Autentikus magyar szövegek

Az alábbi *3. táblázatban* az autentikus magyar szövegek vizsgálati eredményei találhatók. Az egyes oszlopok, akárcsak az előző két táblázat esetében, a két-, három- és négytagú szerkezetek előfordulási gyakoriságát mutatja.

A legtöbb darabszámot felmutató oszlop az autentikus magyar szövegek esetében is a kéttagú szerkezetek csoportja, azonban az egyes szövegekben vizsgált szerkezetek száma korántsem nevezhető kiegyensúlyozottnak. Akad olyan szöveg, amelyben csupán egyetlen, a vizsgálat szempontjából számba vehető szerkezet fordult elő, találhatunk azonban olyan szöveget is az autentikus alkorpuszban, amelyben a talált szerkezetek száma 16.

Elmondható tehát, hogy az eredetileg is magyar nyelven íródott műszaki szövegek nagyobb változatosságról tanúskodnak a fordított magyar szövegekhez képest a szógyakoriságot illetően.

3. táblázat

Autentikus magyar szövegek elemzési adatai

	2 tag	3 tag	4 tag
1. szöveg (21 996)	8 (7-1)	9 (8-1)	1 (1-0)
2. szöveg (31 410)	1 (1-0)	3 (3-0)	0 (0-0)
3. szöveg (22 203)	1 (1-0)	1 (1-0)	6 (6-0)
4. szöveg (20 632)	6 (6-0)	4 (4-0)	14 (14-0)
5. szöveg (15 270)	6 (5-1)	0 (0-0)	0 (0-0)
6. szöveg (15 801)	6 (6-0)	5 (5-0)	0 (0-0)
7. szöveg (24 290)	5 (4-1)	9 (8-1)	7 (5-2)
8. szöveg (15 546)	4 (4-0)	3 (3-0)	5 (5-0)
9. szöveg (29 153)	1 (1-0)	3 (3-0)	2 (2-0)
10. szöveg (10 995)	13 (13-0)	10 (8-2)	11 (10-1)
11. szöveg (24 077)	7 (6-1)	4 (4-0)	4 (4-0)
12. szöveg (17 122)	8 (5-3)	10 (9-1)	4 (4-0)
13. szöveg (19 115)	6 (4-2)	5 (5-0)	1 (1-0)
14. szöveg (23 913)	11 (11-0)	9 (9-0)	5 (5-0)
15. szöveg (24 408)	3 (2-1)	9 (9-0)	5 (5-0)
16. szöveg (30 293)	9 (7-2)	1 (1-1)	0 (0-0)
17. szöveg (16 379)	6 (4-2)	1 (1-0)	5 (5-0)
18. szöveg (29 338)	12 (12-0)	9 (9-0)	6 (5-1)
19. szöveg (18 149)	5 (5-0)	11 (7-4)	5 (4-1)
20. szöveg (16 386)	16 (15-1)	15 (13-2)	6 (6-0)
21. szöveg (18 487)	10 (10-0)	8 (8-0)	4 (4-0)
22. szöveg (31 821)	5 (5-0)	10 (9-1)	10 (10-0)
23. szöveg (12 059)	14 (13-1)	8 (8-0)	4 (4-0)
Átlag	**163 (147-16)**	**147 (134-13)**	**105 (100-5)**

Fontos kiemelni továbbá, hogy az autentikus magyar szövegekben jelentősen több kéttagú kifejezés fordul elő (163), mint a fordított magyar szövegekben (85) – ez pedig nem magyarázható a magyar és az angol nyelv közötti rendszerbeli különbségekkel, inkább a magyar nyelvi norma elvárásaival magyarázható.

(7) Példák 2 tagból álló listaelemekre: *egyes szempontok, fenntartható közlekedés,* gördülési ellenállás, felületi ellenállás

A háromtagú szerkezetek vizsgálatánál hasonló eredmény született: jóval magasabb, több mint kétszeres a háromtagú kifejezések száma a fordított magyar szövegekéhez képest, és a szövegek adatai a három tagot számláló szerkezetek szempontjából 0 és 15 között mozognak.

(8) Példák 3 tagból álló listaelemekre: *erők* és nyomatékok, magnézium bázisú hibrid, fajlagos forgácsolási erők

Ahogyan a forrás- és célnyelvi szövegeket bemutató táblázatok esetében látható volt, úgy az autentikus szövegeknél is a négytagú szerkezeteket bemutató oszlop tartalmazza a legtöbb alacsony értéket, és az összesített darabszám is a legalacsonyabb (105). Ugyanakkor a négytagú szókapcsolatokat illetően is elmondható, hogy jóval nagyobb számban fordulnak elő, mint a fordított magyar szövegekben. Külön kiemelendő továbbá, hogy az autentikus szövegek esetében fordult elő a leggyakrabban a 0 és 1 találatos érték, azonban itt fűzném hozzá, hogy a táblázatok elemzésekor felfigyeltem a jelenségre, miszerint főként az autentikus szövegekre jellemző azon kéttagú terminusértékű kifejezések előfordulása, amelyeket azért nem lehetett a vizsgálat szempontjából figyelembe venni, mert a szerkezet első tagja csupán névelő volt.

(9) Példák 4 tagból álló listaelemekre: *aktuális járatterhelési meteorológiai karakterisztika, állandó amplitúdójú fárasztókísérleti élettartam, hibrid és kompozit anyagok*

A terminusok és a formulaszerű nyelvi elemek eloszlásának tekintetében az autentikus szövegek esetében figyelhető meg legerőteljesebben a terminusok irányába történő eltolódás, különösen a négytagú elemek esetében. Az autentikus szövegekkel egybevetve ugyanis jól látszik, hogy a fordított magyar szövegek jobban támaszkodnak a négytagú formulaszerű kifejezésekre a megfogalmazásban. A műszaki szövegekben gyakori szófordulatok száma rendkívül alacsony értékeket tud felmutatni (a legmagasabb szám 4), és a 0 érték ebben a táblázatban jelentkezik a leggyakrabban. Ezzel kapcsolatban fontos rámutatni arra is, hogy a táblázatban matematikailag nagy esélye lehetett volna a szakmai szófordulatok előretörésének, hiszen az egyes szövegeket eltérő témájukra való tekintettel, egyesével elemeztem.

A táblázatokat összevetve kirajzolódik, hogy a magyarra fordított célnyelvi szövegek szógyakorisági listáin szereplő vizsgált elemek habár nagy számban jelennek meg, alacsonyabb értékeket mutatnak az autentikus szövegekhez képest. Alapvetően kijelenthető továbbá, hogy kiegyensúlyozott számban vannak jelen, beszéljünk akár a két, akár a három vagy a négy tagból álló elemekről. A forrásnyelvi angol szövegek értékei a fordított magyar változataikkal összehasonlítva egymáshoz viszonyítva közel hasonló értékeket adnak, ami arra enged következtetni, hogy a fordítók a többtagú terminusokat, és jellegzetes műszaki fordulatokat igyekeztek megtartani a fordítási munka során.

Az autentikus szövegekben a fentiekkel ellentétben jóval magasabb a különbség az elemadatok tekintetében, tehát itt a legnagyobb az eltérés a szövegek között is. A kiugróan magas értékek mellett a kiugróan alacsony értékek is jelen vannak: a táblázatra tekintve nem található semmilyen szabályszerűség vagy kirajzolódó minta. Az autentikus magyar szövegek egyenetlen képe talán azzal magyarázható, hogy a szabad anyanyelvi szövegalkotásban egyes szövegek írói bátrabban nyúltak a szinonimákhoz, egyesek azonban igyekeztek jól bevett, vagy az adott szakterületen kötelezőnek tekinthető kifejezéseket és

szófordulatokat használni. Ellenben a forrásnyelvi szöveg minden esetben befolyásolja a fordítót, aki a szöveghűség érdekében kevesebb mozgásteret hagy magának a szakkifejezéseket és szakmai szófordulatokat tekintve.

A jelen tanulmány eredményei bizonyos tekintetben hasonló irányt mutatnak az előző kutatással (Nagy 2017). A terminus értékű szavak százalékos aránya az autentikus szövegek vizsgálatában adták a legalacsonyabb értéket, ami szintén azt jelezte, hogy az autentikus szövegek alkotói kevésbé követnek kockázatkerülő magatartást a fordítókhoz képest. Ugyancsak visszatérő jelenség, hogy az autentikus szövegek adják a legszélsőségesebb adatokat mindkét kutatásban, továbbá az, hogy a fordított szövegek igyekeznek többé-kevésbé a forrásnyelvi szövegekkel hasonló képet adni.

6. Összefoglalás

A tanulmány a műszaki szövegek jellemzőit igyekezett feltárni, illetve választ találni arra a kérdésre, milyen hatást gyakorolnak a terminusok az ilyen műfajú szövegek szóhasználatára. A vizsgálatokat követően arra a következtetése jutottam, hogy az eredmények nem konkluzívak, azonban a vizsgált szövegek többsége egy irányba mutat. A tanulmányban különböző műszaki szakszövegek többtagú terminusainak és szószerkezeteinek gyakoriságát tártam fel az *Antconc* nevű számítógépes szoftver gyakorisági listáinak elemzésével. Az egyes szövegcsoportok adatait bemutató táblázatra tekintve megfigyelhető, hogy a kéttagú szerkezetek gyakorisága mutatja a legmagasabb értéket, még annak ellenére is, hogy erre a kategóriára vonatkoztak a legszigorúbb kitételek. Elképzelhető azonban, hogy a kéttagú kifejezések között olyan elemek is előfordulnak, amelyek négytagú kifejezések részét képezik, így az eredmények pontosítása érdekében további vizsgálatok szükségesek.

Az eredményekből arra következtethetünk, hogy a forrásnyelvi szöveg hatására a fordított műszaki szövegek jellegzetes kifejezései nemcsak szavak szintjén végzett, hanem többtagú szószerkezetekre vonatkozó vizsgálat esetén is inkább a forrásnyelvi szöveghez hasonló eredményeket produkálnak, és nem olyan eredmények születnek, mint

az autentikus szövegek esetében. A táblázatok eredményeiből és az elemek csoportokba szelektálását követően az rajzolódik ki, hogy a szövegekben vannak ugyan formulaszerű szakkifejezések, de a szógyakoriságot döntő többségében mégis inkább a terminusok ismétlődése befolyásolja.

A jelen kutatás mindössze előzetes, kvantitatív eredményekkel szolgált a bevezetésben felvetett kutatási kérdést illetően, további hipotézisek megfogalmazását elősegítve, azonban a téma teljes körüljárása és a kutatási kérdés pontosabb megválaszolása érdekében egy jövőbeli elemzés részeként a kvalitatív szövegelemzés – statisztikai vizsgálatokkal kiegészítve – további fontos és hasznos információkkal bővítheti a műszaki szövegekről szóló ismereteket, amihez a Pannónia Korpusz kiváló alapként szolgálhat. Megmutatkozott az elemzések során, hogy pusztán szógyakorisági vizsgálatokkal nem lehetséges pontos képet nyerni a fordított és autentikus szövegek terminuskezelési sajátosságairól – a gépi elemzőprogram viszonylag nagy „zajjal" dolgozik, nem képes elválasztani a formulaszerű kifejezéseket a valódi terminusoktól, szükségessé téve a manuális elemzést (Robin és Seidl-Péch 2019), illetve nem derül fény a két és a négy tagból álló szókapcsolatok esetleges átfedésére. A problémát a már említett kvalitatív vizsgálatokkal lehetséges áthidalni, valamint részletesebb elemzést lehetővé tévő számítógépes programok, például a *Sketch Engine* terminuskezelő alkalmazásának segítségével.

Végezetül kitekintésként megemlíteném ezúttal is a természettudományos szövegek csoportját, mivel jelentős hasonlóságokat mutatnak a műszaki szakszövegekkel, gyakran említik őket rokon szövegtípusként is. A hasonlóságot főként a jelen tanulmányban többször is említett gyakori terminushasználatban láthatjuk, így egy következő kutatás keretében érdekes volna elvégezni terminuskezelési és gyakorisági elemzést természettudományi szövegekből álló párhuzamos és összehasonlítható korpuszon, majd összevetni a műszaki szövegeken végzett vizsgálódás 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. https://doi.org/10.1075/z.64.15bak

Blum-Kulka, S. 1986. Shifts of Cohesion and Coherence in Translation. In: House, J. Blum-Kulka, S. (eds) *Interlingual and Intercultural Communication. Discourse and Cognitionin 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* 14. 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. https://doi.org/10.1093/ijl/1.1.32

Heltai P., Juhász G. 2002. A névmások fordításának kérdései angol–magyar és magyar–angol fordításokban. *Fordítástudomány* 4. évf. 2. szám. 46–62.

Kamenická, R. 2008. Explicitation profile and translator style. In: Pym, A., Perekrestenko, A. (eds) *Translation Research Projects 1.* Tarragona: Intercultural Studies Group. 117–130.

Klaudy K. 1999. Az explicitációs hipotézisről. *Fordítástudomány* 2. évf. 1. szám. 5–22.

Klaudy K. 2004. Az EU-szakszövegek fordításának oktatása. In: Dobos Cs. (szerk.) *Miskolci nyelvi mozaik*. Budapest: Eötvös József Kiadó. 11–24.

Klaudy, K., Károly, K. 2000. The Text-organizing Function of Lexical Repetition in Translation. In: Olohan, M. (ed.) *Intercultural Faultlines. Research Models in Translation StudiesI. Textual and Cognitive Aspects.* Amsterdam: St. Jerome. 143–159. https://doi.org/10.4324/9781315759951-10

Klaudy, K., Károly, K. 2005. Implicitation in Translation: Empirical Evidence for Operational Asymmetry in Translation. *Across Languages and Cultures* Vol. 6. No. 1. 13–29. https://doi.org/10.1556/Acr.6.2005.1.2

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.

Nagy A. L. 2017. A szógyakoriság gépi vizsgálata műszaki szövegeken. *Fordítástudomány* 19. évf. 2. szám. 40–57.

Nattinger, J. R., DeCarrico, J. S. 1992. *Lexical Phrases and Language Teaching.* Oxford: Oxford University Press.

Olohan, M., Baker, M. 2000. Reporting that in Translated English. Evidence for Subconscious Processes of Explicitation. *Across Languages and Cultures* Vol. 1. No. 2. 141–159. https://doi.org/10.1556/Acr.1.2000.2.1

Pawley, A., Syder, F. H. 1983. Two puzzles for linguistic theory: Native selection and nativelike fluency. In: Richard, J. C., Schmidt, R. W. (eds) *Language and Communication.* London: Longman.

Robin E. 2018. *Fordítási univerzálék és lektorálás*. Budapest: ELTE Eötvös Kiadó.

Robin E., Dankó Sz., Götz A., Nagy A. L., Pataky É., Szegh H., Zolczer P. 2016. Fordítástudomány és korpuszkutatás: bemutatkozik a Pannónia Korpusz. *Fordítástudomány* 18. évf. 2. szám. 5–26.

Robin E., Seidl-Péch O. 2019. Alkalmas-e a korpuszalapú módszer a terminuskezelési stratégiák kutatására? Elhangzott: *Korpusz és kontrasztivitás a szakfordítás okta-*

tásában és gyakorlatában. Fordításoktatási szakmai nap. Budapest, KRE BTK. (2019. február 7.)

Seidl-Péch O. 2018. Melyek a (szak)fordító és a fordításkutató munkáját segítő legfontosabb nyelvi korpuszok? In: Robin E., Zachar V. (szerk.) *Fordítástudomány ma és holnap.* Budapest: L'Harmattan Kiadó. 175–191.

Sturcz Z. 2010. Megállapítások és felvetések a műszaki tudományok nyelvéről. In: Dobos, Cs. (szerk.) *Szaknyelvi kommunikáció*. Budapest: Tinta Könyvkiadó.

Toury, G. 1991. What are Descriptive Studies into Translation Likely to Yield apart from Isolated Descriptions? In: van Leuven-Zwart, K. M., Naaijkens, T. (eds) *Translation Studies: The State of the Art.* Amsterdam/Atlanta, GA: Rodopi. 179–192.

Toury, G. 1995. *Descriptive Translation Studies and Beyond*. Amsterdam: John Benjamins. https://doi.org/10.1075/btl.4

Vinay, J. P., Darbelnet J. [1958] 1995. *Comparative Stylistics of French and English. A methodology for Translation.* Amsterdam: John Benjamins. https://doi.org/10.1075/btl.11

Wray, A. 2002. *Formulaic language and the lexicon.* Cambridge: Cambridge University Press.

Weinert, R. 1995. The role of formulaic language in second language acquisition: A review. *Applied Linguistics* Vol. 16. No. 2. 180–205. https://doi.org/10.1093/applin/16.2.180

Források

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, William 2008. *Microsoft® Exchange Server 2007 Administrator's Pocket Consultant, Second Edition*. Washington: Microsoft.

Párhuzamos korpusz – *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. *Bemutatkozik 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.

Összehasonlítható korpusz – Autentikus 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.

Ficzere 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 légiközlekedésben: Az airside, pre-take-off objektum- és folyamatcsoport biztonságintegritása.* Doktori disszertáció. Budapest: BME.

Nagy A. 2014. *Kisméretű légieszközök mozgásfolyamatát meghatározó mérési és szimulációs környezet fejlesztése: siklóernyők fordulási tulajdonságainak elemzése.* Doktori disszertáció. Budapest: BME.

Ozsváth P. 2009. *Magnézium alapú hibrid járműanyagok környezetbarát forgácsolásának optimálása.* 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 belvízi 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.

Szegh, H. 2017. Harmadik kód a tolmácsolásban: létezik tolmácsolási szöveg? *Fordítástudomány* 19. évf. 1. szám. 60–74.

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

1	544	in the
2	439	the camera
3	302	can be
4	279	of the
5	230	will be
6	202	press the
7	185	and press
8	178	live view
9	176	to the
10	170	is selected
11	151	the shutter
12	150	white balance
13	148	custom setting
14	147	control panel
15	147	shutter release
16	144	not be
17	143	selected for
18	138	if the
19	138	when the
20	137	command dial
21	137	g button
22	134	be used
23	131	press j
24	129	focus point
25	128	with the
26	127	number of
27	126	release button
28	125	shutter speed
29	124	nef raw
30	124	on the
31	118	memory card
32	111	be displayed
33	111	can not
34	110	button is
35	102	information on

36	101	multi selector
37	100	for the
38	100	is pressed
39	100	the viewfinder
40	99	note that
41	98	custom settings
42	98	the battery
43	97	displayed in
44	96	choose the
45	93	iso sensitivity
46	93	shooting menu
47	92	built in
48	92	from the
49	91	the following
50	91	using the
51	86	button a
52	86	or to
53	84	and the
54	83	for information
55	82	camera is
56	82	the control
57	81	settings menu
58	81	the flash
59	81	the selected
60	81	to choose
61	80	the current
62	78	is not
63	78	the multi
64	77	do not
65	77	while the
66	76	fn button
67	76	the monitor
68	75	the lens
69	74	a custom
70	74	at the
71	72	the number

72	71	is displayed
73	70	button and
74	70	in flash
75	68	image area
76	68	j to
77	68	main command
78	67	the desired
79	67	the image
80	67	the shooting
81	67	use the
82	66	used to
83	65	rotate the
84	65	the memory
85	64	of shots
86	64	pressing the
87	63	choose a
88	62	press or
89	62	the subject
90	61	by the
91	61	is used
92	60	sub command
93	60	the focus
95	59	picture control
96	59	the built
97	59	the center
98	57	center of
99	57	image sensor
100	57	the main

Az indiai vallási irodalomban használatos beszélőjelölő igéjének magyar fordítása

Bakaja Zoltán

zoltan.bakaja@gmail.com

Bhaktivedanta Hittudományi Főiskola

Kivonat: A *Bhagavad-gītā* szövegében a szanszkrit irodalom néhány más művéhez hasonlóan szerepel az idéző mondatnak az a kimondottan e művekre jellemző formája, amelyet beszélőjelölőnek neveztem el (Bakaja 2019). A jelen tanulmány a *Bhagavad-gītā* beszélőjelölőjében található ige magyar fordítási lehetőségeit vizsgálja. Számba veszi az eddig megjelent kiadások megoldásait, és a Magyar Nemzeti Szövegtár, valamint a Pannónia Korpusz (Robin et al. 2016) segítségével keres választ arra a kérdésre, hogy egy modern magyar fordításban a beszélőjelölő idéző igéjének magyar fordítására a *beszél,* a *mond* vagy a *szól* ige valamely alakja lenne-e a legjobb választás. Mivel számos magyar fordítás angol közvetítőnyelvből készült, a kutatás arra is kiterjed, hogy az angol forrásnyelvi szövegekben a jelölő igéjeként szereplő *said* magyar fordításakor a fordítók a *mondta* és a *szólt* közül melyik igét részesítik előnyben. A kutatás eredménye megerősíti azt az előzetes vélekedésemet, hogy a szanszkrit idéző ige modern magyar nyelvre való lefordítására a *mond* szótőből képzett alakok megfelelőbbek a *szól* tőből képzetteknél.

Kulcsszavak: *Bhagavad-gītā,* beszélőjelölő, idéző ige, szól, beszél, mond

1. A beszélőjelölő igéjének hagyományos fordítása

Egy korábbi cikkemben az általam beszélőjelölőnek nevezett szanszkrit idéző mondat angol és magyar nyelvű fordításaiban található átváltási műveleteket vizsgáltam (Bakaja 2019). (A beszélőjelölő definíciójáról egy későbbi bekezdésben írok.) A kutatási korpusz magában foglalta a *Bhagavad-gītā* néhány angol és az összes magyar nyelvű

Bakaja Zoltán: Az indiai vallási irodalomban használatos beszélőjelölő igéjének magyar fordítása. In: Robin E., Seidl-Péch O. (szerk.) 2020. *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordításkutatás Magyarországon.* Segédkönyvek a nyelvi közvetítésről I. Budapest: ELTE BTK Fordítástudományi Doktori Program, MANYE Fordítástudományi Szakosztály. DOI: https://doi.org/10.36252/Nyelvikozvsegedkonyv1.9

kiadását (lásd a forrásjegyzéket). Feltűnt, hogy noha a fordítók a beszélőjelölő idéző igéjének átültetése során több lehetőség közül is választhattak volna, mind az angol, mind a magyar kiadásokban megfigyelhetünk egy fordítási hagyományt, amelytől csak nagyon ritkán tértek el.

Az említett cikk megírására Klaudy Kingának az a megállapítása ösztönzött, hogy az idéző igéknek a fordítás során történő konkretizálása a magyar ige sajátosságaiból fakad (1999: 45), és az okát a magyar stilisztikai tradíciókban kell keresnünk (1999: 49). Klaudy több munkájában is foglalkozik az idéző igék fordításával (Klaudy 2007; Klaudy és Simigné 2000; Klaudy és Károly 2005), Simignével közös könyve néhány példát is tartalmaz a magyar igék angolra való fordításakor alkalmazott generalizálás szemléltetésére (2000: 182–183). Amikor észrevettem, hogy az egyik angol *Gītā*-kiadás néhol konkretizálta a szanszkrit idéző igét, úgy döntöttem, hogy néhány angol és az összes magyar kiadáson is megvizsgálom, mi történik a beszélőjelölő igéivel.

A beszélőjelölő *vac* gyökből képzett idéző igéje a *Bhagavad-gītā*-ban csak egyes szám harmadik személyű, elbeszélő múlt idejű alakban található meg: *uvāca.* A *vac* magyar jelentését meghatározandó, Körtvélyesi Tibor szanszkrit nyelvtankönyvének gyöktáblázatában a *beszél* ige szerepel (2006: 255), a Monier-Williams (2011) szótárban található angol jelentések pedig a következők: *‘to speak, say, tell, utter, announce, declare, mention, proclaim, recite, describe’*.

Korábbi kutatásomból kitűnt, hogy amikor a fordítók nem konkretizálták az igét, az angol kiadásokban az *uvāca* megfelelőjeként leggyakrabban a *said,* míg a magyarokban a *szólt* szerepel, ezeket tekinthetjük tehát a beszélőjelölő idéző igéjének adott nyelvre jellemző hagyományos fordításának. Ebben a tanulmányban leginkább azt szeretném megvizsgálni, hogy a *Bhagavad-gītā* egy modern magyar kiadásában mennyire célszerű választás a hagyományos *szólt* alak használata.

2. Az ige magyar nyelvre való átültetése során használt fordítási technikák

Kezdjük a nem konkretizált magyar igealakok fordítása során alkalmazott technikák áttekintésével, Klaudy (1999) átváltási műveletekre kidolgozott tipológiája alapján. Ezek között az *uvāca* igével többé-kevésbé ekvivalens alakok (*szólt, szólott, mondta*) egyszerű behelyettesítése mellett megtaláljuk a kihagyást és a grammatikai cserét is. Ahogy azt már említettem, a *szólt* alak a beszélőjelölő igéjének hagyományos magyar fordítása, a *Bhagavad-gītā* majdnem minden hazai kiadásában megtalálható. A *mondta* Dvārakeśa Dāsa magyar fordításának három versszakában fordul elő (Bhaktivedanta é. n. 1.46, 2.1, 11.9), a *szólott* pedig ugyanebben és egy másik kiadásban (Bakos A. és Bakos J. E. 2017) is többször megjelenik. Több magyar *Bhagavad-gītā* fordításban szerepel gyakran a nyomatékosító elem betoldása révén létrejött *így szólt* és *ekképp szólt* kollokáció, míg Baktay Ervin (2013) mindenütt az ige elhagyását, a lexikai kihagyást alkalmazza. Gömöryné Maróthy Margit (1924) és Szabó Ágnes (Sivananda é. n.) grammatikai cserével él, amikor a jelen idejű *szól* alakot használja.

2.1. A beszélőjelölő fogalma

Annak megítéléséhez, hogy vajon minden fordítói megoldás egyformán megfelelő-e, meg kell értenünk, mit is takar pontosan a *beszélőjelölő* terminus. Korábban már említett tanulmányomban így írtam a fogalomról:

> Az idéző mondat azon formáját nevezem beszélőjelölőnek, amelyet a *Mahābhāratában,* az *upaniṣadok* némelyikében (*Haṁsa, Muktika, Rāmarahasya, Śukarahasya*) és a *purāṇákban* használnak, de a *Rāmāyanában* például nem találhatjuk meg. Nem tartozik szervesen a szöveghez, annak többi részétől eltérően nem valamely versmértékben íródott, és néhány [*Bhagavad-gītā*] kiadásban több, másokban kevesebb van belőle. (Bakaja 2018: 757)

A beszélőjelölő szerkezete: *alany + idéző ige* (ez a legtöbb esetben a *vac* gyökből képzett egyes szám harmadik személyű alak: *arjuna uvāca*) vagy *jelző + alany + idéző ige* (*śrī bhagavān uvāca*). Csupán annyit jelez, hogy egy szereplő beszélt, nem utal és nem hívja fel a figyelmet a beszéd tartalmára vagy regiszterére. A nem konkretizált igealakot választó fordításban ezért egy általános jelentésű ige használata által lehet legjobban visszaadni a jelentését.

2.2. A beszélőjelölő igéjének fordítása során használt átváltási műveletek

Mivel a szerkezet feladata a beszélő személyének a megjelölése, nem tekinthetők helyesnek azok a fordítói megoldások, amelyek esetében a fordítók egy határozószó betoldásával megváltoztatták a szerkezet fókuszát, és ennek eredményeképpen az idéző mondat már nem azt hangsúlyozza, hogy *ki beszélt,* hanem azt emeli ki, hogy *mit mondott* (például *így szólt, ezt mondta*). A beszélőjelölő olyanná teszi a művet, mintha egy színdarab szövegkönyvét vagy egy film forgatókönyvét olvasnánk. A bennük szereplő nevek csak a színészeknek szólnak, csak azért vannak rögzítve, hogy az előadók tudják, mikor kell megszólalniuk, az előadás alatt vagy a moziban azonban nem halljuk őket. Emiatt Baktay, Gömöryné és Szabó Ágnes fordítói megoldásai töltik be legjobban a beszélőjelölő funkcióját.

Baktay Ervin (2013) kihagyja az igét, csak a beszélőjelölőben található neveket tartja meg, és ezeket kiskapitális szedéssel, utánuk kettőspontot téve szerepelteti a műben:

> Ardzsuna: […]
>
> Szandzsaja: […]

Mintha csak a *Mahābhārata* megfilmesített változatának szövegkönyvét olvasnánk. Gömöryné Maróthy Margit és Szabó Ágnes azáltal ér el hasonló kontextuális hatást, hogy az idéző igét nem múlt, hanem jelen idejű alakban használja:

Ardzsuna szól: […]

Szandzsaja szól: […][1]

A színházban még akkor is jelen időben hangzanak el a beszélgetések, ha a darab a régmúlt eseményeit idézi fel. Talán nem véletlen, hogy a színésznő Gömöryné élt először ezzel a megoldással.

2.3. Az egyszerű behelyettesítés során használt igék

A beszélőjelölő fordítása során a magyar fordítók három *egyszerű behelyettesítéssel* kapott igealakot használnak a *Bhagavad-gītā* különböző kiadásaiban. Ezek közül a *szólt* és a *szólott* ugyanannak az igének a különféle múlt idejű alakjai, a fordítók tehát többnyire két magyar ige, a *szól* és a *mond* egyes szám harmadik személyű, múlt idejű alakjaival fordították le a szanszkrit *uvāca* szót. Érdekes, hogy egyik magyar *Bhagavad-gītā* beszélőjelölőjében sem találhatunk a Körtvélyesi által megadott *beszél* jelentésből képzett igealakot.

A *beszél* igének a *Magyar értelmező kéziszótár* (Pusztai 2011) *által* megadott nyolc jelentéscsoportja közül hármat vonatkoztathatunk a *Gītā* párbeszédeire:

(1) ‘gondolatait, élményeit élőszóval kifejezi’,

(2) ‘szót vált, tárgyal, társalog vkivel’,

(3) ‘beszédet, előadást tart’.

1 A példában Gömöryné fordításának helyesírását használtam. Szabó munkájában az Ardzsúna és Szandzsája alak szerepel, ezek azonban nem tekinthetők helyesnek.

Ezek közül az elsővel részben megegyezik (a verbális közlésre vonatkozólag) a *mond* igének a szótárban található egyik jelentése: ‘élőszóban v. írásban kifejez, közöl’; a *szól* igének a szótárban először említett jelentése pedig: ‘szavakat mond, beszél’. Véleményem szerint a *Gītā* kontextusában a *beszél* megfelelőbb lenne a beszélőjelölő idéző igéjének fordítására, mert általánosabban utal magára a beszéd aktusára, mint a fordítók által hagyományosan használt magyar igék.

Ha most nekem kellene magyar nyelvre átültetnem a *Bhagavad-gītā* szövegét, a fent említett okokból vagy Baktay megoldását tartanám követendőnek, vagy jelen idejű alakot használnék a beszélőjelölő igéjének a lefordítására. E két megoldás közül azonban egyik sem kötelező átváltási művelet (Klaudy 1999), és a könyvnek több olyan magyar kiadása is megjelent, amelyet valamelyik angol fordítás alapján készítettek. Ezek fordítói nem akartak élni a fentebb említett lehetőségekkel. Az angolból fordított könyvek is követik a jelölőben szereplő ige *szólt* szóval való fordításának hagyományát, noha a szótárban az angol kiadásokban használt *said* megfelelőjeként mindhárom korábban említett magyar igét (*szól, beszél, mond*) megtaláljuk (Országh és Magay 2009).[2]

A *szól* vagy *szólt* használata az adott kontextusban meglehetősen régimódinak tűnhet, ezért egy olyan modern fordításban, amelyben nem cél a szöveg szándékosan régies hangzásúvá tétele, véleményem szerint sem jelen, sem múlt idejű alakban nem fordulhatna elő rendszeresen. Feltételeztem, hogy a *mondja* vagy a *mondta* a beszélőjelölő igéjének fordítására legalkalmasabb általános jelentésű, hétköznapi regiszterű idéző ige. Véleményem megalapozottságának bizonyítása érdekében két szövegkorpuszt hívtam segítségül: a Magyar Nemzeti Szövegtár (a továbbiakban MNSZ) (Várdai és Oravecz 2014) 2_v2.0.5 változatát és a Pannónia Korpuszt (a továbbiakban PC) (Robin et al. 2016).

2 Az angol szavak magyar jelentését mindenütt ebből a szótárból idézem.

3. A vizsgált igék az MNSZ és a PC szövegállományában

A nyelvhasználatról való képalkotás legkézenfekvőbb módszere a szövegkorpuszok vizsgálata. Az idéző igék korpuszalapú kutatásának rendkívül gazdag szakirodalma van. Ebben keleti nyelvekkel kapcsolatos fordítástudományi kutatásokat is találunk (Yu és Guo 2016; Yeganeh és Boghayeri 2014; Ardekani 2002), de szanszkrit idéző igékkel vagy fordításukkal foglalkozó munka még nem került a kezembe. Seidl-Péch Olívia cikke (2018) számos magyar nyelvű korpuszt mutat be, Amba Kulkarni írásából (2016) pedig a szanszkrit műveket magukban foglaló adatbázisokat ismerhetjük meg. Az általam használt két korpusz közül a Magyar Nemzeti Szövegtárra – a magyar nyelv reprezentatív referenciakorpuszára – mérete és az irodalmitól a beszéltnyelvi szövegekig terjedő sokféle benne foglalt szövegtípus miatt esett a választásom. Az MNSZ jelenlegi összetételébe az alábbi táblázat nyújt betekintést. Az adatok millió szóban, százezer szóra kerekítve vannak megadva.[3]

1. táblázat

Az MNSZ összetétele

Alkorpusz	**Szavak száma**
sajtó	364,8 (35,09%)
szépirodalom	80,6 (7,75%)
tudományos	117,9 (11,34%)
személyes	301,1 (28,96%)
hivatalos	99,0 (9,52%)
beszéltnyelvi	76,2 (7,33%)
Összesen	1039,7

A korpuszban megtalálható autentikus vagy fordítás eredményeképp keletkezett szövegek arányáról sem a fentebb hivatkozott cikk, sem a korpusz honlapja nem ad felvilágosítást, ezért megkerestem e-mailben az üzemeltetőket. Válaszukból kiderül, hogy az MNSZ többnyire autentikus magyar szövegeket tartalmaz, a fordítások mennyisége

3 Az MNSZ adatai: http://clara.nytud.hu/mnsz2-dev/stat.html

nem jelentős, „valószínűleg 1 százalék alatti”. Tájékoztatásuk szerint „a ‚off_hu_jrc_.*’ és a ‚press_hu_eunews’ dokumentum nagy valószínűséggel fordítás, a többi nagy valószínűséggel eredeti”. A korpusz 2951 dokumentumából 21 a fentebbi kritériumoknak megfelelőt találtam, ezek között 20 hivatalos, egy pedig sajtószöveg. A három említett ige előfordulásának arányát a teljes korpuszban, és – mivel ezek használatát a *Bhagavad-gītā* szövegében vizsgálom – a szépirodalmi alkorpuszban is kutattam. Ez utóbbi a kapott tájékoztatás szerint nem tartalmaz fordításokat, így lehetőséget nyújt a vizsgált igék tisztán autentikus írásokban való előfordulásának tanulmányozására.

A kiegyensúlyozottság a korpuszok reprezentativitásának fontos jellemzője (lásd Seidl-Péch 2018: 177–180). Az MNSZ-ben a személyes és a sajtószövegek mennyisége mintegy háromszorosa a többi szövegfajtának, így az arányai nem ideálisak, ám mivel kutatásom szempontjából a szépirodalmi alkorpusznak volt igazi jelentősége, nem tekintettem problémának a teljes korpuszra vonatkozó eredmények lekérdezését. A teljes korpusz arányainak ismeretében különösen érdekes az egyes tanulmányozott igéknek a teljes korpuszon és a szépirodalmi alkorpuszon belüli aránya.

A PC szépirodalmi alkorpuszát egyrészt azért vizsgáltam, mert a benne található szövegek többsége 2000 után keletkezett, ezért a modern nyelvhasználatot tükrözi, másrészt ez a korpusz fordítási korpusz, és a benne megtalálható angol–magyar párhuzamos szövegek kutatásával képet kaphatunk arról, hogy az angol *said* igét a szépirodalmi művekben a *beszél,* a *mond* és a *szól* közül mely magyar igével fordítják a leggyakrabban.

Bár a jelen kutatásom során elsődlegesen arra kerestem a választ, hogy a *mond* valóban a leggyakrabban használt, leghétköznapibban hangzó idéző ige-e a magyar nyelvben, és így valóban ennek egyes szám harmadik személyű, múlt idejű alakja-e a legjobb választás az *uvāca* múlt időben való fordításakor, az igék jelen idejű alakjait is bevontam a kutatásba, mivel a beszélőjelölő igéjének jelen időben való fordítása a fenti okfejtés alapján célszerű stratégiának vélhető.

Azt feltételeztem, hogy a három ige közül a *mond* különféle alakjai mindkét korpuszban nagyobb számban fordulnak elő a többinél. Mivel a PC szépirodalmi alkorpuszá-

ból csak 2000 után keletkezett magyar szövegeket vontam be a kutatásba, azt vártam, hogy ezekben a nyomatékosító betoldás nélkül szereplő *mondta* a betoldást szintén nélkülöző *szólt* igéhez képest az MNSZ találatainál is nagyobb arányban fordul elő, alátámasztva azt a vélekedésemet, miszerint egy modern kiadásban az előbbi használata lenne szerencsésebb.

3.1. A vizsgált igék az MNSZ teljes szövegállományában

Az MNSZ használata során az egyszerű keresés lehetőségét választva nemcsak a beírt szóalak előfordulásait kapjuk meg, hanem a rendszer szótőként is keresi a megadott kifejezést, így a szótőből képzett különféle szófajú és ragozású alakjainak a listájához is hozzájutunk. Az egyszerű keresést választva a három szótőből képzett szavak száma és az igék keresése során kapott eredmények összegéhez viszonyított aránya (a százalékos értékeket eltérő közlés híján a többi táblázatban is hasonlóan számoltam ki) a következőképpen alakul:

2. táblázat

A vizsgált igék egyszerű kereséssel kapott lekérdezésének eredménye

beszél	529 152 (15,28%)
mond	2 131 428 (61,55%)
szól	802 327 (23,17%)

A *3. táblázat* soraiban az igék pontos szóalakjára való kereséssel kapott jelen és múlt idejű, egyes szám harmadik személyű alakjainak előfordulása látható, ahogyan azok a beszélőjelölőben szerepelhetnének.

3. táblázat

A vizsgált igék jelen és múlt idejű alakjainak száma az MNSZ-ben

beszél	74 733 (5,28%)
beszélt	112 828 (7,98%)

mondja	238 028 (16,83%)
mondta	749 593 (53,02%)
szól	156 779 (11,09%)
szólt	81 938 (5,79%)

Kutatásom célja az *uvāca* ekvivalenseként használható legáltalánosabb idéző ige megtalálása, ám a három vizsgált magyar ige nem kizárólag idéző funkciót tölthet be a magyar nyelvben. A több millió találat kvalitatív elemzése kivitelezhetetlen, ezért a keresést megpróbáltam úgy leszűkíteni, hogy a találatok a lehető legpontosabban megmutassák a kutatás tárgyát képező igék idéző szerepet betöltő előfordulásainak arányát.

Erre az egyik lehetőség az igék azon alakjainak megkeresése, amelyek előtt valamilyen írásjel található: vessző, felkiáltójel vagy gondolatjel. Az ige után álló kettőspont ugyancsak idéző funkcióra utalhat. Mind közül az ige előtti felkiáltójel keresése adta a legegyértelműbb találatokat, ezeket a soron következő táblázat összegzi. A találatok elváló igekötős alakokat is tartalmaztak, számukat kivontam a kapott értékekből. Az igekötős alakoknak az eredeti eredményben található mennyiségét az igék százalékos aránya után látható zárójelben lévő szám jelzi. Az igekötős alakok kizárása után csupán a *mondta,* a vizsgáltak közül messze a legnagyobb arányban előforduló szó találatai között volt két olyan, amelyben az ige nem idéző funkciót töltött be (e kettőt is kivontam a kapott 111-ből).

4. táblázat

A felkiáltójel után álló idéző igék az MNSZ-ben

beszél	1 (0,54%)
beszélt	0
mondja	58 (31,35%) (1)
mondta	108 (58,37%) (1)
szól	10 (5,40%) (2)
szólt	8 (4,32%) (5)

A vizsgált igék vessző és gondolatjel után álló, valamint kettőspont előtti alakjaiból olyan sok fordul elő a korpuszban, hogy lehetetlen őket kvalitatív vizsgálatnak alávetni, ám e jelek igék előtti, illetve utáni használata nem garantálja, hogy az utánuk, illetve előttük álló igék mindegyike idéző funkciót tölt be. Annak érdekében, hogy képet kaphassak arról, milyen arányban töltenek be idéző funkciót a kutatásom tárgyát képező kifejezések, a korpuszban fellelhető véletlenszerű keverés lehetőségét kihasználva összekevertem a találatokat, majd 30-30-at kiválasztottam közülük, és az így kapott mintát elemeztem. Az elemzés eredményét az alábbi táblázat foglalja össze.

5. táblázat

A vessző és gondolatjel utáni, valamint kettőspont előtti igék

Igék	**Vessző utáni**	**Idéző**	**Gondolatjel utáni**	**Idéző**	**Kettőspont előtti**	**Idéző**
beszél	871 (1,86%)	5/30	151 (0,04%)	19/30	1 284 (1,15%)	21/30
beszélt	989 (2,11%)	15/30	455 (0,11%)	29/30 (1)	3 280 (2,95%)	30/30
mondja	14 694 (31,32%)	20/30 (4)	58 681 (14,04%)	29/30 (1)	20 127 (18,11%)	30/30
mondta	27 888 (59,45%)	28/30 (1)	350 522 (83,89%)	29/30 (1)	71 296 (64,17%)	30/30
szól	1 147 (2,44%)	13/30 (2)	3 008 (0,72%)	27/30 (3)	6 755 (6,08%)	29/30
szólt	1 320 (2,81%)	17/30 (1)	4 996 (1,19%)	24/30 (6)	8 361 (7,52%)	30/30

A fenti táblázat *Idéző* címmel jelölt oszlopaiban az látható, hogy a harminc megvizsgált találatból mennyi volt az igék idéző funkciójú előfordulása. Az ezt jelző tört utáni zárójeles számok az elváló igekötős alakok mennyiségét mutatják. Mivel kutatásom tárgyát a *beszél, mond* és *szól* igekötő nélküli alakjainak vizsgálata képezte, az igekötős alakokat kizártam a további vizsgálatból. A táblázatnak a *szólt* gondolatjel utáni előfordulását bemutató részletében például az ige minden egyes esetben idéző funkciót töltött be, de mivel 6 alkalommal igekötő is társult hozzá, ez a hat nem szerepel az érvényes találatok között, így az idéző funkciójú találatok arányszáma 24/30.

Az *5. táblázat* első változatában feltűnően sokszor fordult elő a gondolatjel utáni *szólt* igekötős alakja (lásd 6. táblázat).

A *szólt* 14 igekötős előfordulása közül 6-6 a *rá* és a *közbe* igekötőt tartalmazta, ezért a keresést úgy is elvégeztem, hogy kizártam e két igekötőt, s ez került az 5. táblázatba.

6. táblázat

A gondolatjel utáni igék a rá és közbe elváló igekötő kizárása előtt

	Gondolatjel utáni	**Idéző**
beszél	151 (0,04%)	19/30
beszélt	455 (0,11%)	29/30 (1)
mondja	58 681 (14,01%)	29/30 (1)
mondta	350 522 (83,70%)	29/30 (1)
szól	3 008 (0,72%)	27/30 (3)
szólt	5 948 (1,42%)	14/30 (14)

Az MNSZ szövegeiből nem szűrhető ki automatikusan egy ige összes igekötős alakja, mert a kereső nem tudja megkülönböztetni, hogy egy adott szó az aktuális mondatban elváló igekötő vagy éppen határozószó-e, ám az egyes igekötők célzott kizárása megoldható. (Noha a szóalakra való keresés során elvileg kizárhatók a szó igekötős formái, a gyakorlatban ez nem működött megbízhatóan, ezért nem használtam.) Az *5. táblázat* tehát a gondolatjel után álló *szólt* ige *rá* és *közbe* igekötőinek kizárásával kapott adatokat tartalmazza.

A táblázatból kitűnik, hogy az igéket az utánuk álló kettősponttal együtt keresve a legnagyobb annak az esélye, hogy a találatban kapott igék idéző funkciót töltenek be, és nem igekötősek, míg ugyanez az esély jóval kisebb, ha az igékre az előttük álló vesszővel együtt keresünk rá. Ez alól csak a *mondja* és a *mondta* képezett kivételt, mivel e két szó az esetek nagy többségében még a vessző utáni keresés esetén is igekötő nélküli idéző ige. Az *5. táblázat* azt is megmutatta, hogy az idéző funkciójú igék megtalálásának a legjobb módszere az előttük álló gondolatjellel, illetve az utánuk álló kettősponttal való együttes keresés. Bár a vesszővel megelőzött igékre is kaphatunk találatokat, ez esetben az eredmény sokkal kevésbé biztos, mint az előbbiekben.

A következő táblázat azt mutatja meg, hogy az egyes igék vessző és gondolatjel utáni, valamint kettőspont előtti előfordulásának összege hogyan aránylik az igéknek a korpuszon belüli összes előfordulásához.

7. táblázat

Az igék három írásjellel együtt vizsgált alakjának aránya teljes korpuszbeli előfordulásukhoz képest

	Teljes korpusz	**Írásjelekkel vizsgált**
beszél	74 733	2 306 (3,08%)
beszélt	112 828	4 724 (4,19%)
mondja	238 028	93 502 (39,28%)
mondta	749 593	449 706 (59,99%)
szól	156 779	10 910 (6,96%)
szólt	81 938	14 677 (17,91%)

A *4–7. táblázatból* kiderül, hogy a hat ige közül a *mondja* és a *mondta* tölt be leggyakrabban és legnagyobb mennyiségben idéző funkciót, őket pedig a *szólt* követi.

A *7. táblázatban* láthatjuk, hogy bár a teljes korpuszban a *beszél* igének több múlt idejű alakja fordul elő, mint a *szól* igének, az idéző funkcióra utaló környezetben ennek az ellenkezője igaz. A *7. táblázat* azt is megmutatja, hogy a *mond* alakjainak száma nemcsak meghaladja a másik két ige azonos idejű alakjainak, sőt azok összegének a számát is, de idéző funkciót is sokkal nagyobb arányban töltenek be, mint a *beszél* és a *szól* vizsgált alakjai.

A beszélőjelölő fontos jellegzetessége, hogy az idézet előtt áll. Írásban a benne található igét csak egy alanyesetű személynév vagy egy melléknév és egy személyre utaló főnév (*śrī-bhagavān* 'szent istenség') előzi meg, és a szöveg fordításaiban kettőspont követi. A korpuszban hasonló találatokat igyekeztem kapni, ezért a kettőspont előtt álló igék kilistázása után a szűrés opciót választva megkerestem azokat az eseteket, amelyekben alanyesetű főnév előzi meg őket. Úgy szűkítettem tovább a keresést, hogy a főnevek

előtt mondatvégi írásjel legyen, így a szerkezet a találatokban a mondat elejére került. Az így kapott minta már elég kicsi volt ahhoz, hogy kvalitatív keresést végezve kiszűrjem a hibás találatokat, illetve azokat az eseteket, amelyekben az ige nem idéző funkciót tölt be. A szanszkritban melléknévvel kezdődő szerkezethez hasonló találatok kinyerése végett a *névelő-melléknév-főnév-ige-kettőspont* keresést is lefuttattam. A végeredmény a *8. táblázatban* látható.

8. táblázat

Az idéző igék a beszélőjelölőével megegyező szerkezetekben az MNSZ-ben

	Főnévvel	**Névelővel**	**Összes**
beszél	4 (0,65%)	0	4 (0,64%)
beszélt	3 (0,48%)	1 (11,11%)	4 (0,64%)
mondja	540 (87,80%)	4 (44,44%)	544 (87,18%)
mondta	45 (7,31%)	2 (22,22%)	47 (7,53%)
szól	18 (2,92%)	1 (11,11%)	19 (3,04%)
szólt	5 (0,81%)	1 (11,11%)	6 (0,96%)

E keresés eredményében is a *mond* alakjai szerepelnek a legtöbb, a *beszél* ige alakjai pedig a legkevesebb találatban. Mivel a *Bagavad-gītā* fordításának lehetőségeit kerestem, és a szépirodalmi regiszter megváltoztathatja a szavak gyakoriságát, a fent ismertetett kutatás lépéseit az MNSZ szépirodalmi alkorpuszában is elvégeztem.

3.2. A vizsgált igék az MNSZ szépirodalmi alkorpuszában

Az MNSZ szépirodalmi alkorpuszát az egyszerű kereséssel megvizsgálva a *9. táblázatban* látható számokat kaptam. Az eredményt a *2. táblázat* adataival összehasonlítva megállapíthatjuk, hogy az igék szótőként való keresésének számai szerint a *mond* alakjai a *szól* alakjainak rovására nagyobb arányban fordulnak elő a szépirodalmi alkorpuszban, mint az MNSZ teljes szövegállományában. A *szól* alakjainak arányszáma olyan mérték-

ben csökkent, hogy a *beszél* alakjai kerültek a gyakorisági lista második helyére, noha a *beszél* alakjainak a másik két igéhez viszonyított aránya az alkorpuszban 0,1 százalékkal kevesebb, mint a teljes korpuszban.

9. táblázat

A vizsgált igék egyszerű keresésének eredménye a szépirodalmi alkorpuszban

beszél	71 361 (15,18%)
mond	330 545 (70,30%)
szól	68 254 (14,52%)

Az MNSZ teljes szövegállományán végzett elemzések után a szépirodalmi alkorpuszban is megvizsgáltam az igék jelen és múlt idejű, egyes szám harmadik személyű alakjainak előfordulását, ahogyan azok a beszélőjelölő különféle fordításaiban szerepelhetnének, a kapott adatok a következő táblázat soraiban láthatók.

10. táblázat

A vizsgált igék jelen és múlt idejű alakjainak száma a szépirodalmi alkorpuszban

beszél	12 414 (5,71%)
beszélt	11 557 (5,31%)
mondja	51 142 (23,51%)
mondta	107 831 (49,57%)
szól	15 588 (7,17%)
szólt	18 988 (8,73%)

A táblázat tanúsága szerint a kutatás tárgyát képező igealakok arányszáma a teljes korpusz adataihoz képest (*3. táblázat*) fordítva változott, mint a szótövekre történő kereséskor, a *szól* alakjai (ha számukat összeadjuk) így a második helyen állnak. Az alkorpuszban a *mond* alakjai is nagyobb arányban fordulnak elő, mint a teljes korpuszban. A szépirodalmi alkorpusz a teljes MNSZ 7,75 százalékát teszi ki. Ennek az aránynak az ismeretében a *11. táblázat* azt mutatja meg, hogy a hat igealak az alkorpuszban a teljes adatbázishoz képest felül- vagy alulreprezentált-e.

11. táblázat

Az alkorpusz igéinek aránya a teljes MNSZ ugyanazon igéinek számához viszonyítva

	Teljes MNSZ	**Alkorpusz**	**Arány**
beszél	74 733	12 414	16,61%
beszélt	112 828	11 557	10,24%
mondja	238 028	51 142	21,48%
mondta	749 593	107 831	14,38%
szól	156 779	15 588	9,94%
szólt	81 938	18 988	23,17%

Mivel kutatási kérdésemre csak az igék idéző funkciójú alakjainak vizsgálata adhatja meg a biztos választ, a teljes korpusz igéihez hasonlóan az alkorpusz igéinek is megkerestem a vessző, felkiáltójel és gondolatjel utáni, valamint kettőspont előtti alakjait.

Az idéző igék felkiáltójel utáni előfordulása a *12. táblázatban* látható eredmények alapján hasonló arányokat mutat a szépirodalmi alkorpuszban, mint a teljes korpuszban (*4. táblázat*). Ez nem meglepő, hiszen a találatok nagy része ebből az alkorpuszból került ki, noha mérete csak 7,75 százaléka az MNSZ teljes méretének.

12. táblázat

A felkiáltójel után álló idéző igék a szépirodalmi alkorpuszban

beszél	1 (0,91%)
beszélt	0
mondja	37 (33,64%)
mondta	62 (56,36%) (1)
szól	3 (2,73%) (2)
szólt	7 (6,36%) (3)

A vessző és gondolatjel utáni, valamint kettőspont előtti igék száma az alkorpuszban is meghaladja a kvalitatív feldolgozást lehetővé tevő nagyságot, ezért itt is hasonlóan jártam el, mint a teljes MNSZ igéinek vizsgálatakor. A *13. táblázat* adatai között az *Idéző*

címmel jelölt oszlopokban az látható, hogy a 30 megvizsgált találatból mennyi volt az igék idéző funkciójú előfordulása. Az ezt jelző tört utáni zárójeles számok az igekötős, és a keresésből kizárt alakok mennyiségét mutatják.

13. táblázat

A vessző és gondolatjel utáni, valamint kettőspont előtti igék a szépirodalmi alkorpuszban

Igék	Vessző utáni	Idéző	Gondolatjel utáni	Idéző	Kettőspont előtti	Idéző
beszél	317 (1,55%)	4/30	33 (0,04%)	12/30 (1)	307 (1,86%)	25/30
beszélt	329 (1,60%)	4/30	67 (0,08%)	22/30 (2)	333 (2,02%)	28/30
mondja	7 184 (35,04%)	27/30 (4)	16 167 (19,73%)	30/30	4 590 (27,81%)	30/30
mondta	11 565 (56,41%)	30/30 (1)	61 465 (75,00%)	30/30	7 127 (43,18%)	30/30
szól	411 (2,00%)	16/30 (3)	997 (1,21%)	21/30 (9)	1 028 (6,23%)	29/30
szólt	695 (3,39%)	16/30 (12)	3 228 (3,94%)	19/30 (11)	3 120 (18,90%)	29/30 (1)

Az *5. táblázathoz* hasonlóan a *szólt* gondolatjel utáni előfordulásainak esetén a *13. táblázat* is a *rá* és *közbe* igekötőt tartalmazó találatok kizárása után kapott eredményt tartalmazza. Az igekötők kizárása előtti számok a *14. táblázatban* láthatók.

14. táblázat

Az gondolatjel utáni igék a rá és a közbe igekötő kizárása előtt a szépirodalmi alkorpuszban

	Gondolatjel utáni	Idéző
beszél	33 (0,04%)	12/30 (1)
beszélt	67 (0,08%)	22/30 (2)
mondja	16 167 (19,55%)	30/30
mondta	61 465 (74,32%)	30/30
szól	997 (1,20%)	21/30 (9)
szólt	3 979 (4,81%)	16/30 (14)

A *13*. és az *5. táblázat* összevetéséből az derül ki, hogy a három ige megfelelő alakjainak egymáshoz viszonyított aránya az alkorpuszban nagyon hasonló a teljes korpuszéban előfordulókhoz. A *szól* alakjainak használata szinte minden esetben gyakoribb, mint a teljes korpuszban, a gondolatjel utáni és kettőspont előtti esetekben az arány növekedése néhol az 50 százalékot is meghaladja. A *beszél* majdnem minden alakjára ennek ellenkezője igaz. A szépirodalmi szövegekben a *szól* igekötős előfordulásának aránya jelentősen nagyobb, mint a teljes korpuszban, a *beszélt* pedig jóval kisebb arányban tölt be idéző funkciót vessző utáni előfordulásaiban. A ritkábban használt igéknek nemcsak a mennyisége kisebb, általában idéző funkcióban is ritkábban jelennek meg, mint a gyakoribbak. A *14. táblázatból* láthatjuk, hogy a *szól* igéhez társuló *rá* és *közbe* igekötők száma a szépirodalmi alkorpuszban is jelentős. A *15. táblázat* a hat igealak teljes alkorpuszon belüli, és potenciális idéző funkciós helyzetben lévő előfordulásának arányát mutatja meg.

15. táblázat

Az igék három írásjellel együtt vizsgált alakjának aránya a szépirodalmi alkorpuszban a teljes alkorpuszbeli előfordulásukhoz képest

	Teljes alkorpusz	**Írásjelekkel vizsgált**
beszél	12 414	657 (5,29%)
beszélt	11 557	729 (6,31%)
mondja	51 142	27 941 (54,63%)
mondta	107 831	80 157 (74,33%)
szól	15 588	2 436 (15,61%)
szólt	18 988	7 043 (37,09%)

A fenti eredményeket a *7. táblázattal* összevetve jól látható, hogy e hat igealak mindegyike nagyobb arányban tölt be idéző funkciót, mint a teljes korpuszban általában. A *mond* szótőből képzett igék szerepelnek leggyakrabban idéző igeként, a *beszél* igéből képzettek pedig a legritkábban.

A szépirodalmi alkorpuszban is megkerestem a beszélőjelölőével megegyező szerkezeteket. Mivel a találati lista a teljes korpuszra vonatkozó kereséshez képest tovább

szűkült, itt is lehetséges volt a kvalitatív vizsgálat, így e lista találatai ugyanolyan pontossággal segíthetnek hozzá a hipotézisem ellenőrzéséhez, mint a *8. táblázat* találatai. A keresés eredményét az alábbi táblázat foglalja össze.

16. táblázat

Az idéző igék a beszélőjelölőével megegyező szerkezetekben a szépirodalmi alkorpuszban

	Főnévvel	**Névelővel**	**Összes**
beszél	2 (2,32%)	0	2 (2,27%)
beszélt	3 (3,47%)	0	3 (3,41%)
mondja	51 (59,30%)	0	51 (57,95%)
mondta	12 (13,95%)	1 (50%)	13 (14,77%)
szól	14 (16,28%)	1 (50%)	15 (17,04%)
szólt	4 (4,65%)	0	4 (4,54%)

A *16*. és a *8. táblázatot* összehasonlítva azt látjuk, hogy a beszélőjelölőével azonos szerkezet[4] néhány ige esetében szinte kizárólag a szépirodalmi alkorpuszban fordul elő, és a *mondja* kivételével minden igére igaz, hogy a szépirodalmi alkorpuszban jóval nagyobb arányban találhatók meg e szerkezetben, mint az a szépirodalmi alkorpusznak az MNSZ teljes méretéhez való arányából következne.

Az MNSZ szövegeiből és annak szépirodalmi alkorpuszából nyert adatok alátámasztották azt a feltételezésemet, miszerint a három ige közül a *mond* szótő alakjai mindkét korpuszban nagyobb számban fordulnak elő a többinél, továbbá rávilágítottak, hogy idéző funkciót is nagyobb arányban töltenek be. A kutatásból az is kiderült, hogy a *mond* ige alakjai a beszélőjelölőével megegyező szerkezetekben is gyakrabban szerepeltek, mint azonos idejű társaik. Az eredmények megerősítettek abban a véleményemben, hogy ha múlt idejű igével akarjuk fordítani, a *mondta* a legjobb választás a beszélőjelölő igéjének megfeleltetésére.

4 A szanszkritban nincs névelő, ezért a magyar mondateleji *névelő-melléknév-főnév-ige* szerkezetet azonosnak tekintem a beszélőjelölő mondat elején található *melléknév-főnév-ige* szerkezetével.

3.3. A vizsgált igék a PC szépirodalmi alkorpuszában

Az MNSZ lehetővé teszi a korpusz online elemzését. Mivel a PC esetében ez a lehetőség egyelőre nem áll rendelkezésre, használatához szoftveres segítségre van szükség. Kutatásomhoz elég volt egy szövegszerkesztő, a Notepad++ 7.5.9. verziójának[5] használata. Ez a program lehetővé teszi az összes megnyitott dokumentumban történő egyidejű keresést, és a lekérdezés egy adott szóalakra is lehetséges benne. A PC szövegei közül kiválasztott művek kutatásához mindez elegendő is volt. A PC még fejlesztés alatt álló szépirodalmi alkorpuszában jelenleg az angolról magyarra fordított regények között találunk megfelelő mennyiségű elemzésre alkalmas szöveget. Adatbázisának áttekintésekor csak egyetlen olyan magyar nyelven íródott regény szerepelt benne, amely 2000 után keletkezett, így ennek alapján nem lehetett messzemenő és általánosítható következtetéseket levonni. A vizsgálatra kiválasztott angol nyelvű regények összterjedelme 612 784 szó. Magyar fordításuk 535 203 szavas korpuszt alkot. A *17. táblázat* mutatja az igék célnyelvi szövegbeli jelen és múlt idejű alakjainak előfordulását, ahogyan azok a beszélőjelölőben szerepelhetnének (a zárójelben álló adatok az egyes igéknek az igék összegéhez viszonyított arányát mutatják).

17. táblázat

A vizsgált igék száma a PC szépirodalmi alkorpuszában

beszél	60 (1,97%)
beszélt	207 (6,78%)
mondja	126 (4,13%)
mondta	2424 (79,45%)
szól	45 (1,47%)
szólt	189 (6,19%)

5 https://notepad-plus-plus.org

A fenti adatokat a *10. táblázattal* összehasonlítva kitűnik, hogy a *beszélt* és a *szólt* kivételével a vizsgált igék egymáshoz viszonyított aránya jelentősen különbözik. Különösen feltűnő a *mondta* gyakoriságának csaknem 30 százalékos növekedése, és a *beszélt* igealakén kívül az összes többi, különösen a *mondja* számarányának csökkenése.

A *17. táblázat* szavai az igék nem idéző funkciójú és elváló igekötős alakjait is magukban foglalják. Ebben a korpuszban is megpróbáltam pontosabbá tenni a találatokat az MNSZ kutatásának leírásában már bemutatott módszerrel. A választott könyvek alkotta korpusz mérete azonban jelen esetben már lehetővé tette a kvalitatív elemzést. Ebben a korpuszban felkiáltójel után álló idéző igék nem fordultak elő. A következő, az igék vessző és gondolatjel utáni, illetve kettőspont előtti pozícióját bemutató táblázat számai már csupán az igekötő nélküli, idéző funkciójú alakokat tartalmazzák.

18. táblázat

A vessző és gondolatjel utáni, valamint kettőspont előtti igék a PC szépirodalmi alkorpuszában

Igék	**Vessző utáni**	**Gondolatjel utáni**	**Kettőspont előtti**
beszél	0	0	0
beszélt	1 (0,92%)	0	2 (1,77%)
mondja	9 (8,26%)	22 (1,05%)	6 (5,31%)
mondta	99 (90,82%)	2067 (98,24%)	93 (82,30%)
szól	0	0	0
szólt	0	15 (0,71%)	12 (10,62%)

A táblázat tanúsága szerint a PC szépirodalmi alkorpuszában a vizsgált igék között a *mondta* alakjai uralják az idéző funkciót, és a jelen idejű idézésre alig találunk példát. Az MNSZ teljes szövegállományában és szépirodalmi alkorpuszában egészen más volt a jelen és múlt idejű idézés előfordulási aránya (*5.* és *13. táblázat*).

Ahogyan az MNSZ vizsgálatánál is megtettem, most bemutatom a három tanulmányozott idéző pozícióban álló ige teljes referenciakorpuszhoz viszonyított arányát, a vizsgálat eredményeit az alábbi táblázat tartalmazza.

19. táblázat

Az igék három írásjellel együtt vizsgált alakjának aránya a PC szépirodalmi szövegeiben a teljes referenciakorpuszbeli előfordulásukhoz képest

	Teljes korpusz	**Írásjelekkel vizsgált**
beszél	60	0
beszélt	207	3 (1,45%)
mondja	126	37 (29,36%)
mondta	2424	2259 (93,19%)
szól	45	0
szólt	189	27 (14,28%)

A táblázat alapján megállapíthatjuk, hogy a *mondja* és a *szólt* teljes referenciakorpuszhoz viszonyított arányszáma az MNSZ szépirodalmi alkorpuszából nyert adatokkal összehasonlítva alacsonyabb igazán jelentősen (*15. táblázat*), az MNSZ teljes szövegállományából nyert adatokhoz képest jóval kisebb a csökkenés (*7. táblázat*). A másik két korpusz adataihoz viszonyítva a *mondta* szerepének változása is feltűnő, a PC szövegeiben szinte csak idéző funkciót tölt be. A kutatás tehát beigazolta azt a feltételezésemet, hogy a PC szépirodalmi alkorpuszában a nyomatékosító betoldás nélkül szereplő *mondta* a betoldást szintén nélkülöző *szólt* igéhez képest is nagyobb arányban fordul elő az MNSZ szövegállományában fellelt arányoknál.

A PC szépirodalmi alkorpuszában egyetlen olyan példát találtam, ahol az ige a beszélőjelölőével azonos szerkezetben volt megtalálható: *A kapitány beszélt: […]* Ehhez a korpusz kettősponttal követett találatai közül egy *Ő mondta: […]* kifejezés áll a legközelebb, de mivel a beszélőjelölőben nem névmás, hanem főnév áll az ige előtt, az ilyen esetek kívül estek a kutatásom körén. Az MNSZ hatalmas adatbázisában is alig fordult elő a beszélőjelölőével azonos igei szerkezet, ezért meglepett, hogy a PC szépirodalmi alkorpuszában egyáltalán előfordult valamilyen találat. Ebből az egyetlen mondatból pedig óvakodnék bármilyen következtetést is levonni.

4. A *said* magyar fordítása

Ahogyan azt a tanulmány első fejezetében említettem, a beszélőjelölő igéjének hagyományos angol fordítása a *said* igealak. Több angolból készült magyar *Bhagavad-gītā*-fordítás is létezik (Gömöryné 1925; Szabó 1941; Sebestyén 2000; Śrīdhara 2003; Maharishi 2016; Bhaktivedanta é. n.; Sivananda é. n.), és ezek a fordítások, amikor nem konkretizálják a beszélőjelölő igéjét, szinte mindig a *szól* valamely alakjával adják vissza a *said* jelentését. A PC szövegei alapján arra is szerettem volna választ kapni, hogy a szépirodalmi kiadványokban mely magyar igével fordítják leggyakrabban a *said* igét. Először megnéztem a kutatásban szereplő igék előfordulását a szépirodalmi alkorpuszban, a vizsgálat eredményét a következő táblázat tartalmazza.

20. táblázat

A kutatott múlt idejű igék mennyisége a PC szépirodalmi alkorpuszában

said	3 838
beszélt	207
mondta	2 424
szólt	189

Ezek a számok, bár azt jelzik, hogy a három vizsgált magyar ige közül valószínűleg a *mondta* áll leggyakrabban a forrásnyelvi *said* helyén, semmiképpen sem bizonyító erejűek, hiszen a magyar szavak más angol kifejezések ekvivalensei is lehetnek. A találatokban szereplő igék nem feltétlenül töltenek be mind idéző funkciót, és a fenti keresés az elváló igekötős, valamint a határozószó beszúrásával kapott alakokat sem különítette el. Hogy bizonyosságot nyerjek, két további vizsgálatot végeztem. Az elsőben összehasonlítottam a *said* és a *mondta* arányát a forrás- és a célnyelvi szövegekben. Ennek eredményét a *21. táblázat* első három oszlopa tartalmazza. A második keresés alkalmával kiválasztottam a kutatásban szereplő művek első és utolsó tíz, idéző funkciójú *said* igét tartalmazó mondatát, és megnéztem, mi szerepel az idéző ige magyar fordításában. A *21. táblázat* 4.

oszlopa azt mutatja meg, hogy hány alkalommal áll *mond* tőből képzett ige a kiválasztott húsz angol mondat magyar megfelelőjében. Ennek az oszlopnak a számai nem tartalmazzák az igekötős és a határozószó betoldásával keletkezett alakokat.

21. táblázat

A said és a mondta aránya a forrás- és a célnyelvi szövegekben

Szerző	*said*	*mondta*	*mond*/20
Addison	79	288 (364, 56%)	10 (50%)
Tolkien	584	419 (71,75%)	14 (70%)
Birch	1 160	473 (40,77%)	8 (40%)
Jacobson	37	502 (93,48%)	17 (85%)
Zusak	347	377 (108,64%)	9 (45%)
Gaiman	546	182 (33,33%)	2 (10%)
Ishiguro	585	183 (31,28%)	9 (45%)

A *22. táblázat* a *said* igének a 140 kiválasztott mondatban található magyar megfelelőjét tartalmazza ábécé és gyakorisági sorrendben.

22. táblázat

A said első és utolsó tíz magyar megfelelője a PC kiválasztott szövegeiben

Ekvivalensek (ábécé)	**Mennyiség**	**Ekvivalensek (gyakoriság)**	**Mennyiség**
1. [kihagyás]	15	1. *mondta*	63
2. állapította meg	1	2. *azt mondta*	16
3. állította	1	3. [kihagyás]	15
4. *azt kérdeztem*	1	4. *mondtam*	5
5. *azt mondta*	16	5. *felelte*	3
6. *azt mondták*	1	6. *feleltem*	3
7. *felelte*	3	7. *jegyezte meg*	3
8. *feleltem*	3	8. *kérdezte*	3

9. *fogalmazott*	1	9. *válaszolta*	3
10. *folytatta*	1	10. *tette hozzá*	2
11. így *folytattam*	1	11. állapította *meg*	1
12. *füstölgött*	1	12. állította	1
13. *híresztelték*	1	13. *azt kérdeztem*	1
14. így *szólt*	1	14. *azt mondták*	1
15. *ijedt meg*	1	15. *fogalmazott*	1
16. *intette*	1	16. *folytatta*	1
17. *jegyezte meg*	3	17. így *folytattam*	1
18. *jelentette ki*	1	18. *füstölgött*	1
19. *kérdezte*	3	19. *híresztelték*	1
20. *kérdeztem*	1	20. így *szólt*	1
21. *magyarázta*	1	21. *ijedt meg*	1
22. *megismételte*	1	22. *intette*	1
23. *megjegyezte*	1	23. *jelentette ki*	1
24. *merengtem*	1	24. *kérdeztem*	1
25. *mondogatták*	1	25. *magyarázta*	1
26. *mondta*	63	26. *megismételte*	1
27. *mondtam*	5	27. *megjegyezte*	1
28. *szerette hívni*	1	28. *merengtem*	1
29. *szólaltam meg*	1	29. *mondogatták*	1
30. örvendezett	1	30. *szerette hívni*	1
31. *tette hozzá*	2	31. *szólaltam meg*	1
32. üdvözölte	1	32. örvendezett	1
33. üdvözölték	1	33. üdvözölte	1
34. *válaszolta*	3	34. üdvözölték	1

A *22. táblázat* alátámasztja a *mond* elterjedtségét megbecslő hipotézisemet. A tanulmányom tárgyát képező hat igealak közül a *mondta* mellett csak a *szólt* található meg a *said* megfelelői között, az is csak betoldás kíséretében, és más olyan ige sem szerepel a táblá-

zatban, amely megközelítené a *mondta* és a többi *mond* alakból képzett szó gyakoriságát. A tanulmányozott szövegekben a *mond* alakjaival való behelyettesítést követően a *kihagyás* a leggyakrabban használt átváltási művelet.

5. A kutatás eredményének összefoglalása

Tanulmányomban a beszélőjelölőben található szanszkrit *uvāca* szó magyarra való fordításának lehetőségeit elemeztem. Szóltam az ige kihagyásának és az igeidő megváltoztatásának lehetőségéről, majd az MNSZ és a PC segítségével azt kutattam, hogy egy modern kiadásban mely magyar lexikai elem lenne a legmegfelelőbb az *uvāca* jelentésének a visszaadására, különös tekintettel arra, hogy számos magyar *Bhagavad-gītā*-fordítás nem a szanszkrit eredetiből, hanem annak angol változatából készült, és ezekben a szanszkrit igét általában a *said* igealakkal fordítják.

Megvizsgáltam, hogy a szanszkrit ige Körtvélyesi által megadott és a magyar kiadásokban fellelhető konkretizálatlan fordításai milyen arányban találhatók meg a korpuszokban, és a fő célom azon feltételezésem alátámasztása volt, hogy egy modern magyar nyelven megfogalmazott kiadásban – amennyiben mindenképpen ragaszkodunk a szerepeltetéséhez – indokolt lenne szakítani az ige átültetésének magyar fordítói hagyományával: ha múlt időben fordítanánk le, a *szólt* igealaknál megfelelőbb lenne a *mondta* használata. Mivel a beszélőjelölő esetében az igeidő megváltoztatását jó fordítói eszköznek tartottam, kutatásomba a jelen idejű alakokat is belevettem. Munkám során a következő eredményre jutottam:

Ahogyan azt feltételeztem, a *mond* a *szól* igénél sokkal gyakoribb, minden táblázatból az derül ki, hogy a tanulmányban szereplő alakjai jóval többször fordulnak elő, mint a *szól* megvizsgált alakjai. A két szépirodalmi korpuszt összevetve azt látjuk, hogy a 2000 utáni szövegekben a vizsgált idéző igék között a *mondta* előfordulásának aránya nőtt, a többié pedig csökkent (*13*. és *18*. *táblázat*). Érdekes, nem várt eredménye a kutatásnak, hogy a beszélőjelölőével azonos szerkezetekben a *mond* és a *szól* szótőből képzett jelen idejű alakok száma jóval nagyobbnak bizonyult a múlt idejűeknél (*8*. és *16. táblázat*). Mindez megerősíti azt a vélekedést, hogy az ige múlt idejű alakjának használata esetén a

mondta lenne a legjobb választás, és meggyőző bizonyítékul szolgált, hogy a jelen idejű fordításra a *mondja* igealak használata javasolható leginkább. A legáltalánosabb szemantikai jelentésű idéző ige alkalmazása azzal is segítheti (a jelölő funkcióját beteljesítve) a beszélő személyére terelni a figyelmet, hogy a maga hétköznapiságával a háttérben marad.

A PC vizsgálata egy másik nem várt eredménnyel is szolgált: a *said* fordításakor a *mondta* és az *azt mondta* használata után az ige kihagyása volt a fordítók által leggyakrabban alkalmazott átváltási művelet. Ez azt mutatja, hogy a kihagyás egyes a beszélőjelölőétől eltérő szerkezetekben is célszerű megoldás lehet. Noha a kutatási kérdésem nem erre irányult, már az MNSZ vizsgálatának eredménye is megerősítette azt a véleményemet, hogy a beszélőjelölő fordításakor a legjobb, ha kihagyjuk az igét. A *Bhagavad-gītā* beszélőjelölőjének alakjai egy hosszú párbeszéd kérdései és válaszai előtt állnak. Bár a jelölőével azonos szerkezet – ha csekély számban is – megtalálható a vizsgált korpuszban, általában csak egy-egy mondatot vezet be, nem szerepel hosszú párbeszédekben. Bár a drámairodalomból jól ismert a beszélő személyének a hosszú párbeszédek során a mondandó előtti feltüntetése, ilyenkor a név mindig ige nélkül, önmagában áll. Elmondhatjuk, hogy ezeknél a műveknél ez az eljárás normának tekinthető, így véleményem szerint jó volna a *Bhagavad-gītā* beszélőjelölőjében az ige elhagyását fordítási normaként elfogadni.

Az ige kihagyásakor érdekes jelenség tanúi lehetünk. Noha a kihagyást az implicitáció körébe sorolják (Klaudy 2004), a beszélőjelölő esetében az ige kihagyása explicitebbé teszi a jelölő funkcióját, így rávilágít arra, hogy nem tartozik szorosan a szöveghez, technikai segédlet csupán. A jelölőre eredetileg azért volt szükség, mert a szanszkrit kéziratokban a helytakarékosság szempontjait szem előtt tartva mindent folyamatosan jegyeztek le, a párbeszédek követhetősége érdekében még akkor is ajánlatos volt a beszélő személyének a szavai előtti megjelenítése, ha a szereplőváltás kiolvasható a szövegből. A gondolatjel kitételének és az új sorba tördelt párbeszédelemeknek a hazai gyakorlata sokszor mind a személynév, mind az idéző ige használatát feleslegessé teszi.

A jövőben érdekes lenne megvizsgálni, hogy a 2000 után keletkezett autentikus magyar szövegekben is hasonló változást láthatunk-e a *szólt* és a *mondta* előfordulásának

arányában, mint az MNSZ és a PC-ből nyert fordítási korpusz összevetésével. Azt is jó lenne látni, hogy egy 20. századi angol–magyar irányú szövegeket tartalmazó fordítási korpuszban változik-e az MNSZ találataihoz képest a két ige előfordulásának aránya.

Irodalom

Ardekani, M. A. M. 2002. The Translation of Reporting Verbs in English and Persian. *Babel* Vol. 48. No. 2. 125–134. https://doi.org/10.1075/babel.48.2.03ard

Klaudy K. 1999. *Bevezetés a fordítás gyakorlatába.* Budapest: Scholastica.

Klaudy K. 2004. Az implicitációról. In: Navracsics J., Tóth Sz. (szerk.) *Nyelvészet és interdiszciplinaritás. Köszöntőkönyv Lengyel Zsolt 60. születésnapjára.* Szeged: Generália. 70–75.

Klaudy K. 2007. Az idéző mondategység igéiről. In: Klaudy K. *Nyelv és fordítás.* Budapest: Tinta Könyvkiadó. 57–68.

Klaudy K., Simigné F. S. 2000. *Angol–magyar fordítástechnika.* Budapest: Nemzeti Tankönyvkiadó.

Klaudy, K., Károly, K. 2005. Implicitation in Translation: Empirical Evidence for Operational Asymmetry in Translation. *Across Languages and Cultures* Vol. 6. No. 1. 13–29.

Körtvélyesi T. 2006. *Szanszkrit nyelvtan.* Budapest: A Tan Kapuja Buddhista Főiskola.

Kulkarni, A. 2016. Sanskrit. In: Hock, H. H., Bashir E. (eds) *The Languages and Linguistics of South Asia: A Comprehensive Guide.* Berlin: De Gruyter Mouton. 744–747.

Monier-Williams, M. 2011. *A Sanskrit-English Dictionary.* Delhi: Motilal Banarsidass Publishers.

Országh L., Magay T. 2009. *Angol–magyar nagyszótár*. Budapest: Akadémiai Kiadó.

Pusztai F. (főszerk.) 2011. *Magyar értelmező kéziszótár*. Budapest: Akadémiai Könyvkiadó.

Robin E., Dankó Sz., Götz A., Nagy A. L., Pataky É., Szegh H., Török G., Zolczer P. 2016. Fordítástudomány és korpuszkutatás: bemutatkozik a Pannónia Korpusz. *Fordítástudomány* 18. évf. 2. szám. 5–26.

Seidl-Péch O. 2018. Melyek a (szak)fordító és a fordításkutató munkáját segítő legfontosabb nyelvi korpuszok? In: Robin E., Zachar V. (szerk.) *Fordítástudomány ma és holnap.* Budapest: L'Harmattan Kiadó. 175–191.

Yeganeh, M. T., Boghayeri M. 2015. *The Frequency and Function of Reporting Verbs in Research Articles Written by Native Persian and English Speakers.* In: Rahimi, A. (ed.) *The Proceedings of 2nd Global Conference on Conference on Linguistics and Foreign Language Teaching. Procedia – Social and Behavioral Sciences* Vol. 192. 582–586. https://doi.org/10.1016/j.sbspro.2015.06.097

Yu H., Guo S. 2016. The Master *said*, the Master *exclaimed*: reporting verbs and the image of Huineng in translations of the *Platform Sutra*. *Asia Pacific Translation and Intercultural Studies* Vol. 3. No. 3. https://doi.org/10.1080/23306343.2016.1228148

Várdai T., Oravecz Cs. 2014. A Magyar Nemzeti Szövegtár egymilliárd szavas új változata. *Magyar Tudomány* 175. évf. 9. szám. 1054–1061.

Források

Abhinavagupta. 2012. *Gītārthasaṁgraha.* Budapest: Perisca Kiadó.

Addison, C. 2013. *A napfény gyermekei.* Szeged: Könyvmolyképző.

Addison, C. 2012. *A Walk Across the Sun.* SilverOak.

Bakaja Z. 2019. Átváltási műveletek a beszélőjelölőnek nevezett szanszkrit idéző mondat angol és magyar nyelvű fordításaiban. *Argumentum* 15. 755–783 https://doi.org/10.34103/argumentum/2019/4

Bakaja Z. (ford.) 2016. *Bhagavad-gítá részlet*. In: Hetényi Zsuzsa (szerk.) *5Pofon: Antológia műfordítás-antológia*. Budapest: ELTE BTK. 8–22.

Bakos A., Bakos J. E. (ford.) 2017. *Bhagavad-gītā.* Budapest: Danvantara Kiadó.

Baktay E. 2013. *A Magasztos szózata.* Budapest: Filosz Kiadó.

Baladeva, V. é. n. *Gītā Bhūṣaṇa.* Chennai: Sri Vaikunta Enterprises.

Besant A., Bhagavân D. (ford.) 1905. *The Bhagavad-gītā.* London, Benares: Theosophical Publishing Society.

Bhaktivedanta Swami Prabhupāda, A. C. 1972. *Bhagavad-gītā As It Is.* New York: Collier Books, London: Collier Macmillan Publishers.

Bhaktivedanta Swami Prabhupāda, A. C. 2001. *A Bhagavad-gītā úgy, ahogy van.* h. n.: The Bhaktivedanta Book Trust.

Bhaktivedanta Swami Prabhupāda, A. C. 2014. *A Bhagavad-gītā úgy, ahogy van.* h. n.: The Bhaktivedanta Book Trust.

Bhaktivedanta Swami Prabhupāda, A. C. 2015. *Bhagavad-gītā As It Is.* h. n.: The Bhaktivedanta Book Trust.

Bhaktivedanta Swami Prabhupāda, A. C. é. n. *Az eredeti Bhagavad-gītā.* Vaduz: The Bhaktivedanta Book Trust.

Bhaktivedanta Swami, A. C. 1968. *The Bhagavad Gita As It Is*. London: Collier – Macmillan Ltd.

Bhaktivinoda 2006. *Srīmad Bhagavad-gītā*. Vrindaban: Rasbihari Lal & Sons.

Birch, C. 2013. *Jamrach vadállatai*. Budapest: Gondolat.

Birch, C. 2011. *Jamrach's Menagerie*. Edinburgh: Canongate.

Gaiman, N. 2013. *The Ocean at the End of the Lane*. New York: William Morrow.

Gaiman, N. 2014. *Óceán az út végén.* Budapest: Agave Könyvek.

Gömöryné Maróthy M. (ford.) 1924. *Bhagavad-Gitá az isteni ének.* Budapest: Légrády Nyomda és Könyvkiadó.

Ishiguro, K. 2006. *Ne engedj el*. Budapest: Palatinus.

Ishiguro, K. 2005. *Never Let me Go*. New York: Alfred A. Knopf.

Karma T. 2003. *A Bhagavad Gítá „A magasztos szózata" (ahogyan én látom)*. Budapest: Berkes Attila magánkiadása.

Kégl S. 1910. *Bhagavadgîtâ.* Budapest: Magyar Tudományos Akadémia.

Lakatos I., Vekerdi J. (ford.) 1987. *A magasztos szózata Bhagavad-gítá*. Budapest: Európa Kiadó.

Macnicol, N. (ed.) 1938. *Hindu Scriptures.* London: Dent. New York: Dutton.

Maharishi, M. 1967. *Bhagavad-Gita: A New Translation and Commentary Chapters 1–6.* London: International SRM Publications.

Maharishi, M. 2016. *A Bhagavad-gítá: Új formátumban és új magyarázattal.* h. n.: Vaszistha Kft.

Nārāyaṇa Goswāmī, Bh. 2011. *Śrīmad Bhagavad-gītā.* Vṛndāvana, New Delhi, San Francisco: Gaudiya Vedanta Publications.

Sebestyén E. (ford.) 2000. *Bhagavad Gita*. Budapest: Sri Sathya Sai Baba Szervezet Magyarországi Központja.

Sivananda *Bhagavad Gítá.* é. n. Budapest: Spirituart Jóga Alapítvány - Sivananda Jógaközpont.

Sridhar Dev-Goswami, Bh. R. 2006. *Śrīmad Bhagavd-gītā: The Hidden Treasure of the Sweet Absolute.* Nabadwip: Sri Chaitanya Saraswat Math.

Śrīdhara Deva Gosvāmī, B. R. 2003. *Bhagavad-Gītā: Az Édes Abszolút rejtett kincse.* h. n.: Harmónia Alapítvány.

Szabó L. (ford.) 1941. A Bhagavadgitá-ból: A Mindenség látomása. In: *Örök barátaink: Szabó Lőrinc kisebb műfordításai*. 348–354. Budapest: Singer és Wolfner.

Szerdahelyi I., Tóth E. (ford.) 1965. *Mahábhárata.* Budapest: Európa Kiadó.

Theodor, I. 2010. *Exploring the Bhagavad Gītā Philosophy, Structure and Meaning.* Farnham: Ashgate Publishing Limited, Burlington: Ashgate Publishing Company.

Tolkien, J. R. R. 2001. *A Hobbit*. Budapest: Európa Kiadó.

Tolkien, J. R. R. 2006. (1937) *The Hobbit*. New York: Harper Collins.

Vekerdi J. (ford.) 1997. *A Magasztos szózata.* Budapest: Terebess Kiadó.

Viśvanātha Cakravartī 2003. *Sārārtha-Varṣiṇī-Tīkā.* Chennai: Sri Vaikunta Enterprises.

Wilkins C. 1785. *Bhagvat-geeta or Dialogues of Kreeshna and Arjoon; in eighteen lectures; with notes.* London: C. Nourse.

Zusak, M. 2015. *A könyvtolvaj*. Budapest: Európa.

Zusak, M. 2005. *The Book Thief*. New York: Alfred A. Knopf.

Internetes hivatkozások

http://clara.nytud.hu/mnsz2-dev/

https://notepad-plus-plus.org

A vonatkozó névmások használata a feliratokban: TED-előadások magyar nyelvű összehasonlítható korpuszának vizsgálata

Malaczkov Szilvia

Malaczkov.Szilvia@uni-bge.hu

Budapesti Gazdasági Egyetem, Külkereskedelmi Kar

Nemzetközi Üzleti Szaknyelvek Tanszék

Kivonat: A magyar nyelvre történő fordítás oktatása során érdemes tudatosítani a fordítóhallgatókkal a különféle nyelvváltozatok megfelelő használatát. A jelen tanulmány célja egy jelenleg is formálódó nyelvváltozat, az írott beszéltnyelviség vagy digilekt fordítástudományi megközelítésű bemutatása. Ez a nyelvváltozat eltér a sztenderd írott és beszélt nyelvi változattól oly módon, hogy mindkettő jellegzetességeit magában hordozza. A kutatás célja az önkéntes TED-fordítók online feliratiban megjelenő nyelvhasználat összehasonlítása az autentikus magyar nyelvű TED-előadások nyelvhasználatával, valamint a sztenderd nyelvhasználattal. A kutatás azt vizsgálja, hogy (1) a vonatkozó mellékmondatok ugyanolyan gyakorisággal fordulnak-e elő az autentikus magyar és a fordított magyar szövegekben, (2) melyek a leggyakrabban használt vonatkozó névmásos kifejezések, és (3) vajon a sztenderd nyelvhasználatot alapul véve eltérő vagy hasonló a vonatkozó névmásos kifejezések használata. A kutatás megállapította, hogy (1) a magyarra fordított szövegek több mellékmondatot tartalmaznak, mint az autentikus magyar nyelvű szövegek, (2) a leggyakrabban használt vonatkozó névmás mindkét korpuszban az *ami(t)* és (3) a sztenderd nyelvhasználattól mindkét szövegtípus eltért.

Kulcsszavak: audiovizuális fordítás, önkéntes fordítók, Pannónia Korpusz, vonatkozó mellékmondat, írott beszéltnyelviség

Malaczkov Szilvia: A vonatkozó mellékmondat használata a feliratokban: TED-előadások magyar nyelvű összehasonlítható korpuszának vizsgálata. In: Robin E., Seidl-Péch O. (szerk.) 2020. *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordításkutatás Magyarországon.* Segédkönyvek a nyelvi közvetítésről I. Budapest: ELTE BTK Fordítástudományi Doktori Program, MANYE Fordítástudományi Szakosztály. DOI: https://doi.org/10.36252/Nyelvikozvsegedkonyv1.10

1. Bevezetés

Az anyanyelvre történő fordítás oktatása során érdemes tudatosítani a fordítóhallgatókkal a nyelvben egyszerre jelenlévő, különféle nyelvváltozatok megfelelő használatát. A nyelvváltozatok megismertetése és későbbi tudatos használata elősegítheti olyan fordított szövegek létrehozását, amelyek az adott szövegtípushoz megfelelően alkalmazkodnak, és elkerülik a fordított szövegekre jellemző nemkívánatos sajátosságokat. Jelen kutatásom célja egy fordított szövegekben (is) megjelenő magyar nyelvváltozat kvantitatív vizsgálata.

Kutatásom első részében a vonatkozó mellékmondatok gyakoriságát vizsgálom autentikus magyar nyelvű TED-előadások átirataiban és önkéntes fordítók által létrehozott magyar nyelvű TED-feliratokban. Hipotézisem az, hogy a fordított szöveg több vonatkozó mellékmondatot tartalmaz, mivel a fordítók az egyértelműsítés elve alapján jobban kifejtik a forrásnyelvi szöveg tartalmát. Azt is megvizsgálom, mely vonatkozó névmásos kifejezések szerepelnek leggyakrabban az autentikus magyar nyelvű, illetve a magyarra fordított szövegekben. Hipotézisem szerint a két alkorpusz szövegei ugyanazokat a vonatkozói névmásos kifejezéseket tartalmazzák, mivel a célnyelvnek megfelelő szövegszerkesztés elve ezt feltételezi. Kutatásom második részében a főnévi vonatkozó névmásos kifejezések használatát vizsgálom. Ezek használata eltérő lehet attól függően, hogy sztenderd vagy nem sztenderd nyelvváltozatot, illetve beszélt vagy írott nyelvet vizsgálunk. Hipotézisem szerint a vizsgált audiovizuális korpuszban a nem sztenderd használat lesz jellemző, mivel szóbeli – azaz az írott, sztenderd változattól eltérő – jegyeket hordoz a fordított felirat.

A jelen tanulmány szempontjából fontos áttekinteni a sztenderd nyelvhasználat, a nyelvváltozatok és a nyelvi norma közötti kapcsolatot, illetve a beszélt és az írott nyelv közötti dichotómiát. A cikk második és harmadik részében ezt a két témakört tekintem át nagy vonalakban az eddigi kutatások tükrében. Mindkét témának óriási szakirodalma van, ezért csak a jelen tanulmány szempontjából fontos átfogó kép kialakítására hozok fel példákat a kutatási terület eredményeiből. A cikk negyedik és ötödik része fordítástudományi megközelítésből mutatja be a vonatkozó mellékmondatok és a főnévi vonatkozó

névmásos kifejezések kutatásának jelentőségét a magyar nyelvben. A cikk hatodik részében magát a kutatást mutatom be, azaz a kutatott korpuszt, a kutatási módszert, a kutatási kérdéseket és hipotéziseket. A cikk a kutatás eredményeinek részletes ismertetésével és a vizsgált terület jövőbeni kutatási irányainak felvázolásával zárul.

2. Nyelvváltozatok, sztenderd nyelvhasználat és nyelvi norma

A leíró nyelvészek és a bizonyos mértékig előíró szemléletet képviselő nyelvművelők máshogy vélekednek az adott nyelvben egyszerre létező különböző nyelvváltozatok hierarchiájáról és értékéről. Vitathatatlan tény, hogy a nyelv változik, és a digitális forradalom korszakában talán még gyorsabban is, mint korábban. A különböző nyelvi változatok – a hagyományos és az újonnan megjelenő formák – sokáig együtt élnek a nyelvben, és ezeket „nyelvi változóknak" (Schirm 2005: 85; Kiss 1996: 62; Labov 1972) hívjuk. A nyelvet közösségek hozzák létre, és az adott közösség kialakítja saját maga számára az általánosan elfogadott nyelvhasználati jellegzetességeket. Aki ezektől a jellegzetes nyelvhasználati formáktól eltér, magára vonja a közösség valamilyen irányú értékítéletét.

Habár egy adott nyelven belül a közösségek létrehozzák a saját nyelvváltozatukat, mindig létezik egy sztenderd nyelvváltozat is, amely a magyar kultúrában a művelt magyar beszélők főként írott szövegekben használt dialektusa, a köznyelv. Ezt a változatot tanítják a magyart mint idegen nyelvet tanulóknak, és ezt a változatot kodifikálják a nyelvtankönyvek normaként (Kiefer 2006: 707). Ezt a sztenderd nyelvváltozatot a leíró nyelvészek ugyanolyan értékűnek tekintik, mint a többi nyelvváltozatot, és nem tekintik ideális nyelvi formának. Kiefer (2006: 709) szerint azonban sokan megijednek a nyelvi változásoktól, és a sztenderdet, azaz a művelt írásbeli nyelvhasználatot tekintik „normának" az európai kultúrában, a nyelv változásában pedig a nyelv romlását látják. Az újítás erejével szemben a hagyománytiszteletet előtérbe állítva, a laikus nyelvművelők a „nyelvhelyességi elveket" éppen ezért sokszor „nyelvhelyességi babonává" változtatják (Szepesy 1986).

A nyelvészek jelenlegi álláspontja szerint a „nyelvi norma” kérdése nem tartozik a nyelvészetre mint tudományterületre. Nádasdy Ádám (2010) *Búcsú a nyelvhelyességtől* című cikkében a következőképpen fogalmazza meg ezt az álláspontot:

> […] a nyelvészek érdeklődése nyelvhelyességi kérdésekkel szemben feltűnően csekély. A legtöbb nyelvész, ha nyelvhelyességi kérdésekben döntőbírónak felhívják, ha őszinte merne lenni, azt felelné: „Én a nyelv történeti fejlődését vizsgálom, törvényeket állapítok meg benne, de bírája nem akarok lenni”. (Nádasdy 2010: 1)

A nyelv „helyes” használatával foglalkozó nyelvművelők – akik között nyelvészeket is találhatunk – az általuk értéket képviselő és védelmet igénylő sztenderd nyelvváltozat fenntartását tűzik ki célul, mivel ez a nyelvváltozat az egész nemzetet mint közösséget egyesíti.

Molnár Mária (2014) tanulmánya a nyelvi norma megítélésének diakrón változását mutatja be. Kutatása alapján a magyar nyelvtudományban megjelenő első tudományos igényű normaleírások (Brassai Sámuel, Simonyi Zsigmond) deskriptív jellegűek voltak, és figyelembe vették a szociális, pragmatikai és kommunikációs tényezőket. Azonban a két világháború között (Gombocz Zoltán normaértelmezését leszámítva) a társadalmi nyelvművelés a morál, az esztétikum és a műveltség kategóriáira támaszkodva valójában előíró normafogalmat vezetett be. Abban a korban „a nyelvi rendszer szabályaihoz való legteljesebb mértékű alkalmazkodás megkövetelése az anyanyelv és a hozzá szorosan kapcsolódó nemzet védelmének kifejezőjévé vált” (Molnár 2014: 52). A szocialista nyelvművelés ugyancsak központi normaként írta elő a társadalmi egységet ilyen módon is szimbolizáló igényes köznyelv használatát. A rendszerváltás után visszatért a nyelvészeti köztudatba a norma szociolingvisztikai-pragmatikai értelmezése, és lassan elfogadottá válik a normapluralitás felőli fogalmi megközelítés, miszerint „a különböző nyelvváltozatot beszélő közösségek a magyar nyelvközösségen belül saját, általuk kialakított normához mérten

használják a nyelvet" (Molnár 2014: 52). Az élő nyelvet a közösségek formálják, és állandó dinamizmusban tartják, ami természetes módon a normák sokszínűségét idézi elő.

A fordítástudomány azonban, az idegennyelv oktatásához hasonlóan, igényelheti a nyelvhelyességi norma előíró tulajdonságát. Igény van a nyelvtanulók és a fordítóhallgatók részéről arra, hogy iránymutatást kapjanak arról, mi a „helyes" és mi a „helytelen" nyelvhasználat. Bartsch (1987: 166–178) és Tolcsvai (1998: 32–3) nyomán a nyelvi norma előíró megközelítését helyezi előtérbe Heltai Pál (2004: 413), akinek definíciója szerint „a norma olyan szokásos nyelvhasználatot jelent, amely normatív ereje révén orientáló mintaként működve előírja, illetve szankcionálja a kívánatos és nemkívánatos nyelvhasználatot". Az adott szövegtípusra leginkább jellemző sajátosságokat megismerve a nyelvet és a fordítás technikáját tanuló hallgatók tudatosan dönthetnek arról, hogy a szokásos nyelvhasználatot az adott szövegtípusban fenntartják, vagy tudatosan – tudományosan megalapozott érvekkel alátámasztva – eltérnek attól. Így például a fordítóhallgatók képzésük során elsajátítják az audiovizuális fordítás jellegzetességeit, többek között a TED-előadások feliratozási normáit, és az adott normától eltérő fordítási megoldásaikat indokolniuk szükséges.

A különféle nyelvváltozatok azonban nemcsak diakrón, azaz történeti változáson mennek át, hanem sajátos nyelvhasználati formák jellemzők a különböző nyelvhasználói közösségekre és nyelvhasználati területekre egy adott korban is. A sztenderd nyelvhasználatot figyelembe véve más-más elvárásokat támaszthat egy adott közösség például a magyar nyelv használatát illetően beszédben, illetve írásban. A *Magyar nyelvhasználati szótár* (Balázs és Zimányi 2007: 30) bejegyzése szerint „beszéd és írás megkülönböztetése a nyelvművelés szempontjából alapvetően fontos, mert jellegéből adódóan más elveket kell érvényesíteni a beszéddel és az írással kapcsolatban". A beszédben, illetve az írásban elfogadott nyelvhasználatot azonban explicit vagy implicit módon – írott szabályok vagy íratlan megállapodások formájában – mindig az adott közösség határozza meg. A közösséghez való tartozás és a nyelvhasználat kapcsolatáról Nádasdy Ádám (2010) a következőket írja:

> A nyelvhelyesség normakövetést, csoporthoz való azonosulást, szerepvállalást jelent — avagy ezek hiányát, tudatos vagy ösztönös megtagadását. Ezek a választások összefüggnek az iskolázottsággal, a személyiséggel, az újhoz és a régihez való viszonyulással, a többségi vagy kisebbségi helyzettel. Nem-nyelvi alapú döntések, melyek a nyelvhasználatban valósulnak meg. Hogy egy adott évtizedben mi a helyes magyarság, azt a közerkölcs szabályozza, tehát konvenció, egyfajta közmegegyezés. (2010: 4)

A nyelvet tehát közösségek hozzák létre, és a közösség által explicit vagy implicit módon elfogadott nyelvhasználati formák, a „helyes" nyelvhasználat betartása biztosítja az adott közösséghez való tartozás érzését. Ez a „helyes" nyelvhasználat eltérhet a sztenderd nyelvváltozattól, így magyar nyelvterületen belül is számos, az adott kisebb közösségre jellemző és a közösség által elfogadott nyelvhasználati forma létezik. A nyelv azonban dinamikusan változik, aminek társadalmi és gazdasági okai is lehetnek. Schirm Anita (2005: 86) szerint egy adott nyelvváltozat fennmaradása függ a presztízstől, azaz a nyelvi értékhierarchiában betöltött helyétől, és egyéb nyelvszociológiai okoktól is (vö. Labov 1972).

A jelen tanulmány célja egy jelenleg is formálódó, új magyar nyelvváltozat fordítástudományi megközelítésű vizsgálata. Ez a nyelvváltozat eltér a sztenderd írott és beszélt nyelvi változattól oly módon, hogy mindkettő jellegzetességeit hordozza. Az önkéntes feliratozó közösség az adott szövegtípuson belül egyértelműen használja ezt a nyelvváltozatot, amelyet (amit?) a nyelvészek és a nyelvművelők közössége talán eltérően ítél meg.

3. A nyelvhasználat változása a digitális korban: írott beszéltnyelviség

Az egyes nyelvváltozatok szinkrón módon is jelen vannak a nyelvben, hiszen az egyes szövegtípusoknál más-más nyelvhasználati elvárásaink vannak. A szövegtípusokat Kiefer (2006: 124) alapján a következő szempontok szerint jellemezhetjük: írott–beszélt, monologikus–dialogikus, mindegyik résztvevő jelen van – csak az egyik résztvevő van

jelen, tervezett–spontán, kifejtő–bennfoglaló, van hagyományozott szerkezete – nincs hagyományozott szerkezete. Az egyik fő megkülönböztető jegy tehát a szóbeli és az írásbeli jellegzetességek elhatárolása, azaz más nyelvhasználati szokások érvényesülnek írott és beszélt nyelvben. A *Magyar Nyelvtan* (Kiefer 2006: 124) a fent említett kategóriákat figyelembe véve tehát két fő szövegtípust különít el: beszéd és írás. Ez a tipológia még nem tartalmazza az akkor már kutatott harmadik lehetséges kategóriát, az írott beszéltnyelviséget.

A digitális korszak, az internet megjelenése egy új eszközt és csatornát iktatott be az emberi kommunikációba. Az internet és más telekommunikációs eszközök által közvetített beszéd és írás befolyásolja a nyelvhasználati szokásokat. Az internet használata új közösségek megjelenését is eredményezte, amelyek a virtuális térben formálódtak, és ezáltal létrehoztak egy új, a virtuális térben elfogadott nyelvhasználati formát.

Bódi Zoltán, az internetes nyelvhasználatról írt általam ismert első magyar nyelvű monográfia szerzője az internetes nyelvhasználatot új nyelvváltozatnak tekinti. Az internetes szövegekre ugyanis a hagyományos, írott szövegekkel ellentétben „a szabályozottság, a normativitás, a szerkesztettség magasabb szintje kevésbé jellemző” (Bódi 2004: 35). Véleménye szerint nem egyértelmű, hogy kommunikációs műfajuk írott vagy beszélt nyelvi kommunikáció-e. Ezért „az interneten jelentkező spontán szövegek köztes kommunikációs műfaját” írott beszélt nyelvnek nevezi. Bódi (Balázs és Bódi 2005: 195) későbbi munkájában az írott beszélt nyelvet a szimbolikus írásbeliség előfeltételének tekinti.

Ez a köztes nyelvhasználati forma kapcsolatban áll a másodlagos írásbeliséggel, amely fogalmat Balázs Géza (Balázs és Bódi 2005) vezette be Walter Ong hármas felosztását továbbgondolva. Az Ong szerinti diakrón felosztásban az elsődleges szóbeliséget és írásbeliséget a másodlagos szóbeliség korszaka követte az élő beszédet közvetítő technikai eszközök (telefon, televízió) megjelenésével. Balázs Géza szerint megjelent a másodlagos írásbeliség is, azaz az informatikai és telekommunikációs eszközökkel közvetített írás. A kettő együtt alkotja az új beszéltnyelviséget, amelyre az átmenetiség jellemző: „beszédközeli (szlenges), képeket is használó írás- és közlésmód, az élőbeszédhez közelítő, de azzal nem azonos beszédszintetizálás” (Balázs és Bódi 2005: 40).

Schirm Anita (2005: 101) a mai magyar nyelv állapotáról készített elemzésében megállapítja, hogy napjainkban új szövegtípusok jelennek meg, mivel „az eddig egymástól mereven elkülönített szóbeli és írott szöveg jegyei […] egyetlen szövegfajtán belül jelennek meg". Az új nyelvi változat, az „írott beszéltnyelviség" kialakulását az új kommunikációs eszközök elterjedésében látja. Meglátása szerint az „írott beszéltnyelvi" szövegek 40 százalékban az írott szöveg sajátosságait, míg 60 százalékban a beszélt nyelv sajátosságait hordozzák. Tehát a hagyományos írott, valamint beszélt nyelvi műfajok mellett megjelent egy harmadik kategória: az írott beszéltnyelviség.

Veszelszki Ágnes (2012) az írásbeliség és a szóbeliség dichotómiájának diakrón áttekintése után a mai helyzet elemzésekor megállapítja, hogy a digitális forradalom nyomán ez a dichotómia fellazult. Veszelszki bevezeti a digilektus fogalmát, amely értelmezésében a számítógép által közvetített írásbeli kommunikáció nyelvhasználati módjának megnevezése. „A digilektus sajátos, más médium által közvetített kommunikációra nem jellemző tulajdonságokkal rendelkező új nyelvváltozat" (2012: 404.). Megállapítja, hogy grammatikai szempontból a digilektus szövegei a beszélt nyelvhez közelítenek. A digilektus kutatásával elsősorban az (inter)netnyelvészet foglalkozik.

Fordítástudományi szempontból érdekes kutatást végzett Assis Rosa (2001), aki a portugál állami, illetve magántulajdonú televízióban megjelenő filmek feliratozását vizsgálta. A feliratozás során médium- (beszéd–írás), csatorna- (vokális–vizuális) és kódváltás (beszélt verbális és nem-verbális nyelv – írott verbális nyelv) is történik. Ebből adódóan a feliratozás egyik fő kihívása a médium által meghatározott regiszter kiválasztása. Egyfelől a forrásnyelvhez hasonló célnyelvi beszélt regiszterhez kell igazodnia, másfelől a célnyelvi írott regiszterhez. Kutatása kimutatta, hogy az állami televízióban a feliratozás az írott regiszter normáit követi, mivel itt a presztízs fenntartása a cél, amelyet a sztenderd, formális írott nyelv képvisel. Assis Rosa (2001: 215) ezt a stratégiát centralizációnak nevezi, azaz a központi normához való igazodásnak. Ha a fordítók nem ezt a normát követik, akkor a „rossz feliratozás" stigmája miatt az állásukat is elveszíthetik. A magántelevízióban viszont megjelenik a beszélt nyelvi regiszter is a feliratozásban, amit Assis Rosa (2001: 219) decentralizációnak nevez, azaz a központi normától való eltérésnek. Ez a

jelenség bizonyíthatja a beszélt nyelv jelenlegi és jövőbeni emelkedő presztízsét, valamint a beszéd és az írás közötti értékhierarchia változását.

A fenti kutatások alapján is látható, hogy napjainkban a beszélt és az írott nyelv közötti éles határ bizonyos szövegtípusok esetén elmosódik, és az írott nyelvi normák javára döntő európai értékhierarchia is változóban van. A nyelvészek már a digitális világban megjelenő új nyelvváltozatokat vizsgálják, de mivel jelenleg is az átalakulás folyamata zajlik, a terminológia még nem konzekvens. A fent említett megnevezések mellett találkozhatunk még a netspeak, a hibrid szöveg, az oraliteralitás, a virtuális vagy digitális textualitás, az írott interaktív regiszter és a virtuális írásbeliség terminussal is.

4. A vonatkozó mellékmondatok és jelentőségük a fordítástudományban

Jelen kutatásomban a vizsgált összehasonlítható korpusz azonos nyelvű és azonos szövegtípusba tartozó feliratokat tartalmaz. Baker (1995) szerint ezek a korpuszok arra szolgálnak, hogy a fordított szövegre jellemző sajátosságokat statisztikai úton kimutathassuk. Kutatásomban tehát autentikus célnyelvi szövegeket vetettem egybe fordított célnyelvi szövegekkel, hogy kimutassam a vonatkozó mellékmondatok előfordulását illető esetleges nyelvhasználati eltéréseket vagy éppen hasonlóságokat.

Korábban hasonló kutatást végzett Klaudy Kinga (2017), aki a fordított szövegekre jellemző szövegsajátosságokat vizsgálta fordítástudományi összehasonlítható korpuszon. Eredményei azt mutatták, hogy többek között az alá- és mellérendelt mondatok nagy száma, a sok összetett mondat és az *ami* használata az *amely* helyett jellemző a fordított szövegekre. Míg Klaudy kutatási korpusza írott szakszövegekből épült fel, a jelen kutatás az audiovizuális szövegek, azon belül is a feliratok elemzésével foglalkozik. A két kutatás eredményeinek összehasonlítása hozzájárulhat a fordított szövegek jellegzetességeinek megismeréséhez.

Baker és Olohan (2000) is készített hasonló kutatást a témában, akik az angol nyelvben megjelenő opcionális *that* kötőszavas mellékmondatokat vizsgálták. A kutatás eredménye kimutatta, hogy a *that* kötőszó nagyobb számban szerepel a fordításokban,

mint az autentikus szövegekben. Ebből az a következtetés vonható le, hogy az angol nyelvben a *that* kötőszavas mellékmondatoknak az elvárásoknál sűrűbb használata fordított szövegre utalhat.

Bánhegyi (2012) a vonatkozó mellékmondatok használatának fordítástudományi megközelítésű vizsgálata során az anyanyelv tudatos használatára hívja fel a figyelmet annak érdekében, hogy a fordítók kiküszöbölhessék a „stilisztikailag vagy éppen nyelvtanilag hibás" (Bánhegyi 2012: 287) közléseket. A fél év alatt általa lektorált, angolról magyarra fordított szövegek 61 százalékában talált helytelen *ami* használatot, 55 százalékában helytelen *amely* használatot, és 71 százalékában helytelen személyátsugárzásból származó *aki* használatot. Véleménye szerint a fordítástudománynak ugyanolyan hangsúlyosan kell kezelnie a nyelvhelyességi kérdéseket, mint a fordítás egyéb területeit. A fordítások nyelvhelyességi értékeléséhez azonban még megfelelő számú nyelvhasználati adatra van szükség, amely igazolhatja például a lektori döntések jogosságát.

A fenti kutatások is bizonyítják, hogy a fordítástudomány számára a vonatkozó mellékmondatok és a vonatkozó névmásos kifejezések használatának vizsgálata fontos adatokkal szolgálhat a fordításoktatás és a fordításértékelés számára is. A fordított magyar szövegekre jellemző jelentős mértékű *ami* használat *amely* helyett (Bánhegyi 2012; Klaudy 2017) jelenleg a sztenderd nyelvhasználattól és az autentikus magyar szövegektől való egyértelmű eltérést mutatja. A nyelv és a (sztenderd) nyelvhasználat diakrón változása azonban új kutatási eredményeket hozhat, és a jelenlegi sztenderdtől eltérő nyelvhasználat akár sztenderddé is válhat a jövőben.

A vonatkozó mellékmondatok használatának gyakorisága már az ómagyar nyelv változásait kutató nyelvészek számára is fontos kutatási terület volt. É. Kiss Katalin *Magyar generatív történeti mondattan* című művében bemutatja azt a változást, amely a vonatkozó mellékmondat használata kapcsán ment végbe az ómagyar nyelvben (2014: 31). Az 1466-ban keletkezett *Müncheni kódex* és az 1516–1519 között keletkezett *Jordánszky kódex* is tartalmazza Máté evangéliumának magyar nyelvű fordítását. Azonban a vonatkozó névmásos kifejezések száma jelentősen eltér a két fordításban: a korábbi fordítás csak

225, míg a későbbi 314 névmást tartalmaz. Ez is bizonyítja a vonatkozó mellékmondatok térnyerését a magyar nyelvben a prenominális, igeneves mondatszerkesztéssel szemben.

Kálmán László (2011) megfogalmazásában „a vonatkozó mellékmondat olyan tagmondat, ami egy mondatban említett dologhoz szorosan kapcsolódik. Vagy a közelebbi azonosítására szolgál (ez a megszorító értelmű vonatkozó mellékmondat), vagy újabb állítást tesz róla (ez a nem megszorító vonatkozó mellékmondat). A vonatkozó mellékmondatot vonatkozó névmások vezethetik be". *A Magyar leíró nyelvtan* (Kálmán 2001: 143) szerint „a vonatkozó mellékmondatot *mindig* vonatkozó névmást tartalmazó vonatkozó kifejezés vezeti be" (szerző kiemelése). Kutatásom első felében az ezen definíciókhoz igazodó vonatkozó mellékmondatok szinkrón gyakoriságát vizsgálom autentikus és fordított magyar nyelvű szövegekben.

5. A főnévi vonatkozó névmásos kifejezések sztenderd használata

Az *Új magyar nyelvtan* (É. Kiss et al. 2003: 95) megállapítja, hogy „a vonatkozó mellékmondat kötőszói pozícióját vonatkozó névmásos kifejezés tölti ki". Nyelvészeti szempontból elmondható még, hogy „kötőszó [...] jeggyel rendelkezik", és „a vonatkozó mellékmondat igéjének valamely vonzatát vagy szabad határozóját is képviseli". A nyelvhasználati és „nyelvhelyességi" szempontok azonban még tovább árnyalhatják ezt a leírást. A főnévi vonatkozó névmásos kifejezések, különösen az *ami* és az *(a)mely* használata talán már a magyar nyelv 1844-es hivatalos államnyelvvé tétele óta foglalkoztatja a nyelvművelés iránt érdeklődőket. A nyelvművelők a sztenderd nyelvhasználat alapján az *ami* használatát az *amely* helyett „nyelvhelyességi hibának" tartják. Kálmán László nyelvész 2011-ben a *Nyelv és Tudományban* írt cikkét alapul véve a főnévi vonatkozó névmásos kifejezések sztenderd nyelvhasználatát a következő táblázatban mutatom be.

A sztenderd nyelvhasználat alapján tehát az *ami* névmás használata abban az esetben indokolt, ha nem személyre és nem főnévre utal vissza.

1. táblázat

A magyar főnévi vonatkozó névmások sztenderd használata (Kálmán 2011 alapján)

Névmás	Mondatszerkezet		Példamondat
amely	nem személyre utaló, névszói alaptagú szerkezet	megszorító értelmű	Kiléptünk (arra) a tisztásra, (a)melyen/amelyiken a ház állt.
		nem megszorító értelmű	Kiléptünk egy tisztásra, (a)melyen egy ház állt.
	rövidítésre használt mutató névmás		Nem a legjobb ajánlatot választották, hanem azt [az ajánlatot], (a)mely/amelyik a legolcsóbb volt.
aki	személyre utaló, névszói alaptagú szerkezet		Fogadták az elnököt, aki rövid látogatásra érkezett.
ami	nem személyre utaló, nem névszói alaptagú szerkezet	az alaptag mutató névmás	Azt tették, amit jónak láttak.
		az alaptag egy tagmondat	Csökkent a kőolaj ára, ami pánikot okozott.
		hiányzik az alaptag	Megtettek értünk, amit csak tudtak.

Amennyiben a főmondatban ugyancsak nem személyre, de főnévre vagy névszói szerkezetre utal vissza a névmás, akkor csak az *amely* használandó.

Szepesy Gyula (1986: 47–50) a „nyelvi babona” egyik mintapéldájának tartja a nyelvművelők azon álláspontját, hogy helytelenítik az ami főnévre történő visszautalását. Szepesy meghatározásában a nyelvi babona tévhit, és terjesztői helytelennek nyilvánítanak „olyan nyelvi eszközöket, amelyek a nyelv rendszere szempontjából teljesen kifogástalanok” (1986: 3). Szepesy felsorol számos népnyelvi, köznyelvi és irodalmi példát arra, hogy az *ami* bizony használatos főnévre való visszautalásnál is. Sőt, még a nyelvművelők maguk is használják az *amely* helyett az *amit*, amire példát is hoz.

Azon a véleményen van, hogy az *(a)mely* használata a köznyelvben már teljesen ismeretlen. Az *ami* gyakori használata a beszélt nyelvben viszont azt eredményezheti, hogy a beszélők a kollokviális stílussal azonosítják, és emiatt az írott nyelvből száműzik a hiperkorrekció hatására.

„Vádjai, *miket* az iskola ellen emel” (Tolnai Vilmos: Nyelvőr, 32:100). „A másik hiba, *amin* gyakran megütközhetünk, az, hogy” (Balassa József: *A magyar nyelv könyve*) (Szepesy 1986: 48)

A *Magyar nyelvhasználati szótár* (Balázs és Zimányi 2007: 246–247) szerint már több évtizede tart a beszélt köznyelvből az *amely* kiszorulása, és ez a folyamat már az írott nyelvet is elérte. A szótár írói szerint azonban választékosabb és igényesebb lenne a beszédünk, ha megtartanánk benne az *amely–ami* kettősségét.

Nádasdy Ádám (2001) szerint az *ami* használata az *amely* helyett inkább stilisztikai, mint nyelvhelyességi kérdés. Abban az összetett mondatban, ahol az előzményfőnév nem személy, az *amely/ami/amelyik* névmás is használható. „Az *amely* az irodalmi-konzervatív, az *ami* a semleges, az *amelyik* a fesztelen stílusértékű. Mármost akinek emilyen a lelkülete, az az irodalmit fogja semlegesnek érezni, akinek meg amolyan, az a fesztelent.” A választásban a személyes preferencia mellett a szövegtípus is segítségünkre lehet, mivel attól függően, hogy „írott-e vagy beszélt, irodalmi vagy értekező, formális vagy fesztelen”, más-más névmás használata lehet ott és akkor a megfelelőbb (2001: 41).

Összefoglalásként elmondható, hogy a főnévi vonatkozó névmásos kifejezések használata, illetve „helyes” használata nem mai eredetű kérdés. Az *amely* névmást mindig is a művelt és főként írott nyelvi regiszterrel azonosították, az ugyanabban a funkcióban használt *ami*-t pedig a fesztelenebb beszélt nyelvi regiszterrel. A jelenlegi változás Nádasdy Ádám (2001) szerint abban áll, hogy már a kötöttebb szerkesztésű írott nyelvben is megjelent az *ami* használata az *amely* rovására. Ennek fő okát nem a nyelv „romlásában” látja, hanem a *mely* kérdőszónak a magyar nyelvből való szinte teljes eltűnésével magyarázza. Mivel „a vonatkozó névmások mind *a* + kérdőszó alakúak …, az *amely* alól a *mely* eltűnése kihúzza a talajt” (2001: 41). Ezt Nádasdy „terapeutikus változásnak” nevezi, azaz a nyelvi rendszeren esett sérülést a nyelv maga így próbálja orvosolni.

6. A kutatás bemutatása

Kutatásom célja az önkéntes TED-fordítók magyar nyelvű online feliratiban megjelenő nyelvhasználat összehasonlítása az autentikus magyar nyelvű ismeretterjesztő TED-előadások szövegtípusra jellemző nyelvhasználatával. Veszelszki Ágnes (2012) az online szövegek jövőbeni kutatási irányát abban látja, hogy „szisztematikus korpuszállítással" és „anyaggyűjtéssel" megállapítsuk a beszélt nyelvi elemek „statisztikai gyakoriságát" és a „digitális kommunikációs korpusz" adatait összehasonlítsuk egy „beszélt nyelvi korpusz" adataival (2012: 409). Kutatásom tükrözi ennek a szempontrendszernek a főbb elemeit.

Összehasonlítható kutatási korpuszomat az ELTE Fordítástudományi Doktori Program fordítástudományi kutatásokra szolgáló Pannónia Korpuszában (Robin et al. 2016, Robin és Szegh 2017) található szövegekből építettem. A korpusz ismeretterjesztő TED-előadások autentikus magyar nyelvű átiratait és fordított feliratait tartalmazza: tíz autentikus magyar nyelvű TED-előadás átiratát hasonlítottam össze tíz angolról magyarra fordított TED-előadás feliratával. Az autentikus audiovizuális szövegek időbeli hossza percben kifejezve: 134:29. A fordított audiovizuális szövegek időbeli hossza percben kifejezve: 138:39. Az összehasonlítható szinkrón korpusz tehát ugyanabba a doménbe tartozó, ugyanolyan hosszúságú szövegeket tartalmaz, megfelelve Baker (1995) korpuszépítési szempontjainak. A feliratok pontos terjedelmi adatait az alábbiakban a *2. táblázat* tartalmazza.

A korpuszban az adatgyűjtés és az adatok feldolgozása félautomatikus módon történt. Az *AntConc* számítógépes korpuszelemző program segítségével először kigyűjtöttem a szövegekben található vonatkozó mellékmondatokat, amelyeket vonatkozó névmás vezet be, és összehasonlítottam a gyakorisági előfordulásukat az autentikus és a fordított szövegekben.

Ezután a főnévi vonatkozó névmásos kifejezéseket vizsgáltam meg a szövegekben, és manuálisan sztenderd vagy nem sztenderd nyelvhasználati kategóriákba soroltam őket az *1. táblázat* alapján. A kapott számszerű eredményeket végül párhuzamba állítottam egymással.

2. táblázat

Az autentikus átiratok és a fordított feliratok szavainak száma

Autentikus szövegek		**Fordított szövegek**	
Balogh	842	Ariely	2572
Baritz	1124	Biddle	1530
Freund	2053	Evans	1656
Gyurkó	2212	Goldstein	2191
Hankiss	2542	Jones	1547
Joós	852	Kemp-Robertson	1608
Mányai	1086	Sandel	1443
Nemes	1589	Shields	561
Szentesi	1047	Tapscott	2067
Szvetelszky	2242	Veitch	902
Összesen	**15589**	**Összesen**	**16077**

Három kutatási kérdést (K) fogalmaztam meg, és hozzájuk kapcsolódóan három kutatási hipotézist (H) állítottam fel.

K1: Vajon a vonatkozó mellékmondatok ugyanolyan gyakorisággal fordulnak elő az autentikus és a fordított szövegekben?

K2: Vajon ugyanazok a leggyakrabban használt vonatkozó névmásos kifejezések az autentikus és a fordított szövegekben?

K3: Vajon a sztenderd nyelvhasználatot alapul véve eltérő vagy hasonló a vonatkozó névmásos kifejezések használata az autentikus és a fordított szövegekben?

H1: A fordított szövegek több vonatkozó mellékmondatot tartalmaznak majd, mivel a fordítók az egyértelműsítés elve alapján jobban kifejtik a forrásnyelvi szöveg tartalmát.

H2: Az autentikus és a fordított szövegek ugyanazokat a vonatkozói névmásos kifejezéseket használják a leggyakrabban, mivel a célnyelvnek megfelelő szövegszerkesztés elve ezt feltételezi.

H3: Az autentikus (átirat) és a fordított szövegekre (felirat) is a nem sztenderd nyelvhasználat a jellemző a szóbeli interferencia miatt.

A TED feliratozói közösség saját nyelvhasználati elvárásokat is megfogalmazott, amelyeket a feliratozók megpróbálnak betartani.[1] A kutatás szempontjából a fontosabb irányelvek a következők: 1. informális stílus használata a formális (*academic*) stílus helyett; 2. modern kifejezések használata a hagyományossal szemben; 3. személyes stílus a semleges személytelen helyett; 4. globális nyelvi kifejezések használata a regionális helyett. A TED-közösség feliratozási elvárása tehát az, hogy az önkéntes fordítók a fordított feliratokban az adott nyelv mai, élő, modern használatához igazodjanak. Ezek az elvárások erősíthetik a szóbeliségre jellemző nyelvi megformálás gyakoriságát a TED-feliratokban.

7. Kutatási eredmények

A kutatás első felében az autentikus magyar nyelvű és a fordított magyar nyelvű feliratokban a vonatkozó mellékmondatok gyakoriságát vizsgáltam. Az *AntConc* korpuszelemző programmal végzett szólistás keresés segítséget nyújtott a ragozott és ragozatlan alakban álló vonatkozó névmásos kifejezések kigyűjtéséhez. Az autentikus szöveg 46, míg a fordított szöveg 53 különböző vonatkozó névmásos kifejezést tartalmazott. A különbség oka főként a ragozott alakok eltérő száma volt: az *aki* névmás például az autentikus szövegben ötféle ragozott alakban fordult elő, míg a fordított szövegben hétféle alakban. Az adott névmást és annak ragozott alakjait nem együttesen, hanem külön-külön vizsgáltam a pontosabb eredmények érdekében. A vonatkozó mellékmondatok konkordanciaprogrammal végzett ellenőrzés utáni számadatait az alábbiakban a *3. táblázat* mutatja.

1 https://translations.ted.com/Portal:Magyar

3. táblázat

Vonatkozó mellékmondatok gyakorisága a korpuszban

	Vonatkozó mellékmondatok száma	**Eltérés**
Autentikus magyar átiratokban	308	
Fordított magyar feliratokban	403	+31%

Az eredmények alapján a fordított szövegek 31 százalékkal több vonatkozó mellékmondatot tartalmaznak. Ez a különbség eredhet abból, hogy a fordító megpróbál egyértelműsíteni, ezért a mondatokat kisebb szegmensekre bontja. Eredhet emellett abból is, hogy a forrásnyelvi szöveg tartalmát a fordító ilyen módon teszi könnyebben feldolgozhatóvá a célközönség számára. Természetesen az a lehetőség sem zárható ki, hogy maga az angol forrásnyelvi szöveg tartalmazott több mellékmondatot.

Az alárendelő mondatok fordított szövegekben való gyakori használatát egyes nyelvművelők nem tekintik összeegyeztethetőnek a szokásos nyelvhasználattal, mivel a magyart inkább „mellérendelő nyelvnek” tartják. A *Nyelvművelő kéziszótár* (Grétsy és Kemény 2005: 511) szerint azonban ez a megállapítás „nyelvhelyességi babonának” tekinthető. Az indoeurópai nyelvekkel összehasonlítva viszont „lényegesen gyakoribb összetett mondatainkban a mellérendelő viszony”. A *2. táblázat* eredménye alapján a vizsgált alárendelő mondattípus száma a fordított feliratokban magasabb, ennek oka akár forrásnyelvi interferencia is lehet.

A kutatás következő részében arra kerestem a választ, hogy a fent említett 46, illetve 53 vonatkozó névmásos kifejezés közül vajon melyeket használják leggyakrabban az autentikus és a fordított szövegek. Az eredményeket a következő táblázatban összesítettem. A táblázatból egyértelműen kiolvasható, hogy a leggyakrabban használt vonatkozó névmás mindkét alkorpuszban az *ami* és ragozott alakja, az *amit*.

A leggyakrabban használt vonatkozó névmások között az *amely* sehol sem jelenik meg.

4. táblázat

A leggyakoribb vonatkozó névmásos kifejezések

Autentikus magyar előadás átiratokban				**Fordított magyar feliratokban**			
1	*ami*	63	32%	1	*amit*	75	28%
2	*amikor*	35	18%	2	*amikor*	55	20%
3	*aki*	29	15%	3	*ami*	47	17%
4	*amit*	23	12%	4	*akik*	28	10,2%
5	*ahol*	22	11%	5	*ahogy*	26	9,6%
6	*akik*	15	7%	6	*ahol*	25	9,2%
7	*ahogy*	10	5%	7	*aki*	16	6%
Össz.		**197**	**100%**	**Össz.**		**272**	**100%**

Sőt, az *ami* és az *amit* használata annyira dominál a szövegekben, hogy kettejük együttes használata a leggyakrabban használt vonatkozó névmások majdnem 50 százalékát adja. Ezt az adatot leolvashatjuk az alábbi, *5. táblázat* összegzéséből is.

A beszélt nyelvi előadásanyagok átirata a százalékos adatok alapján körülbelül ugyanolyan arányban használja az élő beszédre inkább jellemző *ami* névmást, mint az előadások írott fordítása. Az eredmények alapján úgy tűnik, hogy az ismeretterjesztő előadások fordított feliratai megőrzik az élőnyelvi beszéd jellegzetességeit, legalábbis a vonatkozó névmások tekintetében.

A kutatás további részében azt is megvizsgálom, hogy az *ami* sztenderd vagy nem sztenderd használata jelenik-e meg a két szövegcsoportban.

A kutatás eredményei a vonatkozó névmások használatának egy másik érdekességére is rámutatnak. Észrevehető (lásd *4. táblázat*), hogy ugyanazokat a vonatkozó névmásokat alkalmazza mindkét szövegcsoport, csak egy kissé más sorrendben a százalékban kifejezett gyakoriság alapján.

5. táblázat

Az ami és az amit vonatkozó névmás előfordulási gyakorisága

Autentikus magyar előadás átiratokban				**Fordított magyar feliratokban**			
1	*ami*	63 db	32%	1	*amit*	75 db	28%
4	*amit*	23 db	12%	3	*ami*	47 db	17%
Össz.		**86 db**	**44%**	**Össz.**		**122 db**	**45%**

6. táblázat

Vonatkozó névmásos kifejezések autentikus és fordított szövegekben

Autentikus szövegek			**Fordított szövegek**	
1	*ami*		1	*amit*
2	*amikor*		2	*amikor*
3	*aki*		3	*ami*
4	*amit*		4	*akik*
5	*ahol*		5	*ahogy*
6	*akik*		6	*ahol*
7	*ahogy*		7	*aki*

A *6. táblázat* adatai azt igazolják, hogy az írott beszéltnyelviség elemei jelennek meg a TED-feliratozásban. Az autentikus magyar nyelvű előadások ugyanazokat a vonatkozó névmásos kifejezéseket tartalmazzák leggyakrabban, mint a magyarra fordított, írott feliratok.

A kutatás következő részében a főnévi vonatkozó névmásos kifejezések használati gyakoriságát és „helyességét” vizsgáltam, azaz a sztenderd nyelvhasználathoz történő igazodást vagy az attól való eltérést. Az *amely* és *mely* névmást külön kezeltem, mivel valószínűsíthető volt, hogy a *mely* ritkábban fordul elő a szövegekben. A vonatkozó névmások általában *a*+kérdőszó alakúak (Nádasdy 2001), így a *mely* névmás atipikus formája a kisebb gyakorisági előfordulás felé mutat. Az *amelyik* névmást is külön kezeltem, mivel

Nádasdy (2001) szerint napjainkban az *ami/amely* pozícióharc mellett az *amelyik* névmás elterjedése *aki* helyett a másik jelentős nyelvi változás.

7. táblázat

Főnévi vonatkozó névmások autentikus magyar szövegekben

	Összes db	**Sztenderd nyelv-használat**	**Nem sztenderd nyelvhasználat**	**Nem sztenderd nyelvhasználat (%)**
aki+	55	55	0	0 %
ami+	120	41	79	66%
amely+	14	14	0	0%
mely+	4	4	0	0%
amelyik+	4	3	1	25%

A *7. táblázat* az autentikus magyar szövegekben előforduló főnévi vonatkozó névmásos kifejezések használati gyakoriságát mutatja. A + jel arra utal, hogy az adott névmás ragozott alakjait is figyelembe vettem az adatok kigyűjtésekor. Az *aki* személyre utaló névmás jelentős számban szerepelt a szövegekben, és az előadók mindenhol a sztenderd nyelvhasználatnak megfelelően alkalmazták. Személyátsugárzásos használatra tehát nem volt példa, a beszélők személytelen dolgokra vagy intézményekre való utaláskor nem használták *ami/amely* helyett.

Ugyancsak megfelelően használta minden előadó a *mely* névmást, bár használati gyakorisága elenyésző volt. A *mely* az élőbeszédhez hasonlító feliratokban már alig jelenik meg, ami utalhat arra, hogy ebben a szövegtípusban már nem rendszeres a használata. Aki viszont mégis használja, tudatában van a szokásos használatával.

Az *ami* és az *amely* használatában releváns különbség mutatkozik, mivel az *ami* előfordulási gyakorisága nyolcszor nagyobb, mint az *amely* névmásé. Ha a sztenderd nyelvhasználatot nézzük, akkor az *amely* használata szinte mindenhol megfelel a nor-

máknak. A *mely* névmáshoz hasonlóan ebben az esetben is megállapíthatjuk, hogy aki az *amely* névmást egyáltalán alkalmazza, tudatában is van a „helyes" használatának.

Az *ami* alkalmazását illetően a művelt társadalmi rétegbe tartozó előadók már 66 százalékban eltértek a sztenderd nyelvhasználattól. A főnévre történő visszautalásoknál is *ami*-t használtak, ahogyan az alábbi példák is mutatják:

(1) „[…] egy értékrend a mai világban, ami szerintem alternatívát jelenthet"
(2) „[…] 66%-át teszi ki a pletyka, ami azért nem rágalom"
(3) „]…] Az első tagolt beszéd, ami minden emberi közösséget jellemez"
(4) „[…] és a gazdasági tevékenység, amit eddig folytattam"

Ez is bizonyítja, hogy a művelt beszédben az *ami* használata már jelentős mértékben átvette az *amely* funkcióit. A magyarul beszélő előadók, akik jellemzően egyetemi végzettséggel rendelkező szakemberek, kutatók és oktatók, már 66 százalékban az *amely* szerepébe bújtatják az *ami* névmást.

Az *amelyik* névmást alig használták a beszélők, ráadásul a négy előfordulásból egyszer (lásd 6. példa) nem a sztenderd használatnak megfelelően.

(5) „[…] pirossal van az a nyúlványa, amelyikkel gátol"
(6) „Kell ehhez egy ún. szuperpolip, amelyik minden medence felett állva"

További kutatás szükséges ahhoz, hogy az *amelyik* szó használatát a magyar nyelvben jobban leírhassuk. Jelenleg az állapítható meg, hogy élőbeszédben nem gyakran használatos, és használati szokásaival sincsenek tisztában a beszélők. Nádasdy Ádám (2001) szerint viszont az *amelyik* épphogy egyre inkább terjed a magyar nyelvben. Idézi az 1961-es akadémiai nyelvtant, amely az *amelyik* névmás beszélt nyelvben való elburjánzásától tart. Ezt a feltételezést a jelen kutatás eredményei nem tudják alátámasztani. Nádasdy szerint az *amelyik* elterjedése a magyar nyelv rendszerének átalakulásához vezethet:

> […] ugyanis e névmás érzéketlen a személy/nemszemély különbségre. (*A betegség, amelyik...; A nővér, amelyik...*) Az *ami/amely* pozícióharc ezt nem érinti, hiszen az a nemszemély kategórián belül zajlik. Az *amelyik* viszont a hagyományos kétkategóriás rendszer helyett – mely például a franciával (*qui/que*) vagy az angollal (*who/which*) analóg – olyan egykategóriás rendszer felé mutat, mint az olasz (*che*), a német (*der/die/das*) vagy az orosz (*kotóryj*).

A jelen tanulmányban bemutatott korpuszalapú kutatás adatai szerint ez a rendszerszintű változás egyelőre még nem történt meg a magyar nyelvben. A fordított magyar szövegekben a főnévi vonatkozó névmásos kifejezések használati gyakorisága (*8. táblázat*) csak az *ami* és *amely* esetén mutat eltérést az autentikus szövegeknél kapott eredményekhez képest.

8. táblázat

Főnévi vonatkozó névmások fordított magyar szövegekben

	Összes db	**Sztenderd nyelvhasználat**	**Nem sztenderd nyelvhasználat**	**Nem sztenderd nyelvhasználat (%)**
aki+	52	51	1	2%
ami+	160	75	85	53%
amely+	29	29	0	0%
mely+	7	7	0	0%
amelyik+	2	0	2	100%

Ennek egyik oka lehet a vonatkozó mellékmondatok fordított szövegekben megmutatkozó nagyobb gyakorisága.

Az *aki* gyakorisága és használati jellemzői megegyeznek az autentikus szövegek vizsgálatakor kapott eredményekkel. Egyetlen példát találunk a személyátruházásra. A *Magyar nyelvhasználati szótár* (Balázs és Zimányi 2007: 247) szerint: „Újabb fejlemény

az *aki* névmás olyan használata, amikor személyekből álló csoportokra vonatkoztatják". Erre a használatra egyetlen előfordulást ad a vizsgált korpusz:

(7) „Csak néhány példa azokból a cégekből, akik vásárolnak a műanyagunkból."

A *mely* és *amely* használatát illetően itt is megállapíthatjuk, hogy aki alkalmazza, ismeri is a szokásos használatot. Az egyetlen különbség a gyakoriságban figyelhető meg: kétszer több ezen névmások előfordulása a fordított szövegekben. Az írott beszéltnyelvi szöveg tehát még próbálja megőrizni az *(a)mely* használatát mint a beszélt nyelvvel szembeni saját jellegzetességet, de az *ami* névmáshoz viszonyított aránya, amely jelen adatok szerint a használati gyakoriság egy ötöde, jóval alacsonyabb.

A vonatkozó mellékmondatok számának növekedését a fordított szövegben az *ami* gyakoribb használata is mutatja. A vizsgált korpuszban az *ami* használatának 30 százalékos növekedése figyelhető meg a fordított szövegekben (160 db) az autentikus magyar szövegekhez képest (120 db). A nem sztenderd használat 53 százalék, azaz az *amely* helyett az írott beszélt nyelv is már jelentős mértékben az *ami* névmást használja.

Az *amelyik* használata a fordított szövegekben sem jelentős, ráadásul ezek is eltérnek a sztenderd használattól. A két előforduló példa a következő:

(8) „[…] egyfajta kiterjedt, globális, megosztott főkönyv, amelyik sok millió számítógépen fut egyszerre"
(9) „[…] ha egy olyan levél érkezik, amelyik így kezdődik"

Az *amelyik* használata az adatok alapján a vizsgált korpuszra nem jellemző. Az autentikus (A) és a fordított (F) szövegekben előforduló főnévi vonatkozó névmásos kifejezések használati gyakorlatiságát a következő összesített táblázat mutatja.

Az eredmények nem mutatnak jelentős különbséget a főnévi vonatkozó névmásos kifejezések sztenderd vagy nem sztenderd százalékos előfordulásában az autentikus és a fordított, a beszélt és az írott nyelvben.

9. táblázat

Főnévi vonatkozó névmások autentikus (A) és fordított (F) magyar szövegekben

	Főnévi vonatkozó névmások									
	aki+		**ami+**		**amely+**		**mely+**		**amelyik+**	
	A	**F**	**A**	**F**	**A**	**F**	**A**	**F**	**A**	**F**
Összes	55	52	120	160	14	29	4	7	4	2
Sztenderd	55	51	41	75	14	29	4	7	3	0
Nem sztenderd	0	1	79	85	0	0	0	0	1	2
Nem sztenderd %	0%	2%	66%	53%	0%	0%	0%	0%	25%	100%

É. Kiss Katalin (2003: 99) szerint már „nem érezzük, hogy az *ami*, *amely* és *amelyik* azonos vagy különböző kategóriájú-e […]. A vonatkozó névmások közötti különbségek elhomályosulása arra utal, hogy a vonatkozó névmási rendszer egyszerűsödőben van”. A jelen kutatás eredményei is ezt támasztják alá.

8. Összefoglalás

A kutatás eredményei igazolták a kutatási kérdések alapján megfogalmazott hipotéziseket. Kimutatható volt, hogy a fordított szövegek több vonatkozói mellékmondatot tartalmaznak, mint az autentikus magyar nyelvű szövegek. Az alárendelő mellékmondatok nagyobb számát valószínűleg a fordítás folyamatában bekövetkezett fordítói döntések idézték elő, azaz az egyértelműsítésre és a könnyebb feldolgozhatóságra való törekvés. Habár a számbeli eltérést a forrásnyelvi szövegek felépítése is befolyásolhatta, korábbi kutatások (Klaudy 2017) a fordított szövegekre jellemző lehetséges sajátosságként említik az összetett mondatok gyakori használatát. A jelen kutatás eredményei is ezt a megállapítást támasztják alá.

Az autentikus magyar nyelvű előadások és a magyarra fordított angol nyelvű előadások hasonlóan alkalmazták a vonatkozói névmásos kifejezéseket. Ennek hátterében egyrészről a célnyelvi normának megfelelő szövegszerkesztés igénye állhat, azaz a for-

dított szöveg az autentikus magyar nyelvű szövegekhez alkalmazkodik. Másrészről ez a hasonlóság már szövegtípusra jellemző sajátosság is lehet, mivel az ismeretterjesztő előadások feliratai átvették a beszélt nyelv sztenderd változatára jellemző szóbeli sajátosságokat. Ezek a feliratok az írott beszéltnyelviség jegyeit hordozzák, azaz mind a beszédre, mind az írásra jellemző sajátosságokkal rendelkeznek. A beszélt nyelvre utal az *ami* névmás jelentős nem sztenderd használata, viszont az írott nyelvre utal az *(a)mely* gyakoribb megjelenése. Míg korábbi fordítástudományi kutatások (Klaudy 2017, Bánhegyi 2012) a magyarra fordított szövegek nyelvhelyességi normától való eltérésére mint helytelen használatra hívták fel a figyelmet, a jelen kutatás adatai azt mutatják, hogy a magyarra fordított TED-feliratokban maga az írott beszéltnyelviség, a sztenderd magyar nyelvváltozattól való eltérés jelentkezik normaként.

A fordítástudomány számára érdekes kutatási irány lehet más nyelvi változók vizsgálata is. Schirm Anita (2005: 93–99) a magyar nyelv jelenlegi állapotáról írt cikkében felsorol például az írott beszéltnyelvben megjelenő szóbeliségre utaló más jegyeket is. Kutatási terület lehet a *-ban/ben* toldalék átalakulása *-ba/be* formává, a *határozó+hogy* kötőszavas (*valószínűleg, hogy; remélhetőleg, hogy*) szerkezetek elterjedése vagy a leggyakrabban használt szófajok vizsgálata.

A nyelv dinamikus változását társadalmi, technikai és gazdasági folyamatok eredményezik. A digitális forradalom korszakában az internet és a telekommunikációs eszközök szinte globális elterjedése és használata hatással van a társadalom és a nyelv alakulására egyaránt. A digitalizáció megváltoztathatja a nyelvhasználatot, és átalakíthatja a fennálló nyelvváltozatok presztízsrendszerét. A jelen kutatás eredményei arra a tendenciára hívják fel a figyelmet, hogy az eddigi beszélt és írott nyelv közötti dichotómia inkább egy skála két végpontjára utal, mivel a beszéltnyelvi normák már az írott regiszterre is jellemzők bizonyos szövegtípusokban.

Az anyanyelvoktatásban, a fordításoktatásban és a lektorképzésben is tudatosítani szükséges a nyelv dinamikus jellegét. A hallgatóknak a későbbi munkájuk megfelelő elvégzéséhez meg kell ismerniük a nyelvben létező különböző nyelvváltozatokat, és szövegtípustól, szituációtól függően képesnek kell lenniük eldönteni, melyik nyelvváltozat

használata a megfelelő. A jelen kutatás eredményei az audiovizuális fordítás és feliratozás oktatása során nyújthatnak támpontot a fordítói döntések meghozatalához. A TED-előadások feliratozása során az írott beszéltnyelviség tűnik a normának, így az ehhez tartozó stílusjegyek használatához érdemes igazodni a feliratozóknak.

Irodalom

Assis Rosa, A. 2001. Features of oral and written communication in subtitling. In: Gambier Y., Gottlieb, H. (eds) (*Multi)Media translation: Concepts, Practices and Research*. (Benjamins Translation Library). Amsterdam/Philadelphia: John Benjamins Publishing Company. 213–222. https://doi.org/10.1075/btl.34.26ass

Baker, M. 1995. Corpora in Translation Studies: An Overview and Some Suggestions for Future Research. *Target* Vol. 7. No. 2. 223–243. https://doi.org/10.1075/target.7.2.03bak

Baker, M., Olohan, M. 2000. Reporting *that* in translated English. Evidence for subconscious processes of explicitation? *Across Languages and Cultures* Vol. 1. No. 2. 141–158. https://doi.org/10.1556/Acr.1.2000.2.1

Balázs G., Bódi Z. (szerk.) 2005. *Az internetkorszak kommunikációja. Tanulmányok*. Budapest: Gondolat–INFONIA.

Balázs G., Zimányi Á. (szerk.) 2007. *Magyar nyelvhasználati szótár*. Celldömölk: Pauz–Westermann.

Bánhegyi M. 2012. Néhány gondolat a fordítói és lektori munkáról a magyar vonatkozó névmások használatának jelenkori változása és a nyelvhelyesség tükrében. In: Pintér, T., Pődör D., P. Márkus K. (szerk.) *Szavak pásztora. Írások Magay Tamás tiszteletére*. Szeged: Grimm Kiadó. 287–293.

Bartsch, R. 1987. *Norms of Language: Theoretical and Practical Aspects*. London/New York: Longman.

Bódi Z. 2004. *A világháló nyelve. Internetezők és internetes nyelvhasználat a magyar társadalomban.* Budapest: Gondolat Kiadó.

É. Kiss K. (szerk.) 2014. *A magyar generatív történeti mondattan*. Budapest: Akadémiai Kiadó.

É. Kiss K., Siptár P., Kiefer F. 2003. *Új magyar nyelvtan*. Budapest: Osiris Kiadó.

Grétsy L., Kemény G. (szerk.) 2005. *Nyelvművelő kéziszótár*. Budapest: Tinta Könyvkiadó.

Heltai P. 2004. A fordító és a nyelvi normák I. *Magyar Nyelvőr* 128. évf. 4. szám.

Kálmán L. 2011. Ami és amely. *Nyelv és Tudomány.* 2011. január 26.

Kálmán L. (szerk.) 2001. *Magyar leíró nyelvtan*. *Mondattan I*. Budapest: MTA–ELTE Elméleti nyelvészet szakcsoport.

Kiefer F. (szerk.) 2006. *Magyar nyelv*. Budapest: Akadémiai Kiadó.

Kiss J. 1996. *Társadalom és nyelvhasználat. Szociolingvisztikai alapfogalmak*. Budapest: Nemzeti Tankönyvkiadó.

Klaudy K. 2017. Eredeti magyar szakszöveg vs fordított magyar szakszöveg. Elhangzott: MANYE-MANYSI Konferencia. *Magyar szaknyelvek a Kárpát-medencében*. Nemzetközi magyar szaknyelvi konferencia. 2017. január 18–20. Budapest: Pázmány Péter Katolikus Egyetem.

Labov, W. 1972. *Sociolinguistic Patterns*. Oxford: Blackwell.

Molnár M. 2014. A nyelvi norma a magyar nyelvtudományban. *Hungarológiai évkönyv* 15. évf. 1. szám. 41–54.

Nádasdy Á. 2001. A burjánzó vonatkozó. *Magyar Narancs* 28. szám. 41.

Nádasdy Á. 2010. Búcsú a nyelvhelyességtől. *Élet és Irodalom* 54. évf. 32. szám. 1–4.

Robin E., Dankó Sz., Götz A., Nagy A. L., Pataky É., Szegh H., Török G., Zolczer P. 2016. Fordítástudomány és korpuszkutatás: bemutatkozik a Pannónia Korpusz. *Fordítástudomány* 18. évf. 2. szám. 5–26.

Robin E., Szegh H. 2018. A Pannónia Korpusz audiovizuális alkorpusza. In: Dróth J. (szerk.) *Gépiesség és kreativitás a fordítási piacon és az oktatás különböző szintjein.* Budapest: KGRE, L'Harmattan Kiadó. 93–110.

Schirm A. 2005. Jelentés a mai magyar nyelvről – 2005. In: Balázs G. (szerk.). *Jelentés a magyar nyelvről 2000–2005*. Budapest: Akadémiai Kiadó. 83–116.

Szepesy Gy. 1986. *Nyelvi babonák*. Budapest: Gondolat.

Tolcsvai Nagy G. 1998. *A nyelvi norma*. Budapest: Akadémiai Kiadó.

Veszelszki Á. 2012. A digilektus beszélt nyelvi jellemzői. In: Bárdosi V. (szerk.). *Tanulmányok. Nyelvtudományi Doktori Iskola. Asteriskos 1. Eötvös Loránd Tudományegyetem Bölcsészettudományi Kar doktori iskolák tanulmányai.* Budapest: ELTE BTK. 399–417.

Források

Autentikus magyar nyelvű TED-feliratok

Balogh T. 2016. Az improvizáció stratégiája. TEDxDanubia. http://www.tedxdanubia.com/videos/az-improvizacio-strategiaja-balogh-tamas-a-zenerol-tamas-balogh-tedxdanubia-2016

Baritz S. 2010. Nézőpont, Változás. TEDxDanubia. http://www.tedxdanubia.com/videos/nezopont-valtozas:-baritz-sarolta-laura-at-tedxdanubia-2010

Freund T. 2011. Agyhullámok és kreativitás. TEDxDanubia. http://www.tedxdanubia.com/videos/tedxdanubia-2011-freund-tamas-agyhullamok-es-kreativitas

Gyurkó S. 2013. Luke, én vagyok az apád. TEDxDanubia. http://www.tedxdanubia.com/videos/luke-en-vagyok-az-apad:-gyurko-szilvia-at-tedxdanubia-2013

Hankiss E. 2011. Életstratégiák a bizonytalanság korában. TEDxDanubia. http://www.tedxdanubia.com/videos/tedxdanubia-2011-hankiss-elemer--eletstrategiak-a-bizonytalansag-koraban

Joós A. 2015. A tanári pálya márpedig menő. TEDxDanubia. http://www.tedxdanubia.com/videos/a-tanari-palya-marpedig-meno-andrea-joos-tedxdanubia

Mányai R. 2017. A tárgy-kultúra robbanás. TEDxDanubia. https://www.youtube.com/watch?v=A9w-jcJ8GZk

Nemes O. 2014. Mit kezdjünk a fiatalokkal? TEDxDanubia. http://www.tedxdanubia.com/videos/mit-kezdjunk-a-fiatalokkal:-nemes-orsolya-at-tedxdanubia-2014

Szentesi É. 2016. Visszaút a rák halálos ítéletéből. TEDxDanubia. http://www.tedxdanubia.com/videos?performer=2824

Szvetelszky Á. 2011. A pletyka tudománya és művészete. TEDxDanubia. http://www.tedxdanubia.com/videos/tedxdanubia-2011-szvetelszky-zsuzsa-a-pletyka-tudomanya-es-muveszete

Fordított magyar nyelvű TED-feliratok

Ariely, D. 2012. Mitől szeretjük a munkánkat? TEDxRiodelaPlata. https://www.ted.com/talks/dan_ariely_what_makes_us_feel_good_about_our_work?language=hu

Biddle, M. 2011. A műanyagot is újrahasznosíthatjuk! TEDGlobal. https://www.ted.com/talks/mike_biddle?language=hu

Evans, Ph. 2013. Hogyan alakítják át az adatok az üzletet. TED@BCG. https://www.ted.com/talks/philip_evans_how_data_will_transform_business/up-next?language=-hu

Goldstein, D. 2011. A jelenbeli és jövőbeli énünk csatája. TEDSalon NY. https://www.ted.com/talks/daniel_goldstein_the_battle_between_your_present_and_future_self/transcript?language=hu

Jones, V. 2010. A műanyag okozta gazdasági igazságtalanság. TEDxGreatPacificGarbagePatch. https://www.ted.com/talks/van_jones_the_economic_injustice_of_plastic/up-next?language=hu

Kemp-Robertson, P. 2013. Bitcoin. Verejték és Tide. A márkázott valuták jövője. TEDGlobal. https://www.ted.com/talks/paul_kemp_robertson_bitcoin_sweat_tide_meet_the_future_of_branded_currency/up-next?language=hu

Sandel, M. 2013. Miért ne bízzuk a piacra a társadalmi életet. TEDGlobal. https://www.ted.com/talks/michael_sandel_why_we_shouldn_t_trust_markets_with_our_civic_life/up-next?language=hu

Shields, A. 2015. Hogyan keressük az élet nyomait idegen bolygókon? TED2015. https://www.ted.com/talks/aomawa_shields_how_we_ll_find_life_on_other_planets/up-next?language=hu

Tapscott, D. 2016. Hogy alakítja át a blokklánc a pénz és az üzleti világot. TEDSummit. https://www.ted.com/talks/don_tapscott_how_the_blockchain_is_changing_money_and_business/up-next?language=hu

Veitch, J. 2015. Ez történik, mikor válaszolunk egy spam emailre! TEDGlobal>Geneva. https://www.ted.com/talks/james_veitch_this_is_what_happens_when_you_reply_to_spam_email/up-next?language=hu

Harmadik kód a tolmácsolásban: létezik tolmácsolási szöveg?

Intermodális korpuszok és esettanulmány

Szegh Henriett

laparisienne.gov@gmail.com

Eötvös Loránd Tudományegyetem

Fordítástudományi Doktori Program

Kivonat: A jelen tanulmány egy 2017-ben megkezdett, három részből álló kutatási projekt második szakaszának eredményeiről számol be. A projekt célja, hogy megállapítsam, vajon létezik-e a fordítási szöveg (Károly 2007) analógiájára tolmácsolási szöveg is, ha pedig igen, milyen univerzális tulajdonságok jellemzik. Feltételezésem szerint a tolmácsolt szövegek, mind forrásnyelvi eredetijükhöz, mind a fordításokhoz, mind az eredetileg célnyelven elhangzott szövegekhez képest egyszerűbbek, implicitebbek, kevésbé redundánsak, ez pedig leginkább a lexikai kihagyásokban nyilvánul meg. A háttérben valószínűleg maga a tolmácsolási mód, a tolmácsolás mentális műveleteinek komplex karaktere, a tolmácsolás során megjelenő időkorlát és a rövid távú memória limitált tárolási kapacitása áll. A kutatás első szakaszában egy kétnyelvű párhuzamos, eredeti hangzó szövegeket és azok tolmácsolt átiratát tartalmazó korpusszal dolgoztam (Szegh 2017). Második lépésként intermodális elemzést végeztem, megvizsgáltam ugyanazon forrásnyelvi szöveg tolmácsolt és fordított változatát is. Mindkét esetben azonos módszertan szerint haladtam: először statisztikai vizsgálatnak vetettem alá a szövegeket, majd a kvalitatív elemzés során igyekeztem a számadatok mögé nézni, és felfedni, mi áll a felszíni szövegjellemzők és a tolmácsok döntéseinek hátterében. Az eredmények az első és a második szakaszban is megerősítették előzetes hipotézisemet. A kutatás folytatásaként autentikus szövegekkel tervezem összehasonlítani a tolmácsolt szövegeket, a kapott adatokat és az azokból levont következtetéseket pedig egy következő cikkben kívánom ismertetni.

Szegh Henriett: Harmadik kód a tolmácsolásban: létezik tolmácsolási szöveg? Intermodális korpuszok és esettanulmány. In: Robin E., Seidl-Péch O. (szerk.) 2020. *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordításkutatás Magyarországon.* Segédkönyvek a nyelvi közvetítésről I. Budapest: ELTE BTK Fordítástudományi Doktori Program, MANYE Fordítástudományi Szakosztály. DOI: https://doi.org/10.36252/Nyelvikozvsegedkonyv1.11

Kulcsszavak: tolmácsolási szöveg, fordítási szöveg, intermodális korpusz, párhuzamos korpusz, összehasonlítható korpusz

1. Bevezetés: korpusznyelvészet és leíró fordítástudomány

A modern korpusznyelvészet kezdetei a 60-as évekig nyúlnak vissza, amikor Francis és Kučera megalkották a gépileg is olvasható, egymillió szövegszót tartalmazó, kimondottan nyelvészeti célokkal összeállított Brown Corpust. A korpusznyelvészet kifejezést azonban csak 1984-től használjuk terminusként, ebben az évben adta ki ugyanis Jan Aarts és Willem Meijs *Corpus Linguistics: Resent Developments in the Use of Computer Corpora in English Language Research* címet viselő tanulmánykötetét. A korpusznyelvészet fókuszában a korpusz áll, amely a Magyar Tudományos Akadémia meghatározása szerint:

> […] ténylegesen előforduló írott, vagy lejegyzett beszélt nyelvi adatok gyűjteménye. A szövegeket valamilyen szempont szerint válogatják és rendezik. Nem feltétlenül egész szövegeket tartalmaz és nem csak tárháza a szövegeknek, hanem tartalmazza azok bibliográfiai adatait, bejelöli a szerkezeti egységeket (bekezdés, mondat). Emellett pedig feltünteti a szavak mellett szófaji kódjukat is.[1]

A korpuszokat a mai modern kor kutatói már elektronikus formában tárolják, így téve lehetővé nagy mennyiségű szöveg feldolgozását, valamint a gyors, rugalmas hozzáférést. Bár a definícióban az szerepel, hogy a korpuszt felépítő szövegeket előre meghatározott szempontok szerint és céllal válogatják, egyes nézetek szerint korpusznak nevezhető maga a világháló is (Sinclair 1991). A szigorú kritériumok szerint összeállított, kiegyensúlyozott korpuszok előnye egyrészről, hogy hatalmas adatbázist biztosítanak az empirikus alapokon nyugvó nyelvészeti elemzésekhez, másrészről pedig, éppen ennek a nagy mennyiségű adatnak köszönhetően megbízható és reprezentatív eredményeket produkálnak (Bieber et al.1994).

1 http://mnsz.nytud.hu/

Baker (1995) a korpuszok három típusát különbözteti meg: többnyelvű, párhuzamos és összehasonlítható. A többnyelvű korpusz két vagy több különböző nyelven írt, autentikus szövegek egynyelvű korpuszából áll. Használata lehetővé teszi, hogy egyes nyelvi elemeket és jellemzőket, mintázatokat saját természetes környezetükben vizsgáljunk, hozzájárulva a gépi fordítórendszerek teljesítményének növeléséhez, de kitűnő alapot jelenthet a kontrasztív nyelvészek munkájához is. A párhuzamos korpuszok az eredeti forrásnyelvi szöveget és annak fordítását tartalmazzák. Segítségükkel a nyelvpárspecifikus fordítói viselkedést vizsgálhatjuk, hasznuk a fordítóképzésben és a gépi fordítórendszerek kialakításában, fejlesztésében is megkérdőjelezhetetlen. Az összehasonlítható korpuszok tulajdonképpen a párhuzamos és a többnyelvű korpuszok között képeznek átmenetet. Két alkorpuszból állnak, az egyik fordításokat tartalmaz a célnyelven, míg a másik autentikus, eredetileg is a célnyelven született szövegekből épül fel. Az összehasonlítható korpuszok jelentősége elsősorban a fordított szövegek mintázatának vizsgálatában mutatkozik meg (Baker 1995: 234).

A korpusznyelvészet fejlődésével párhuzamosan a fordítástudományban is mérföldkőnek bizonyultak a 70-es évek, amikor is a sokáig uralkodó előíró szemléletet felváltotta a leíró szemlélet, amely már nem az eredeti szövegekhez méri a fordításokat, hanem önálló, kutatásra érdemes tárgyként tekint rájuk. Ez a nézőpont teljesen más fényben tünteti fel az ekvivalenciát. Toury (1980: 39) szerint, ha egy szöveg egy másik fordítása, feltételezhető, hogy a kettő között egyenértékűség van. Nem létezik abszolút ekvivalencia. A kutatás fókusza is megváltozik, és a kommunikációs normák felé fordul. Minden nyelvnek és kommunikációs helyzetnek megvannak a normái, így a fordított szövegnek az adott nyelven és kommunikációs helyzetben érvényes normáknak kell megfelelnie. Az érvényességet pedig a relevancia elve határozza meg: „a fordításnak az adott célközönség, a fordítás tervezett használója számára az adott szituációban optimálisan relevánsnak kell lennie (Heltai 2005: 167).”

2. A korpuszalapú fordítástudomány

A korpuszalapú fordítástudomány a 90-es évek végén új kutatási irányként jelent meg a fordítástudományon belül. Mona Baker már akkor előrevetítette, hogy az eredeti és a fordított szövegek korpuszának elérhetősége a fordított szövegek természetének, jellemzőinek feltárását teszi lehetővé a fordításkutatók számára. A korpuszalapú fordítástudomány három alapvető jellemzője a célorientált megközelítésekkel való elméleti kapcsolata, a sajátos módszertan, valamint az ebben a sajátos módszertanban rejlő egyedülálló lehetőség, amely segítségével a fordítás megkülönböztető nyelvi jellemzői, a fordítási univerzálék kutathatók (Baker 1996). A fordítási univerzálék a fordítás folyamatának elkerülhetetlen velejárói, hatásukra a célnyelven létrehozott produktum, a fordítási szöveg, „fordításízű" lesz: nem ellenkezik ugyan a célnyelv szabályaival, de eltér a célnyelvi szövegek normáitól.

Az első fordítási univerzálé, amelyet többen is megfogalmaztak, az egyértelműségre, a nagyobb explicitségre irányuló tendencia. A fordítástudományban először Blum-Kulka (1986) végzett empirikus kutatásokat az explicitációt illetően. Szövegelemzések segítségével tanulmányozta a fordítás során bekövetkező kohéziós és koherenciaeltolódásokat, majd a vizsgálatok eredményeképpen megalkotta az explicitációs hipotézist: „A forrásnyelvi szöveg értelmezése során a fordító olyan célnyelvi szöveget hozhat létre, amely redundánsabb, mint a forrásnyelvi szöveg. Ez a redundancia a célnyelvi szöveg kohézív explicitségének növekedésében mutatkozhat meg" (Blum-Kulka 1986: 19; a szerző saját fordítása). Baker megfogalmazása szerint a fordítás univerzális jellemzői tipikusan a fordított szövegekben lelhetők fel, és nem a nyelvrendszerek interferenciájának eredményeképpen jönnek létre (Baker 1993: 243). Tanulmányában négy univerzálét sorol fel: elsőként az explicitációt, azaz az explicitség szintjének jelentős emelkedését a forrásszöveghez és az eredetileg célnyelven született írásokhoz képest. Ezt követi az egyértelműsítés irányába mutató tendencia, vagyis az egyszerűsítés, amely leggyakrabban az átlagos mondathossz rövidülésével, szavak, kifejezések kihagyásával, módosításával, alacsonyabb lexikai sűrűséggel, alacsonyabb szótípus/szövegszó (*type/token*) aránnyal jár együtt. Baker harmadik univerzáléja a konzervativizmus (*conservatism*), amely tulajdon-

képpen nem más, mint a hagyományos nyelvhasználat előnyben részesítése, a célnyelvi sajátosságok túlzott hangsúlyozása. Végül a kiegyenlítődés (*levelling out, equalizing*), amely a fordítási szövegnek a kontinuum középpontja felé tartó mozgása. Ezen univerzálé értelmében az adott fordítási korpuszhoz tartozó szövegek jobban hasonlítanak egymásra a lexikai sűrűséget, szótípus/szövegszó arányt és az átlagos mondathosszúságot illetően, mint az eredeti szövegek.

Laviosa (2009) az *Encyclopedia of Translation Studies* fordítási univerzálékról szóló szócikkében hat univerzálét említ: az egyszerűsítést, az explicitációt, a normalizációt (Baker 1993, konzervativizmus), az ismétléskerülést (*avoidance of repetitions*), a forrásnyelvi szövegtulajdonságok átvételét (*discourse transfer*), amely Toury (1995) interferenciáról szóló törvényével egyenértékű, és a célnyelvi lexikai egységek sajátos megoszlását (*distinctive distribution of target-language items*). Az enciklopédia későbbi kiadásában már szerepel Toury (1995) növekvő standardizációról megfogalmazott törvénye is, ennek értelmében a fordítások nyelvi szempontból kevésbé változatosak, mint az eredetileg célnyelven írt szövegek.

Chesterman 2004-es tanulmányában az időközben felmerülő fogalmi nehézségek tisztázására irányuló kísérletként két típusú univerzálét különböztet meg: a forrásnyelvi szövegekhez képest eltérést jelentő F-univerzálékat (*S-universals*) és a fordítások és az autentikus, célnyelven írt szövegek közötti különbségekben megmutatkozó C-univerzálékat (*T-universals*). Az előbbi csoportba tartozik a szövegek terjedelmének növekedése, az interferencia törvénye, a növekvő standardizáció törvénye, az explicitációs hipotézis és az ismétlések kerülése, míg az utóbbiak között szerepel az egyszerűsítés, a konvencionalizáció, a szokatlan lexikai mintázatok, a célnyelvi egyedi elemek alulreprezentáltsága.

3. A korpuszalapú tolmácsolástudomány

A korpuszalapú tolmácsolástudomány a korpuszalapú fordítástudománnyal szemben több éves, sőt évtizedes lemaradást mondhat magáénak, amelynek oka elsősorban az adatokhoz való hozzáférés nehézségében és az átírás időigényességében keresendő. 1965 és 2000

között mindössze kilenc tolmácsolási korpusz született (Oléron és Nanpon 1965; Dejean Le Féal 1978; Chernov 1978; Lederer 1981; Shlesinger 1998; Pöchhacker 1994; Kalina 1998; Setton 1997; 1999), és csak 1998-ban fogalmazta meg Shlesinger a gondolatot, mely szerint érdemes lenne a korpuszalapú módszertant a tolmácsolás vizsgálatára is kiterjeszteni, ezáltal létrehozva egy új kutatási irányt a fordítástudományon belül. A 2000-es évek technikai fejlődése elősegítette a felvetés megvalósítását, és ugrásszerűen megnőtt a tolmácsolási korpuszok száma. Hozzá kell ugyanakkor tenni, hogy ezen korpuszok az esetek többségében disszertációkhoz készültek (Dirkier 2001; Vuorikoski 2004; Monacelli 2005), vagy saját felhasználásra hozták létre őket, azaz nagy nyilvánosság számára nem elérhetőek (Cencini 2000; Fumagalli 1999–2000).

A jelenleg elérhető legjelentősebb tolmácsolási korpuszok: a Center for Integrated Acoustic Information Research of Nagoya University 1999 és 2003 között épített, közel egymillió szövegszót tartalmazó japán–angol, angol–japán CIAIR szinkrontolmácsolási korpusza (Tohyama et al. 2006); az összesen 117 295 szövegszót tartalmazó háromnyelvű (olasz, angol, spanyol) EPIC (*European Parliament Interpreting Corpus*) (Bendazzoli és Sandrelli 2005); a 36 000 szövegszót számláló TIC (*Television Interpreting Corpus*) (Cencini 2000); a 2249i korpusz (Russo et al. 2018); a konferenciatolmácsolás direkcionalitásának vizsgálatára létrehozott DIRSI korpusz (Bendazzoli és Sandrelli 2009); a professzionális tolmácsok metaforák tolmácsolására alkalmazott stratégiáinak elemzésére készült IMITES (*Interpretación de la Metáfora entre ITaliano y ESpanol*) (Spinolo 2018); a világ legnagyobb televíziós tolmácsolási korpusza, a CorIT (Falbo 2012); a 2008-as futballvilágbajnokság sajtókonferenciáinak átiratát tartalmazó FOOTIE korpusz (Sandrelli 2012); a japán Nara Institute of Science and Technologie által épített angol–japán, japán–angol NAIST korpusz (Neubig et al. 2018) valamint a konszekutív tolmácsolásokat tartalmazó CEIPPC (Chinese–*English Interpreting for Premier Press Conferences Corpus*) (Wang 2012).

A fordítás- és tolmácsolástudomány tehát hosszú ideig külön utakon fejlődött, mindkét tudományág a maga párhuzamos és összehasonlítható korpuszaival igyekezett feltárni a fordítási és tolmácsolási szövegre jellemző sajátos tulajdonságokat. 2004-ben

viszont Schäffner *Translation Research versus Interpreting Research: Kinship, differences and prospects for partnership* című cikkében felvetette, hogy a fordítás és a tolmácsolás különbségeinek ellenére a közös alapok, a mindkét területet foglalkoztató, hasonló ismeretelméleti, módszertani, szélesebb intézményi és társadalomtudományi vonatkozások miatt a két tudományág együttműködéséből számos előnye származna a kutatótársadalomnak (Gile 2004). Shlesinger (2009) Schäffner gondolataival összhangban elsőként javasolta a tolmácsolt és eredeti hangzó szövegek egynyelvű összehasonlítható korpuszának fordított szövegekkel való kiegészítését, azaz egy intermodális korpusz létrehozását, majd ezt a gyakorlatban is megvalósította.

> „[…] a fordítástudósok sokat tanulhatnak az (írott) fordítás folyamatáról és magáról a fordított szövegről azáltal, hogy többet megtudnak a tolmácsolásról – és a tolmácsolástudósok is levonhatnak következtetéseket a fordításnak erről a nagy nyomás alatt zajló formájáról, miközben a lassabb, jobban megfigyelhető (írott) fordítás folyamatát és eredményét vizsgálják" (Shlesinger és Ordan 2012: 44; a szerző saját fordítása).

4. Intermodális korpuszok és intermodális korpuszokon végzett univerzálé kutatások

Shlesinger 2008-as kutatásához kísérleti körülmények között gyűjtött anyagot használt. Hat profi tolmács-fordítót kért fel arra, hogy előbb tolmácsolják, majd (3 évvel később) fordítsák le ugyanazt az angol forrásnyelvi szöveget anyanyelvükre, héberre. Az intermodális korpusz, amelyen a kutatást végezte, megközelítőleg 10 000 szövegszót tartalmazott. A kísérlet egy sor fontos különbséget fedett fel a két modalitás szövegprodukciója között, elsősorban a lexikai változatosság, valamint a lexiko-grammatikai, stiláris és pragmatikai jellemzők tekintetében, amelyek mind a tolmácsolás szóbeliségét tükrözik. A szótípus/szövegszó arány (*TTR*), azaz a lexikai változatosság kulcsmutatója tekintetében azt találták, hogy a vizsgálatban részt vevő mind a 6 adatszolgáltatónál alacsonyabb volt a tolmácsolt, mint a fordított outputban.

A beszédrészek eloszlása szintén jelentős eltérést mutatott: a névmások száma jóval magasabb volt szinkrontolmácsolás során, mint fordításban, ami Shlesinger (2008) szerint a héberben az oralitás egyértelmű jele. A lexikai elemek választása szintén illeszkedett ebbe a szóbeliséget tükröző tendenciába. Tolmácsolás során a vizsgálati személyek az informális, kollokviális kifejezéseket részesítették előnyben, valamint jóval gyakrabban fordult elő, hogy a célnyelven is ismert és használt idegen szavak eredeti formájukban kerültek bele a tolmácsolt szövegbe, míg fordításnál inkább az adott szó standard célnyelvi megfelelőjét igyekeztek megkeresni.

Bár a 2012-es kutatás, amelyet Schlesinger Ordannal közösen végzett, ugyanazokra a kérdésekre kereste a választ, mint a négy évvel korábbi, két fontos különbséget mégis meg kell említeni. Az egyik, hogy míg a korábbi egy tudományos kísérlet eredményeképpen született, a 2012-es autentikus, konferenciákon rögzített adatokat használt, valamint hogy a kutatópáros kibővítette a vizsgálódás kereteit egy újabb aspektussal. Ezúttal nemcsak írott fordításokkal, hanem eredetileg héber nyelven elhangzott spontán beszédekkel is egybevetették a tolmácsolt szövegprodukciót. Ennek megfelelően a korpusz három alkorpuszból állt, amelyek mindegyike 24 000 szót tartalmazott.

A jelen cikkben csak a két modalitásra, azaz a fordításra és a tolmácsolásra vonatkozó eredményeiket ismertetem. A beszédrészeket tekintve a kutatók vizsgálták a névmások és a határozószók számát. Mindkettő jóval gyakrabban fordult elő a tolmácsolt (névmások: 1524; határozószók: 1401), mint a fordított szövegekben (névmások: 659; határozószók: 851). A korábbi kutatás nem vizsgálta sem a tulajdonneveket, sem az indulatszavakat, ezúttal azonban ezen a téren is meggyőző eredmények születtek. A tulajdonnevek 73 százalékkal kevesebb alkalommal fordultak elő tolmácsolásban, mint az írott fordításban. Ez összhangban van a névmások gyakoribb használatával a szinkrontolmácsolás során, hiszen a tulajdonneveket általában névmásokkal váltják ki (Meyer 2008). Az indulatszavak 383 százalékkal jelentek meg gyakrabban a tolmácsolt szövegekben, mint az írottban. Az eredmények a tolmácsolás beszélt nyelvi jellegét bizonyítják.

A CECIC (*Chinese–English Conference Interpreting Corpus*), vagyis a kínai–angol konferenciatolmácsolási korpusz is azzal a céllal jött létre, hogy segítségével tanul-

mányozhassák a tolmácsolási szöveg jellemzőit, a tolmácsolási normákat és a tolmácsolás kognitív folyamatait. A korpusz három, összesen 500 000 szövegszót tartalmazó alkorpuszból áll: sajtókonferenciák angol nyelvű korpusza, sajtókonferenciák tolmácsolásának kínai–angol párhuzamos korpusza, a kínai kormány munkajelentéseinek kínai–angol párhuzamos korpusza. A korpuszon Hu és Tao (2013) vizsgálták a fordítás két univerzális jellemzőjének, a normalizációnak és az explicitációnak a megjelenését a tolmácsolt szövegekben. A kutatók egy tipikus angol nyelvtani szerkezet, a passzív megfogalmazás tolmácsolt, fordított és eredeti angol nyelven írt szövegekben való előfordulásának statisztikai elemzésével igyekeztek megállapítani, hogy jellemző-e a normalizáció a tolmácsolt szövegekre. Az explicitáció jelenlétét pedig az opcionális *that* kötőszó és a *to* főnévi igenevet bevezető partikula használatának elemzésével próbálták igazolni. Az opcionális *that* kötőszót a főnévi és névmási mellékmondatok bevezetésére használjuk. Ugyanakkor a tárgyként vagy névmási mellékmondatként funkcionáló főnévi mellékmondat esetén gyakran elmarad. A *to* partikula főnévi igeneves szerkezeteket vezet be, két vagy több főnévi igenév esetén azonban előfordul, hogy csak az első főnévi igenév előtt jelenik meg. Ez esetben is kizárólag a fordított és tolmácsolt szövegekre vonatkozó adatokra szorítkoznék az eredmények ismertetésekor, amelyek világosan mutatják, hogy az angol passzív szerkezet (1.66-szor gyakrabban), valamint az opcionális *that* (háromszor gyakrabban) és a *to* partikula (3,8-szor gyakrabban) is jóval nagyobb gyakorisággal fordult elő a tolmácsolt, mint a fordított szövegekben, következésképpen a tolmácsolt szövegek jelentősebb tendenciát mutattak a normalizációra és az explicitációra, mint a fordítások. A kutatók azzal magyarázzák a gyakori előfordulásokat, hogy egyrészről a tolmácsok szeretik kihangsúlyozni az új információkat, amire leginkább a szenvedő szerkezet alkalmas, másrészről, hogy a tolmácsolásban a *that* kötőelemet és a *to* főnévi igenevet bevezető partikulát leginkább a hallgatóság szövegfeldolgozását megkönnyítő diskurzusjelölőként használják.

He, Boyd-Graber és Daumé (2016) japán–angol fordított és szinkrontolmácsolt szövegek korpuszán végeztek vizsgálatokat. A fordított szövegeket saját maguk gyűjtötték, a tolmácsolt szövegeket a CIAIR (*Center for Integrated Acoustic Information Research*) által épített szinkrontolmácsolási korpusz biztosította. Előzetes hipotézisük szerint a tolmácsolási stratégiákat két nem kizárólagos kategóriába sorolták. Az elsőbe a késede-

lem (*delay minimization*), a másikba a memóriaigény minimalizálására irányuló stratégiák (*memory footprint minimization*) tartoztak. A kísérlet eredményei szerint a tolmácsok által leggyakrabban használt stratégiának a szegmentálás, a passzivizálás, az általánosítás és az összefoglalás bizonyultak. A vizsgálati személyek az esetek 73 százalékában szegmentáltak. Ennek során a forrásnyelvi mondatokat több rövidebb részre bontották, és megpróbálták az új információt az előző részleges fordításba beépíteni a *which is* vagy egy kötőszó használatával. A passzivizálást nemcsak a normalizálásra való törekvés motiválhatja, de hasznos stratégia lehet a *head-final* nyelvekről *head-initial* nyelvekre történő tolmácsoláskor. Ahelyett hogy a tolmács megvárná, míg a forrásnyelv egyik központi eleme a megnyilatkozás végén megjelenik, passzív szerkezetté alakítva a célnyelvi mondatot szinte késedelem nélkül követheti a beszélőt. Ennek ellenére a kísérlet csak az esetek 57,1 százalékában igazolta a passzív szerkezet használatát. Oka minden bizonnyal az, hogy a passzivizálás tanult képesség, vagyis nem mindenki képes nehézségek nélkül alkalmazni. A kísérletet végzők azt is megfigyelték, hogy a tolmácsok inkább a mondat lényegére összpontosítanak, és ennek kifejezésére az esetek többségében inkább gyakori szavakat használnak, mert ezek lehívása a rövid távú memóriából kevesebb időt vesz igénybe. A tolmácsoknak óriási információmennyiség kezelésével kell megküzdeniük munkájuk során, ezért hatékonyan kell a jelentést értelmezniük és átültetniük. Ez indokolja, hogy gyakran összefoglalnak, kevésbé fontos szavakat vagy esetleg egész mondatokat kihagyhatnak, azaz egyszerűsítenek elsősorban akkor, amikor jelentősen lemaradnak a beszélő mögött. Ez utóbbi két stratégiát is gyakran alkalmazták a tolmácsok, annak ellenére, hogy nem a tolmácsolási feladat elválaszthatatlan velejárói, inkább a korlátozott rövid távú memória melléktermékei.

Az egyszerűsítést Silvia Bernardini, Adriano Ferraresi, Maja Milicevic (2016) vizsgálták mediált (tolmácsolt és fordított) szövegekben a kétirányú (angol, olasz) intermodális EPTIC korpusz (*European Parliament Translation and Interpreting Corpus*) segítségével. Laviosa-Braithwaite (1998) fordított szövegekre alkalmazott módszertanának segítségével elemezték a szövegeket. Négy paraméter mentén haladtak: lexikai sűrűség, átlag mondathossz, alapszókincs és a szövegekben leggyakrabban előforduló szavak (*list head*). Megállapították, hogy az EPTIC-ben található tolmácsolt szövegek következetesen

egyszerűbbek, mint fordított megfelelőik. Az egyszerűsítés paramétereit azonban a vizsgálat eredményei szerint különböző módon alkalmazzák a nyelvek.

1. táblázat

Intermodális korpuszokon végzett kutatások eredményei

	Kutató neve/korpusz neve	**Korpusz nyelve, mérete**	**Eredmények a tolmácsolási nyelvre vonatkozóan**
1.	Shlesinger, saját korpusz (2008)	angol-héber, kb. 10 000 szövegszó	**TTR**: alacsonyabb a tolmácsolt szövegben > **oralitás** **Beszédrészek eloszlása**: névmások száma magasabb tolmácsolt vs. fordított szövegben > **oralitás** **Lexikai elemek választása**: informális, kollokviális kifejezések, idegen szavak gyakoribb használata tolmácsolt vs. fordított szövegben > **oralitás**
2.	Shlesinger és Ordan, saját korpusz (2012)	angol-héber, 72 000 (48 000) szövegszó	**Beszédrészek eloszlása**: névmások száma magasabb tolmácsolt vs. fordított szövegben > **oralitás** **Tulajdonnevek száma:** kevesebb tolmácsolt vs. fordított szövegben
3.	Hu és Tao, CECIC (2013)	kínai, angol, 500 000 szövegszó	**Passzív szerkezetek száma**: magasabb tolmácsolt vs. fordított szövegben > **normalizáció** **Opcionális „that" és „to" száma:** magasabb tolmácsolt vs. fordított szövegben magasabb tolmácsolt vs. fordított szövegben > **explicitáció**
4.	Bernardini, Ferraresi és Milicevic, EPTIC (2016)	angol, olasz, spanyol, 180 000 szövegszó	**egyszerűsítés**
5.	He, Boyd-Graber és Daumé, CI-AIR (2016)	japán, angol, kb. 1 millió szövegszó	**szegmentálás, passzivizálás, általánosítás, összefoglalás**

Míg az angolra tolmácsolók inkább támaszkodnak pusztán lexikai forrásokra (szívesebben használnak gyakori szavakat és szövegen belüli ismétléseket), addig az olaszra tolmácsolók szintaktikai eszközökkel is élnek (rövidebb mondatok és több funkciószó) egyszerűsítés esetén. Az áttekintést megkönnyítendő, a fenti, *1. táblázatban* foglalom össze a fentiekben részletesen ismertetett szakirodalom eddigi, intermodális korpuszokon végzett kutatásokkal alátámasztott eredményeit a tolmácsolt szövegekre vonatkozóan.

5. A kutatás bemutatása

Kutatásom célja annak feltárása, hogy létezik-e tolmácsolási szöveg, azaz olyan, a tolmácsolási folyamat eredményeképp létrejövő, sajátos tulajdonságokkal rendelkező szövegtípus, amely megfelel ugyan a célnyelv szabályainak, mégis – hol jobban, hol kevésbé – érezhető rajta, hogy eredetileg nem ezen a nyelven született, azaz eltér a célnyelvre jellemző normától. A 2017-ben megkezdett kutatási projekt első szakaszában az eredeti (angol és magyar) hangzó szövegek és azok tolmácsolt változatának átiratát vetettem egybe. A számadatok egyértelműen kimutatták, hogy a tolmácsolt szövegek egyszerűbbek, implicitebbek és kevésbé redundánsak, mint a forrásnyelvi hangzó változat. A kvalitatív kutatás mindezt nemcsak megerősítette, de egyúttal magyarázatként is szolgált: a háttérben a tolmácsolás során megjelenő időkorlát, a tolmács azon törekvése, hogy ideális távolságot tudjon tartani a beszélő mögött, ne maradjon le túlságosan, valamint a memória telítettségének elkerülésére irányuló szándéka áll.

A második szakaszban intermodális elemzést végeztem, amelynek során ugyanazon hangzó szövegek fordított és tolmácsolt verzióját vetettem egybe. A teljesség és a korábbi kutatások eredményeivel való összehasonlíthatóság érdekében ezúttal is végeztem kvantitatív és kvalitatív elemzést is. Előbb tehát a szövegek mondatainak, szavainak számát, az átlagos mondathosszúságot, a névmások számát vizsgáltam, majd az explicitációval és implicitációval járó fordítási műveletekre koncentráló kontrasztív elemzést végeztem. Az eredményektől azt vártam, hogy megerősítsék előzetes hipotézisemet, miszerint a tolmácsolt szövegek nemcsak a forrásnyelvi eredetihez, de a fordításokhoz képest is egyszerűbbek, illetve hogy természetük sokkal inkább beszéltnyelvi, mint fordítási sajátosságokkal rendelkezik.

5.1. A korpusz

A kutatás második szakaszának alapját a jelenleg 15 millió szövegszót tartalmazó Pannónia Korpusz (Robin et al. 2016) tolmácsolási alkorpuszának öt-öt eredeti hangzó szövege (angol és magyar), valamint azok fordított változata (magyarról angolra és angolról

magyarra) és tolmácsolt átirata (magyarról angolra és angolról magyarra) adta. A hangzó szövegek EP-képviselők autentikus felszólalásai, míg a valós körülmények között zajló tolmácsolást az Európai Parlament tolmácsai biztosítják. Az Európai Parlament hivatalos munkahelyein minden üléstermet a nemzetközi szabványoknak megfelelő tolmácsfülkékkel szereltek fel: műszaki problémák esetén konferencia technikusok állnak a tolmácsok rendelkezésére. A fordítások a hangzó szövegek standardizált változatának fordításai, ezeket ugyancsak az EP hivatalos fordítói biztosították, akiknek a munkakörülményeit is hasonló, professzionális és mindenkinek ugyanazokat a lehetőségeket nyújtó eszköz- és háttérrendszer jellemzi.

Az adatgyűjtés módszere nem volt interaktív, mivel az adatszolgáltatók nem játszottak aktív szerepet az adatgyűjtésben, illetve az adatok elemzésében és értelmezésében (Gile 1998: 71), mint ahogyan nincsenek tudatában annak sem, hogy beszédprodukciójukat kutatási céllal elemzéseknek vetjük alá – a jelen vizsgálat alapjául szolgáló szövegek az Európai Parlament nyilvánosan elérhető archívumából kerültek ki. A felszólalások mindegyike 2000 után hangzott el, hosszúságuk hasonló (körülbelül 1 perc). Annak köszönhetően, hogy minden forrásnyelvi szöveget más parlamenti képviselő ad elő, és más tolmács alakít át célnyelvi szöveggé, kizárható az eredményeket esetlegesen befolyásoló egyéni preferenciák és stílus hatása.

5.2. Kutatási kérdések és előzetes hipotézisek

A kutatás egyik fő kérdése, hogy léteznek-e a fordítási univerzálék mintájára tolmácsolási univerzálék, azaz minden tolmácsolt szövegre jellemző közös tulajdonságok. Mivel mind a fordított, mind a tolmácsolt szövegprodukció egy forrásnyelv és egy célnyelv között zajló közvetítési folyamat eredménye, feltételeztem, hogy lesznek olyan univerzálék, amelyek mindkettőre jellemzőek. Az a tény azonban, hogy az egyik írott, a másik szóbeli folyamat eredménye, úgy gondolom, hogy szükségszerűen eltéréseket okoz a két szövegtípus között. Feltételezésem szerint a szóbeliséggel járó korlátok, mint az időkorlát vagy a rövid távú memória limitált tárolási kapacitása, nagyban befolyásolhatja a mind-

két szövegre jellemző közös tulajdonságok megjelenését és arányát, például az eddigi kutatások szerint a kihagyások, a szegmentálás (He et al. 2016) és az egyszerűsítés (Shlesinger 2008; Bernardini et al. 2016) előfordulása jóval gyakoribb a tolmácsolt, mint a fordított szövegekben. A kihagyásokra további magyarázatokkal szolgálhat a fordítási és a tolmácsolási helyzet különbsége: tolmácsoláskor a kommunikációs kontextus azonnal a befogadó rendelkezésére áll, így az ebből kikövetkeztethető információk különösebb gond nélkül elhagyhatók. Ha ez mégis problémát okoz, a hallgatónak lehetősége van a feldolgozáshoz szükséges részleteket megszerezni. A fordításnál mindez késleltetetten jelentkezik. A fordított szöveg olvasója egészen más kontextusban találkozik a szöveggel, mint a fordító, így az ő megértésének elősegítése érdekében jóval kevesebb részlet hagyható el, mint a tolmácsolás során.

A kutatás másik fő kérdése, hogy az ontológia (tolmácsolt vagy autentikus) vagy a modalitás (beszélt vagy írott) gyakorol-e erősebb hatást a tolmácsolt szövegekre. Hipotézisem szerint, amelyet elsősorban Shlesinger korábbi kutatásának (2008) eredményeire alapozok, ez utóbbi játszik fontosabb szerepet. A beszélt nyelv alapvető jellemzői a megjelenési formában, a spontaneitásban, a kötetlenségben és egyes beszédrészek használatában mutatkoznak meg. Ami a megjelenési formát illeti, ahogy azt a neve is mutatja, a beszélt nyelv elsősorban hangzó formában valósul meg. Az írott nyelvi szövegeket is fel lehet olvasni, de ez csupán másodlagos megvalósulás. A spontaneitás azt jelenti, hogy a beszéd legtöbbször mindenféle előkészítés nélkül történik, azaz a mondanivalón való gondolkodás, vagy másképp a konceptuális tervezés, a gondolatok megfogalmazása, és az artikuláció nagyjából szinte teljesen egy időben történik. A kötetlenség pedig abban nyilvánul meg, „hogy a beszélő a beszédtevékenység során nem támaszkodik korábban leírt vagy hallott szövegekre, vagyis a beszélt nyelv szövegfüggetlen szövegalkotás révén jön létre” (Lanstyák 2009: 15).

A beszélt nyelv múlékony, egyszeri, visszavonhatatlan. Bár javításra, módosításra van lehetőség, az elhangzottakat nem tudjuk kitörölni. A beszélt nyelv továbbá egy adott beszédhelyzetbe ágyazott, vagyis a beszélő a beszédhelyzetre támaszkodva nem fejezi ki verbálisan mondanivalója egyes részleteit. Hasonlóképpen a szupraszegmentális esz-

közökkel (hangszínnel, hanglejtéssel…) kifejezhető, illetve a testbeszéddel, mimikával, gesztusokkal pótolható tartalmak is elmaradhatnak. Az információ visszakeresése hangzó szöveg esetén nehézkes, ilyenkor csak a memóriára támaszkodhatunk. A megjelenési formát, spontaneitást, kötetlenséget és kontextus függőséget tekintve tehát kétség nem fér hozzá, hogy a tolmácsolt szöveg a hangzó beszéd jellemzőit hordozza magán. A további tulajdonságait (beszédrészek, redundancia) a jelen kutatás keretei között vizsgálom.

5.3. A kutatás módszerei

Az érvényes és megbízható következtetések érdekében kétféle elemzésnek vetettem alá a kutatás alapjául szolgáló szövegeket. Az első lépésben kvantitatív jellegű, statisztikai elemzést végeztem, amelynek során a tolmácsolt és a fordított szövegek terjedelmét, mondataik számát és az átlagos mondathosszúságot, a névmások számát és a szótípus/szövegszó arányt (*TTR*) vizsgáltam. Mindezt a Microsoft Office szószámlálójának és a Wordsmith Tools 6.0[2] verziójának segítségével végeztem.

A statisztikai vizsgálatokat kontrasztív szövegelemzés követte. Ennek során Klaudy (1997) tipológiáját alkalmazva először is megállapítottam, milyen átváltási műveleteket végeztek a tolmácsok (lexikai konkretizáció, generalizáció, felbontás, kihagyás, betoldás, áthelyezés, csere, teljes átalakítás és kompenzálás; grammatikai konkretizáció, generalizáció, felemelés, lesüllyesztés, kihagyás, betoldás, felbontás, összevonás, áthelyezés és csere), majd az azonosított műveleteket Klaudy (1999) alapján explicitációs és implicitációs kategóriákba soroltam. A grammatikai és lexikai áthelyezések és cserék esetén a két kategória egyikébe történő besorolás alapját az adta, hogy a műveletek eredményeképpen vajon „hangsúlyosabb, egyértelműbb, kifejtettebb közlés jött létre, vagy ennek az ellenkezője (Robin 2018: 99). Végezetül az explicitációval és implicitációval járó átváltási műveleteket Robin (2018) alapján további három kategóriába soroltam: kötelező szabálykövető, fakultatív normakövető, fakultatív stratégiakövető explicitációs és implicitációs műveletek. A kötelező kategóriába azok a műveletek tartoznak, amelyeket a két

2: https://www.lexically.net/wordsmith/

nyelv rendszerbeli különbségei miatt kell elvégezni, és amelyek figyelmen kívül hagyása nyelvtanilag helytelen és értelmetlen mondatot eredményezne. Ezt a kategóriát pontosan kötelező jellege miatt hagytam ki a vizsgálatból, ezek a műveletek ugyanis nem befolyásolják a tolmácsolási szöveg mint önálló szövegtípus jellemzőit. Kizárólag a másik két kategóriára fókuszáltam: a fakultatív jellegű normakövető műveletekre, vagyis azokra a beavatkozásokra, amelyek olyan gyakran fordulnak elő, hogy normaként értelmezhetőek, valamint a stratégiakövető műveletekre, amelyek segítségével olyan jelentés fogalmazódik meg expliciten a célnyelvi szövegprodukcióban, amely a forrásnyelvi eredetiben nem volt benne, vagy impliciten került megfogalmazásra, esetleg kevésbé volt hangsúlyos. Az elemzések megbízhatóságát kettős kódolással (Károly 2002) biztosítottam, egyik kollégám segítségét kérve, aki az előre egyeztetett tipológia alapján ugyancsak elvégezte a szövegek elemzését, az egyeztetés során pedig feloldottuk az eltéréseket.

6. Eredmények: statisztikai elemzés

A következő fejezetben az intermodális korpuszon végzett statisztikai elemzést mutatom be.

6.1. Statisztikai elemzés: szavak száma

Az intermodális korpuszon végzett statisztikai elemzés első lépéseként a szavak számát vizsgáltam. Az eredményeket a *2. táblázat* foglalja össze.

A *2. táblázat* első számadatai az öt-öt szöveg együttes szószámát jelzik, míg a zárójelben felsoroltak szövegenként mutatják az adott szövegek szavainak számát. Az eredmények alapján a következő megállapításokat tehetjük.

Az eredeti írásbeli megnyilatkozások összességében és többségükben rövidebbek, mint a szóbeliek, hiszen utóbbiak alapjában véve több ismétlést, redundáns elemet tartalmaznak.

2. táblázat

Szavak száma a vizsgált szövegekben

Eredeti angol írott	813 (154; 159; 172; 160; 168)	**Eredeti angol hangzó**	828 (159; 160; 178; 157; 174)
Angolról magyarra fordított	722 (132; 147; 162; 137; 144)	**Angolról magyarra tolmácsolt**	572 (87; 126; 117; 120; 122)
Eredeti magyar írott	632 (105; 132; 127; 145; 123)	**Eredeti magyar hangzó**	666 (108; 145; 140; 149; 124)
Magyarról angolra fordított	876 (136; 165; 176; 200; 199)	**Magyarról angolra tolmácsolt**	835 (150; 162; 238; 151; 134)

A tolmácsolt és fordított szövegek esetében viszont, legyen szó bármely nyelvi irányról, ennek éppen a fordítottja igaz. Hiába szóbeli megnyilatkozás a tolmácsolt szöveg, mégis az esetek többségében redukáltabb a fordítottnál. Az ok egyrészről, hogy a tolmács folyamatosan az idő nyomása alatt dolgozik, ügyelnie kell arra, nehogy túlságosan lemaradjon a beszélő mögött, másrészről memóriája kapacitása is korlátozott (nincs lehetőség a szöveg újraolvasására), így igyekszik elkerülni annak túltelítettségét. Mindkét esetben fennáll az információvesztés kockázata, ezért a tolmács egyrészt feladata megkönnyítése érdekében, másrészt pragmatikai okokból (Pym 2008) (a koherens, érthetőbb szövegprodukció érdekében) egyszerűsít, vagyis az üzenet szempontjából nem releváns, redundáns vagy a kontextusból kikövetkeztethető elemeket elhagyja, esetleg összevon, azaz a több hosszabb mondatban elhangzott információt egy rövidebb mondatban, kevesebb szóval foglalja össze. A pragmatikai kihagyások anélkül valósíthatók meg, hogy veszélyeztetnék a kommunikációs aktus alapvető céljait (Pym 2008). Az alábbi példában a szövegkörnyezetből kikövetkeztethető információ elhagyását láthatjuk:

(1) Eredeti: „And in my view *these reports before the plenary today***,** fail to address the fundamental problems with the current system of structural funding.”

(1a) Fordított: „Véleményem szerint a *mai plenáris ülés elé került jelentések* nem foglalkoznak a strukturális finanszírozás jelenlegi rendszerének alapvető problémáival."

(1b) Tolmácsolt: „Szerintem *a jelentések* nem válaszolnak a jelenlegi strukturális ö re [alapok] ö rendszerének ö alapvető problémáira."

Gile (1995) szerint léteznek olyan kommunikációs helyzetek, ahol a külső körülmények mentális telítettséget okoznak (előadó gyors beszédtempója, információsűrűség, a beszélő erős akcentusa, esetleg helytelen nyelvtani- és szóhasználata), így akadályozzák a tolmácsot abban, hogy teljes egészében visszaadja az elhangzott információt. Ha azonban a tolmács a mentális telítettségi szint alatt dolgozik, akkor az indokolatlan kihagyások és hibák egyetlen oka a forrásnyelvi szöveg komplexitása lehet. Az információvesztéssel járó kihagyásra az alábbi mondatokban láthatunk példát:

(2) Eredeti**:** „Gond ez ma Európában? Hasonlóan beteges az a gondolkodás is, amely megtiltaná a szomszédnak, hogy megemlékezzen a napról, amikor elvesztette vagyonát, földjét, esetleg családtagjait. Erre is volt példa a napokban, amikor az első világháborút lezáró trianoni békeszerződésre emlékeztünk június 4-én. A kirekesztő, elnémító, múltat letagadó, egy régió létét, fejlődését megtiltani akaró gondolkodásra van ellenszer: egy nyitott, befogadó, kulturális értékeket megbecsülő, kisebbségek jogait tiszteletben tartó, erős Európa."

(2a) Fordított: *„Is this a problem in Europe today? The way of thinking that would forbid a neighbour to commemorate the day when he had lost his fortune, land, or perhaps even his relatives, is just as morbid.* We had an example of that in the past days as well, when we commemorated the Trianon Peace Treaty ending World War I on 4 June. *There is a remedy for the way of thinking that excludes and shuts up people, denies the past, and wants to impede*

*the development and existence of a region***:** a *strong* Europe that is *open* and *inclusive*, that appreciates cultural values, and respects the rights of minorities."

(2b) Tolmácsolt: „…There have been other problems in the last few days. We've [we] organised on June 4th we organised the commemoration of the Versailles treaty which put an end to the first world war … But ehm we have problems and there are means of tackling them but we must respect the values of all minorities and everyone's cultural values. Thank you."

A *2. táblázatból* az is kiolvasható, hogy a magyarra fordított és tolmácsolt szövegek rendre kevesebb szót tartalmaznak, mint az angolra fordított és tolmácsolt szövegprodukciók. Ennek hátterében két magyarázat is állhat. Az egyik, hogy anyanyelvre történő fordításkor és tolmácsoláskor a nyelvi közvetítők magabiztosabban alkalmazzák a kihagyást és összevonást, mint idegen nyelvre történő fordítás és tolmácsolás esetében. Nem szabad megfeledkeznünk a két nyelv tipológiai különbségeiről sem, amelyek szintén befolyásolják a szövegek terjedelmét. A magyar agglutináló nyelv, nyelvtani viszonyait a szótőhöz illesztett toldalékokkal fejezi ki, míg az angol izoláló, vagyis minden funkciót külön szó ábrázol. Ennek megfelelően a magyarra tolmácsolt és fordított szövegek szavainak száma az esetek nagy százalékában kevesebb, mint az angolra tolmácsolt és fordított szövegek szavainak száma.

6.2. Statisztikai elemzés: mondatok száma

A 3. táblázat első számadatai az öt-öt szöveg együttes mondatszámát jelzik, míg a zárójelben felsoroltak szövegenként mutatják az adott szöveg mondatainak számát. Az eredmények alapján a következő megállapításokat tehetjük. Az eredeti írásbeli szövegek összességében és többségében is kevesebb mondatot tartalmaznak, mint a szóbeli megnyilatkozások; egyenként vizsgálva a szövegeket a mondatok száma közel azonos vagy kevesebb az írott szövegekben, csupán két kivétellel.

3. táblázat

Mondatok száma a vizsgált szövegekben

Eredeti angol írott	33 (8; 8; 7; 6; 4)	**Eredeti angol hangzó**	42 (9; 8; 7; 9; 9)
Angolról magyarra fordított	37 (9; 9; 7; 6; 6)	**Angolról magyarra tolmácsolt**	44 (8; 10; 8; 6; 12)
Eredeti magyar írott	35 (7; 6; 8; 6; 8)	**Eredeti magyar hangzó**	41 (8; 8; 8; 7; 10)
Magyarról angolra fordított	33 (7; 6; 7; 6; 7)	**Magyarról angolra tolmácsolt**	36 (5; 8; 11; 5; 7)

Ennek magyarázata szintén lehet a szavak számánál hivatkozott indok, miszerint a szóbeli megnyilatkozások több ismétlést, redundáns elemet tartalmaznak. Ugyanakkor meg kell jegyeznünk, hogy a hosszú, összetett mondatok nagyobb feldolgozási erőfeszítést igényelnek, ezért a szóbeli megnyilatkozások, amikor a befogadónak azonnal fel kell dolgoznia a hallottakat, több mondatot tartalmazhatnak.

A tolmácsolt és fordított szövegeknél ugyanez látható. A tolmácsolt szövegek mindkét nyelvi irányban összességében és többségében több mondatot tartalmaznak, vagyis az adott információt a tolmácsok kevesebb szóval (lásd *1. táblázat*), ugyanakkor több mondatban fogalmazzák meg, mint fordító társaik. Ilyenkor az írásban gyakran előforduló hosszú körmondatokat a könnyebb érthetőség és feldolgozás érdekében a tolmács több mondatra bontja fel, vagyis szegmentál, ahogyan az alábbi példa is mutatja.

(3) Eredeti: „The WikiLeaks files have now given an even clearer picture of the grossest abuses of human rights that have taken place there – the widespread use of waterboarding, other forms of brutal torture to extract confessions and information – and they also revealed that hundreds of people were held in Guantánamo with no real evidence against them."

(3a) Fordított: „A WikiLeaks által kiszivárogtatott aktáknak köszönhetően most még világosabb képet kaphatunk az ott elkövetett legnagyobb emberi jogi

jogsértésekről – a szimulált vízbefojtás és a brutális kínzás más formáinak széles körű alkalmazásáról a vallomások és információk kicsikarása érdekében -, és ezek az akták arra is rávilágítanak, hogy a guantanamói fogolytáborban emberek százait tartják fogva valódi bizonyítékok nélkül."

(3b) Tolmácsolt: „A Wikileaks fájlok világosabb képet adtak a lehető legsúlyosabb emberjogi sértésről. Például vízbe lenyomták az érintetteket, kínzással próbálták őket szóra bírni. Kiderült az is, hogy több száz embert tartottak fogva Guantanamón minden bizonyíték nélkül, ami ellenük szólt volna."

A szövegekre lebontott adatok mutatják, hogy időnként előfordul a fenti eljárás ellenkezője is. Három esetben a tolmács kevesebb mondatot használt, vagyis az eredetiben több mondatban elhangzó információt tömörítette, és egy mondatban fejezte ki. Ennek oka a beszélő mögötti lemaradás minimalizálása vagy a memóriatelítettség elkerülése. Mondatok összevonását figyelhetjük meg az alábbi 4. példában is.

(4) Eredeti magyar: „Nagyon helyes és társadalmilag igazságos az a javaslat, hogy hozzunk létre egy átmeneti támogatási kategóriát. Elfogadhatatlan azonban a kohéziós támogatásoknak a makrogazdasági feltételrendszerhez kötése, mivel a régiókat olyanért büntetnénk, olyan kormányzati politikáért, amelyre nem tudnak befolyást gyakorolni."

(4a) Angolra fordított: „The proposal that we should establish a transitional support category is indeed correct and socially fair. However, attaching cohesion support to a system of macro-economical conditions is unacceptable, because we would be punishing the regions for a governmental policy on which they have no influence."

(4b) Angolra tolmácsolt: „It is fair and right to say … that there ought to be a transitional method found >fund< but it is not acceptable but cohesion objectives and macro-economic objectives need to be bound inextricably together

because that would punish regions for something that they don't have no control over."

6.3. Statisztikai elemzés: átlagos mondathosszúság

Az átlagos mondathosszúság kiszámításához a Wordsmith Tools 6.0 gépi elemzőprogram statisztikai funkcióját használtam. Az eredményeket a *4. táblázat* foglalja össze.

A táblázatban látható adatok megerősítették az eddigi megállapításokat. Mivel a tolmácsok az esetek többségében és összességében kevesebb szót használnak több mondatban, míg a fordítók több szót sűrítenek kevesebb mondatba, a tolmácsolt szövegek átlagos mondathosszúsága csak kevesebb lehet, mint a fordított szövegeké, következésképpen egyszerűbbek azoknál. A nyelvi irányok szerinti eredmények ismételten azt mutatják, hogy anyanyelvre történő direkt fordítás és tolmácsolás során a nyelvi közvetítők bátrabban, magabiztosabban egyszerűsítettek.

4. táblázat

Átlagos mondathosszúság a vizsgált szövegekben

Eredeti angol írott	24,70 szó (19,63; 19,5; 24,71; 26,5; 42,5)	**Magyarra fordított**	20,11 szó (14,78; 18,25; 23,29; 23; 24)
Eredeti magyar írott	18 szó (14,86; 22; 15,75; 24,17; 15,38)	**Angolra fordított**	26,61 szó (19,43; 27,67; 25,14; 33,5; 28,43)
Eredeti angol hangzó	19,76 szó (18; 19,88; 25,43; 17,33; 19,44)	**Magyarra tolmácsolt**	13,64 szó (12,29; 14,11; 14,63; 20,17; 10,17)
Eredeti magyar hangzó	16,17 szó (13,38; 18,13; 17,25; 21,29; 12,4)	**Angolra tolmácsolt**	24,56 szó (29,6; 20,38; 22,45; 36,5; 21,83)

Az intermodális összehasonlítás eredményein kívül, a táblázatból az is világosan kiolvasható, hogy a hangzó szövegek átlagos mondathosszúsága minden esetben (az autentikus és a tolmácsolt szövegekben is) kevesebb, mint az írott szövegeké. Ennek oka minden bizonnyal a feldolgozás megkönnyítése, hiszen itt nincs lehetőség a mondatok újraolvasására.

Továbbá jól látható a két nyelv rendszerbeli különbsége is. Az agglutináló magyar szinte minden esetben kevesebb szót tartalmazó mondatokkal dolgozik, mint az izoláló angol.

6.4. Névmások száma

Shlesinger és Ordan (2012) korábbi vizsgálataiból kiderült, hogy a tolmácsok jelentősen több névmást használnak, mint fordító társaik, gyakran a tulajdonneveket is személyes névmással helyettesítik az egyszerűsítés érdekében. A jelen kutatás adatai egyértelműen alátámasztják az eredményeiket, ahogyan az 5. táblázatban látható.

5. táblázat

Névmások száma a vizsgált szövegekben

		Névmások száma	
Eredeti írott szövegek	134	A: 70 (9, 17, 19, 21, 4).	M: 64 (10, 15, 12, 11, 16)
Eredeti hangzó szövegek	139	A: 77 (13, 20, 19, 18, 7)	M: 62 (11, 15, 13, 8, 15)
Fordított szövegek	146	A-M: 69 (6, 15, 22, 14, 12)	M-A: 77 (9, 29, 13, 10, 16)
Tolmácsolt szövegek	172	A-M: 61 (5, 12, 22, 13, 9)	M-A: 111 (16, 22, 38, 23, 12)

A lenti példa jól illusztrálja a különbséget. Az eredeti hangzó szöveg (beszélt nyelv), amikor visszautal a már korábban elhangzott *Regionális Fejlesztési Bizottságra*, akkor tulajdonképpen személyes névmást használ, de mivel a magyarban normakövető művelet azok elhagyása – csak hangsúlyos helyzetben van helyük a mondatban –, csak az igéből (*látja*) következtethetünk rá.

A fordító ügyel az írott nyelv eleganciájára, így nem névmással helyettesíti a tulajdonnevet, ugyanakkor lerövidíti a *Regionális Fejlesztési Bizottság* kifejezést, és helyette csak

a *Bizottság* szót használja, hiszen a szövegkörnyezetből így is mindenki számára teljesen világos, hogy ugyanazon testületről beszélünk. A tolmács még egyszerűbben fogalmaz. Mivel a beszélő is a Bizottság tagja, így a *we* névmás tökéletesen megfelel a tolmácsolási helyzetben hosszú, időt igénylő kifejezés helyettesítésére, amelynek ismétlése a forrásnyelvi szöveg hallgatása során zavaró lehet, és akár még következményekkel is járhat: előfordulhat, hogy miközben a tolmács ezt a hosszú kifejezést ismétli, nem hallja pontosan a beszélő szavait, esetleg zavarja a feldolgozásban, így legrosszabb esetben információvesztés következhet be.

(5) Eredeti: „*A Regionális Fejlesztési Bizottság* megtárgyalta és örömmel támogatja az előterjesztést. Úgy látja, hogy elég körültekintő, és jó felkészülést biztosít nemcsak magára a trialógusra, hanem a későbbiekben az egész költségvetési folyamatra."

(5a) Fordított: „Mr, President, *the Committee on Regional Development* has debated and gladly supports the proposal. *The Committee* finds it to be quite prudent and believes that it provides appropriate preparation not only for the trilogue itself, but also for the whole budgetary process in the future."

(5b) Tolmácsolt: „On behalf of *the Regional Development Comittee* I can say that we have discussed the proposal and we can endorse it. *We* believe that it was carefully drafted to prepare us for the trialogue and also for the entire budgetary process."

Látható az is, hogy a hangzó szövegekben magasabb a névmások száma, ez a beszélt nyelvi jelleget támasztja alá. Ha pedig vetünk egy pillantást a nyelvi irányok szerinti eredményekre, ismét bebizonyosodik, hogy az angolban a legtöbb esetben kötelező a névmás kitétele, míg a magyarban ezt az ige személyragja fejezi ki. Nem kötelező ugyan a névmás elhagyása, de normakövetőnek számít, és csak akkor szükséges, ha hangsúlyos pozícióban áll.

6.5. Szótípusok és szövegszók aránya (*Type/token ratio, TTR*)

A szótípusok (a szövegben előforduló különböző szavak) és szövegszók (az összes szó) aránya hasznos statisztikai adat, fontos információkkal szolgál a szövegek lexikai változatosságát, szókincsét illetően. Minél magasabb a szótípusok aránya az összes szövegszó számához képest, annál változatosabb a szöveg szókészlete. A *TTR* összességében mindkét nyelvi irányban alacsonyabb értéket mutat a tolmácsolt szövegekben, mint a fordítottban – jóllehet angolról magyarra fordított, illetve tolmácsolt szövegek esetén nem számottevő a különbség –, ami alátámasztja Shlesinger (2008) eredményeit és következtetését: a tolmácsolt szövegek a szótípusok és szövegszók aránya tekintetében egyszerűbbek, mint a fordítottak.

6. táblázat

A szótípusok és szövegszavak aránya a vizsgált szövegekben

Eredeti angol írott	46,50 (69,43; 60,26; 63,01; 64,78; 67,06)	**Magyarra fordított**	63,67 (76,69; 73,29; 72,39; 79,71; 75,69)
Eredeti magyar írott	65,24 (80,77; 75; 75,4; 67,59; 78,86)	**Angolra fordított**	45,67 (69,12; 62,65; 62,5; 57,71; 61,31)
Eredeti angol hangzó	45,90 (68,52; 58,49; 61,24; 68,59; 65,14)	**Magyarra tolmácsolt**	63 (79,07; 74,02; 76,07;73,55, 76,23)
Eredeti magyar hangzó	63,50 (81,31; 72,41; 74,64; 67,79; 78,23)	**Angolra tolmácsolt**	41,92 (62,84; 57,67; 53,44; 58,22; 65,65)

Ha azonban szövegekre lebontva vizsgáljuk az adatokat, akkor ezt csak a magyarról angol nyelvre történő fordítás és tolmácsolás esetén látjuk megvalósulni. A másik nyelvi irányban a TTR magasabb a fordított és tolmácsolt szövegekben. Ennek hátterében az állhat, hogy a nyelvi közvetítők anyanyelvre történő tolmácsoláskor magabiztosabban és bátrabban bánnak a kihagyásokkal, a redundancia csökkentésével, melynek következtében nőhet a TTR.

Bár ez egy intermodális összehasonlítás, mégis érdemes egy pillantást vetni az eredeti szövegek adataira. Az eddigi kutatásoknak megfelelően kisebb a TTR az angol és magyar autentikus hangzó szövegalkotásban, mint írásban, következésképpen a kisebb

TTR lehet a szóbeliség jellemzője is. Ha a két fenti eredményt együttesen szeretnénk értelmezni (tolmácsolásban is alacsonyabb, hangzó szövegben is) akkor megállapíthatjuk a tolmácsolt szöveg a hangzó szöveg ezen sajátosságát mindenképp magán viseli. Továbbá az írott és hangzó eredmények nagyon hasonlóak a fordított és tolmácsolt eredményekhez, így lehetséges, hogy a különbség nem feltétlenül csak a közvetítés módjának, hanem a szóbeliség és írásbeliség sajátosságainak tudható be. Ugyanakkor az is jól látható, hogy a hasonlóság az eredeti angol, angolra fordított és tolmácsolt valamint az eredeti magyar, magyarra tolmácsolt és fordított szövegek között is fennáll, azaz beszélhetünk az egyes nyelvekre jellemző tulajdonságokról is. Továbbá nem mehetünk el észrevétlenül amellett az egyértelműen látszó tény mellett sem, hogy az eredetihez képest mind a tolmácsolt, mind a fordított szövegek szótípus/szövegszó aránya csökkent, ami minden bizonnyal a nyelvi közvetítés következménye, és így a nyelvi közvetítés univerzáléja is.

6.6. A statisztikai eredmények összefoglalása

A statisztikai elemzés során öt adatot vizsgáltam: a szavak, mondatok, névmások számát, az átlagos mondathosszúságot és a szótípus/szövegszó arányt. A célom kettős volt. Elsősorban arra voltam kíváncsi, vannak-e a tolmácsolt szövegeknek általános, univerzális tulajdonságai, felfedezhetőek-e bizonyos mintázatok, amelyek alapján kimondhatjuk, hogy létezik saját nyelvi jellemzőkkel rendelkező tolmácsolási szöveg, másodsorban pedig a modalitás, illetve ontológia hatásának erősségét vizsgáltam. A tolmácsolt szövegekre (a fordított szövegekkel összevetve) vonatkozó, statisztikai adatok alapján levont következtetéseket az alábbi táblázatban foglalom össze.

A táblázat adatai alapján megállapítható, hogy a tolmácsolt szövegek egyszerűbbek, mint a fordítások, és inkább a beszélt nyelv tulajdonságait viselik magukon.

Jól látható az is, hogy a tolmácsolási szöveg univerzáléi is a beszélt nyelv jellemzői felé mutatnak.

6. táblázat

Statisztikai adatok összefoglalása a kutatás két fő kérdése szempontjából

	Univerzálé	**Ontológia vagy modalitás**
Szavak száma	kihagyás > egyszerűsítés	nem redundáns ≠ beszélt nyelv kontextusfüggő = beszélt nyelv
Mondatok száma	több mondat > szegmentálás	beszélt nyelv
Átlagos mondathosszúság	rövidebb > egyszerűsítés	beszélt nyelv
Névmások száma	több	beszélt nyelv
TTR	alacsonyabb > egyszerűsítés	beszélt nyelv

7. Kontrasztív elemzés

A következőkben a kontrasztív, kvalitatív szövegelemzés eredményeit ismertetem a fordítói műveletek tükrében (*8. táblázat*) és a leggyakoribb fordítói műveleteket illetően (*8. táblázat*), azokon belül is lexikai illetve grammatikai műveletekre vonatkozóan.

7.1. Átváltási műveletek száma

A kontrasztív szövegelemzés első lépéseként Klaudy (1997) átváltási műveletekre kidolgozott tipológiája alapján azonosítottam a fordított és tolmácsolt szövegekben fellelhető fakultatív – normakövető és stratégiakövető – műveleteket. Az eredmények a *8. táblázatban* láthatók, ahol a fakultatív műveletek és az egyes szövegek adatai is összesítve szerepelnek.

Korábbi kutatások (Makkos és Robin 2012; Robin 2018) kimutatták, hogy a kötelező műveletek többsége grammatikai, mivel nyelvspecifikus műveletekről van szó. A jelen elemzés kizárta vizsgálata köréből a kötelező műveleteket, csak a fakultatívakra koncentrált, így a grammatikai műveletek száma eleve alacsonyabb lett.

8. táblázat

Átváltási műveletek száma a vizsgált szövegekben

	A-M	M-A	Összesen
Átváltási műveletek száma fordított szövegekben	151	134	285
ebből lexikai	116 (77%)	72 (53%)	188 (66%)
ebből grammatikai	35 (23%)	62 (47%)	97 (34%)
Átváltási műveletek száma tolmácsolt szövegekben	263	293	556
ebből lexikai	216 (82%)	225 (77%)	441 (79%)
ebből grammatikai	47 (18%)	68 (23%)	115 (21%)

A táblázatból kiolvasható, hogy a tolmácsok jóval több, majdnem kétszer annyi fakultatív átváltási műveletet végeztek (556), mint a fordítók (285), vagyis jobban átalakították az eredeti hangzó szöveget. A fakultatív műveleteken belül a lexikai műveletek voltak többségben mindkét modalitás esetén, de a tolmácsolásnál kiugróan magas értékekről beszélhetünk. Ennek magyarázata valószínűleg abban rejlik, hogy tolmácsolásnál az információ feldolgozása és átadása az elsődleges feladat, az információt pedig a tartalmas lexikai elemek hordozzák. A normakövető és stratégiakövető grammatikai elemek inkább a szövegszervezést és a feldolgozhatóság megkönnyítését célozzák, létrehozzák a logikai kapcsolatokat, segítenek a kohézió és így a koherencia megteremtésében. Írásban erre nagy szükség van, de tolmácsolás esetén a kontextus, az adott beszédhelyzet hozzáférhetősége segít, így nincs olyan nagy szükség felszíni kohéziós elemekre – a hallgató össze tudja kapcsolni a logikai összefüggéseket.

7.2. A leggyakoribb műveletek száma

Az elemzés következő lépéseként Klaudy (1997) tipológiáját követve kategorizáltam a lexikai és grammatikai műveleteket: lexikai konkretizáció, generalizáció, felbontás, kiha-

gyás, betoldás, áthelyezés, csere, teljes átalakítás és kompenzálás; grammatikai konkretizáció, generalizáció, felemelés, lesüllyesztés, kihagyás, betoldás, felbontás, összevonás, áthelyezés és csere. A 9. táblázat a négy leggyakoribb műveletet mutatja a fordításban és a tolmácsolásban előbb összességében, majd nyelvi irányonként lebontva.

Fordításban az első két helyen lexikai műveletek állnak, a harmadik és negyedik helyre azonban grammatikai műveletek kerültek, ami talán nem is oly meglepő, hiszen az írott szöveg tudatosabb, igényesebb, szabályosabb szerkesztésű, „nyelvi illemszabályai" szigorúbbak, jóval fontosabbak a nyelvi kifejezőeszközök, mint a hangzó szövegek esetében.

. Ezzel szemben tolmácsolásnál kizárólag lexikai műveletek találhatóak. A grammatikai műveletek, azon belül is a csere komoly erőfeszítést, a forrásnyelvi szövegtől való elvonatkoztatást igényel.

A fordítónak ez nem jelent akkora gondot, mint az időkényszer alatt dolgozó, az erőfeszítéseit kiegyensúlyozni igyekvő tolmács számára. Ugyanez lehet az oka annak is, hogy a lexikai konkretizáció a fordításnál az első, míg tolmácsolásnál csak a negyedik helyen található. A tolmács ennek ellenkezőjét teszi, általánosít: mivel gyakran nem áll rendelkezésére elegendő idő a megfelelő szó keresésére, a memóriájából legkönnyebben lehívható szót aktiválja.

Mindez megerősíti He, Boyd-Graber és Daumé (2016) fentiekben ismertetett kutatásának eredményeit.

A táblázatból az is jól látható, hogy míg a fordításban az explicitáció (betoldás, konkretizáció), addig a tolmácsolásban az implicitáció (kihagyás, általánosítás) az uralkodó műveleti kategória.

Fordításnál az első négy leggyakoribb műveletet összesen 188 alkalommal alkalmazták a fordítók, ebből – mivel a táblázat nem részletezi, hogy a grammatikai csere adott esetekben az explicitáció vagy az implicitáció kategóriájába tartozik-e – 148–188 alkalommal explicitáltak (79–100%).

9. táblázat

Leggyakoribb műveletek száma a vizsgált szövegekben

Leggyakoribb műveletek száma fordításban		**Leggyakoribb műveletek száma tolmácsolásban**	
lexikai konkretizáció	78	lexikai kihagyás	201
	A-M: 52, M-A: 26		A-M: 106, M-A: 95
lexikai betoldás	52	lexikai betoldás	92
	A-M: 29, M-A: 23		A-M: 41, M-A: 51
grammatikai csere	40	lexikai általánosítás	70
	A-M: 23, M-A: 18		A-M: 31, M-A: 44
grammatikai betoldás	18	lexikai konkretizáció	60
	A-M: 19, M-A: 17		A-M: 26, M-A: 29

A tolmácsok az összesen 423 elvégzett műveletből 271 alkalommal implicitáltak (66%), és csak 152 esetben explicitáltak (34%). Vagyis a tolmácsolt szövegek jóval implicitebbek, mint a fordítottak, de az explicitáció jelenségétől sem mentesek.

Tolmácsolás esetén első helyen a lexikai kihagyás áll. A tolmácsok kihagynak bizonyos szavakat, mondatrészeket, vagy akár egész mondatokat is a szövegből, ha felesleges, redundáns, beszédhelyzetből kikövetkeztethető információról van szó. Így több idejük jut a feldolgozásra és a memória telítődése sem következik be olyan hamar. A fordítóknak általában elég idő áll rendelkezésükre a feldolgozáshoz, és memóriájukat sem kell olyan intenzitással használniuk, mint a tolmácsoknak, hiszen bármikor visszaolvashatják a fordítandó szöveget, ezért jóval ritkábban élnek ezzel a lehetőséggel. A tolmácsolási szöveg tehát egyfajta redukált szöveg, ettől függetlenül mégis betölti funkcióját, ha a befogadóhoz elér az üzenet. A második és a negyedik helyen első pillantásra furcsamód két explicitációval járó művelet került, összesen 152 előfordulással. Mit sem bizonyít jobban ez a kettősség, mint a Shlesinger (2008) által vizsgált és megállapított tényt: a tolmácsolás is a fordítás egy fajtája, vagyis a modalitás is hatással van rá, de a beszélt nyelvi jelleg jóval erősebben befolyásolja.

Érdekes eredményeket hoztak a nyelvi irányok szerinti lebontások. Nem anyanyelvre történő fordítás során minden átváltási műveletből kevesebbet hajtottak végre a fordítók, mint anyanyelvre történő fordítás során, azaz ismét bebizonyosodott az a megállapítás, mely szerint az anyanyelv nagy szerepet játszik a nyelvi közvetítők magabiztosságában. Tolmácsolásnál csak a kihagyás műveletével bántak magabiztosabban anyanyelvre tolmácsolásnál: jobban érezték a kontextust, több kockázatot vállaltak. Angolra tolmácsolásnál inkább betoldottak, konkretizáltak és általánosítottak, azaz magyaráztak. Az idegen nyelvre történő tolmácsolás esetén láthatóan több műveletet végeztek, mint a fordítók, talán mert jobban igyekeztek, törekedtek az üzenet átadására, a feldolgozás megkönnyítésére.

8. Összefoglalás

A projekt első, párhuzamos szakaszának eredményei szerint a tolmácsolt szövegek az eredeti forrásnyelvi megfelelőjükhöz képest egyszerűbbek, implicitebbek. Ezek a jellemzők a modalitások összehasonlítása során ismételten bizonyítást nyertek. A statisztikai elemzés eredményei egyértelműen mutatják, és így alátámasztják Shlesinger (2008) illetve Bernardini (2016) és társai eredményeit, hogy – mind a szavak, mind a mondatok száma, az átlagos mondathosszúság, valamint a szövegtípusok és szövegszók aránya alapján – a tolmácsolt szövegek a fordítottakhoz képest is egyszerűbbek, kevésbé változatosak. A fenti adatok, kiegészülve a névmások számának összehasonlításából származó eredményekkel a modalitás elsődleges hatását, a tolmácsolt szövegek beszélt nyelvi jellegét hangsúlyozzák. A kontrasztív egybevetés hasonló eredményeket hozott. A tolmácsok jóval gyakrabban élnek az implicitáció kategóriájába tartozó átváltásokkal (kihagyás, általánosítás), mint fordító társaik, következésképpen a tolmácsolt szövegek implicitebbek. A redundancia tűnik a hangzó szöveg egyetlen olyan tulajdonságának, melyen a tolmácsolási szöveg nem osztozik. A kontrasztív összehasonlítás azt is lehetővé tette, hogy a számadatok mögé nézzünk, és megkíséreljünk magyarázatot találni a miértekre. A leggyakoribb oknak a memória feldolgozási kapacitásának korlátai és az időkényszer bizonyultak. A projekt harmadik szakaszában továbbra is a tolmácsolt szövegek jellemzőit igyekszem kutatni, de

akkor eredeti, a forrásnyelven elhangzott beszédek átiratához fogom hasonlítani őket. A jelen kutatás eredményeit, valamint az eredmények alapján megfogalmazódó hipotéziseket az általánosíthatóság érdekében mindenképpen érdemes több autentikus, fordított és tolmácsolt szöveg felhasználásával, illetve más nyelvpárok esetében is megismételni, így kerülhetünk közelebb a tolmácsolási szöveg meghatározásához.

Irodalom

Aarts, J., Meijs, W. 1984. *Corpus Linguistics: Recent Developments in the Use of Computer Corpora in English Language Research*. Amsterdam: Rodopi. https://doi.org/10.1111/j.1467-968X.1983.tb01200.x

Altman, J. 1994. Error analysis in the teaching of simultaneous interpretation: A pilot study. In: Lambert, S., Moser-Mercer, B. (eds) *Bridging the gap: Empirical research in simultaneous interpretation*. Amsterdam/Philadelphia: John Benjamins. 25–38. https://doi.org/10.1075/btl.3.05alt

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. https://doi.org/10.1075/z.64.15bak

Baker, M. 1995. Corpora in Translation Studies: An Overview and Some Suggestions for Future Research. *Target* Vol.7. No. 2. 223–243. https://doi.org/10.1075/target.7.2.03bak

Baker, M. 1996. Corpus-based translation studies: the challenges that lie ahead. In: Somers, H. (ed.) *Terminology, LSP and Translation: Studies in Language Engineering in honour of Juan C. Sager*. Amsterdam Philadelphia: John Benjamins. 175–186. https://doi.org/10.1075/btl.18.17bak

Barik, H. C. 1994. A description of various types of omissions, additions and errors of translation encountered in simultaneous interpretation. In: Lambert, S., Mo-

ser-Mercer, B. (eds) *Bridging the gap: Empirical research in simultaneous interpretation*. Amsterdam and Philadelphia: John Benjamins. 121–137. https://doi.org/10.7202/001972ar

Bendazzoli, C., Sandrelli, A. 2005. An approach to corpus-based interpreting studies: Developing EPIC (European Parliament Interpreting Corpus). In: Gerzymisch-Arbogast, H., Neuert, S. *EU-High-Level Scientific Conference Series MuTra 2005 – Challenges of Multidimensional Translation: Conference Proceedings.* Saarbrücken: Saarland University. 149–160.

Bendazzoli, C., Sandrelli, A. 2009. Corpus-based Interpreting Studies: Early work and future prospects. *Revista Tradumatica* No. 7. L'aplicació del corpus lingüístics a la traducció. https://doi.org/10.1.1.601.7973

Bernardini, S., Ferraresi, A., Milicevic, M. 2016. From EPIC to EPTIC – Exploring simplification in interpreting and translation from an intermodal perspective. *Target* Vol 28. No. 1. 61–86. https://doi.org/10.1075/target.28.1.03ber

Biber, D., Conrad, S., Reppen, R. 1994. Corpus-based approaches to issues in applied linguistics. *Applied Linguistics* Vol. 15. 169–89. https://doi.org/10.1093/applin/15.2.169

Blum-Kulka, S. 1986. Shifts of Cohesion and Coherence in Translation. In: House, J. Blum-Kulka, S. (eds) *Interlingual and Intercultural Communication. Discourse and Cognitionin Translation and Second Language Acquisition*. Tübingen: Narr. 17–35.

Cencini, M. 2000. *Il Television Interpreting Corpus (TIC). Proposta di codifica conforme alle norme TEI per transcriziioni di eventi di interpretazione in televisione*. Unpublished dissertation. Advanced School for Translators and Interpreters (SSLMIT). University of Bologna at Forlí.

Cencini, M. 2002. Ont he importance of an encoding standard for corpus-based interpreting studies. Extending the TEI scheme. *inTRAlinea Special Issue* http://www.intralinea.org/specials/article/1678

Chernov, G.V. 1978. *Teoriya ipraktika sinkhronnogo perevoda*. Moscow: Mezhdunarodnyye otnosheniya.

Chesterman, A. 2004. Beyond the Particular. In: Mauranen, A., Kujamaki, P. (eds) *Translation Universals: Do they exist?* Amsterdam: Benjamins. 33–49. https://doi.org/10.1075/btl.48.04che

Dejean LeFéal, K. 1978. *Lectures et improvisations*. Unpublished doctoral dissertation. Université de Paris III.

Diriker, E. 2001. *Contextualising Simultaneous Interpreting: Interpreters in the Ivory Tower?* Unpublihed doctoral dissertation. Bogazici University, Turkey.

Falbo, C. 2012. CorIT (Italian Interpreting Corpus): Classification Criteria. Television. In: Staniero, F., Falbo, C. (ed.) *Breaking Ground in Corpus-Based Interpreting Studies.* Bern: Peter Lang. 157–185. https://doi.org/10.3726/978-3-0351-0377-9

Fumagalli, D. 1999–2000. *Alla ricerca dell'interpretese. Uno studio sull'interpretazione consecutiva attraverso la corpus linguistics.* Unpublished MA Thesis. Advanced School for Translators and Interpreters (SSLMIT), University of Trieste.

Gile, D. 1995. *Basic concepts and models for interpreter and translator training*. Amsterdam and Philadelphia: John Benjamins. https://doi.org/10.1075/btl.8

Gile, D. 1998. Observational studies and experimental studies in the investigation of conference interpreting. *Target* Vol. 10. Nr. 1. 69–93. https://doi.org/10.1075/target.10.1.04gil

He, H., Boyd-Graber, J., Daumé III, H. 2016. Interpretese vs. Translationese: The Uniqueness of Human Strategies in Simultaneous Interpretation. In: Knight, K., Nenkova, A., Rambow, O. (eds) *Proceedings of the 2016 Conference of the North Ameri-*

can Chapter of the Association for Computational Linguistics: Human Language Technologies. 971–976. https://doi.org/10.18653/v1/N16-1

Heltai P. 2005b. A fordító és a nyelvi normák III. *Magyar nyelvőr* 129. évf. 2. szám. 165–172.

Hu, K., Tao, Q., 2013. The Chinese-English Conference Interpreting Corpus: Uses and Limitations. *Meta* Vol. 58. No. 3. 626–642. https://doi.org/10.7202/1025055ar

Kalina, S. 1998. *Strategische Prozesse beim Dolmetschen: theoretische Grundlagen, empirische Fallstudien, didaktische Konsequenzen.* Tübingen: Narr Verlag.

Károly K. 2002. Az alkalmazott nyelvészeti kutatások néhány alapvető módszertani kérdéséről. *Alkalmazott Nyelvtudomány* 2. évf. 1. szám. 77–87.

Károly K. 2007. *Szövegtan és fordítás*. Budapest: Akadémiai Kiadó.

Lanstyák, I. 2009. *A magyar beszélt nyelv sajátosságai*. Bratislava: Stimul.

Klaudy K. 1997. *A fordítás elmélete és gyakorlata*. Budapest: Scholastica.

Klaudy K. 1999. Az explicitációs hipotézisről. *Fordítástudomány* 2. évad. 1. kötet. 5–22.

Laviosa-Braithwaite, S. 1998. Universals of Translation. In: Baker, M. (ed.) *Routledge Encyclopedia of Translation*. London: Routledge. 288–291.

Laviosa, S. 2009. Universals. In: Baker, M. (ed.) *Encyclopedia of Translation Studies.* London: Routledge. 306–311.

Lederer, M. 1981. *La traduction simultanée. Expérience et théorie*. Paris: Minard.

Makkos A., Robin E. 2012. Fordítói műveletek a kompetencia függvényében. In: Horváthné Molnár K., Sciacovelli A. (szerk.) 20 *Az alkalmazott nyelvészet regionális és globális szerepe: alkalmazott nyelvészeti kutatások az EU magyar elnökség évében.* A MANYE kongresszusok előadásai 8. Budapest–Szombathely–Sopron: MANYE–NYME. 305–318.

Meyer, B. 2008. Interpreting Proper Names: Different Interventions in Simultaneous and Consecutive Interpreting? *trans-kom* Vol. 1. No. 1. 105–122.

Monacelli, C. 2005. *Surviving the Role: A corpus-based study of self-regulation in simultaneous interpreting as perceived through participation framework and interactional politeness*. PhD dissertation. Heriot Watt University.

Neubig, G., Hiroaki, S., Sakriani, S., Satoshi, N., Tomoki, T. 2017. The NAIST Simultaneous Translation Corpus. In: Russo, M., Bendazzoli, C., Defranq, B. (eds) *Making way in corpus-based interpreting studies.* Singapore: Springer. 205–215. https://doi.org/10.1007/978-981-10-6199-8_11

Oléron, P., Nanpon, H. 1965. Recherches sur la traduction simultanée. *Journal de psychologie normale et pathologique* Vol. 62. No. 1. 73–94.

Pöchhacker, F. 1994. Simultandolmetschen als komplexes Handeln. *Language in Performance* Vol. 10 Tübingen: Narr. https://doi.org/10.1075/target.7.1.16gil

Robin, E. 2018. *Fordítási univerzálék és lektorálás.* Budapest: ELTE Eötvös Kiadó.

Robin, E., Dankó, Sz., Götz, A., Nagy, A. L., Pataky, É., Szegh, H., Zolczer, P. 2016. Fordítástudomány és korpuszkutatás: bemutatkozik a Pannonia Korpusz. *Fordítástudomány* 18. évf. 2. szám. 5–26.

Russo, M., Bendazzoli, C., Defranq, B. 2018. Making way in corpus-based interpreting studies. *New Frontiers in Transation Studies*. Singapore: Springer. https://doi.org/10.1007/978-981-10-6199-8

Sandrelli, A. 2012. Interpreting Football Press Conferences.: The FOOTIE Corpus. In: J. Kellett Bidolli, C. (ed.) *Interpreting across Genres: Multiple Research Perspectives.* Trieste: EUT Edizioni Universitá di Trieste. 78–101.

Schäffner, C. 2004. Translation Research versus Interpreting Research: Kinship, differences and prospects for partnership. In Schäffner, C. (ed.) *Translation Research and*

Interperting Research. Traditions, Gaps and Synergies. Clevedon, Buffalo, Toronto: Multilingual Matters. 10–34.

Setton, R. 1997. *A pragmatic model of simultaneous interpretation*. Unpublished doctoral thesis. Chinese University of Hong-Kong.

Setton, R. 1999. *A cognitive-pragmatic analysis of simultaneous interpretation*. Amsterdam: John Benjamins. https://doi.org/10.7202/001886ar

Shlesinger, M. 1998. Corpus-based Interpreting Studies as an Offshoot of Corpus-based Translation Studies. *Meta* Vol. 43 No. 4. 486–493. https://doi.org/10.7202/004136ar

Shlesinger, M. 2008. Towards a definition of Interpretese. An intermodal, corpus-based study. In: Hansen, G., Chesterman, A., Gerzymisch-Arbogast, H. (eds) *Efforts and Models in Interpreting and Translation Research. A tribute to Daniel Gile*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/btl.80.18shl

Shlesinger, M., Ordan, N. 2012. More spoken or more translated? Exploring a known unknown of simultaneous interpreting. *Target* Vol. 24 No. 1. 43–60. https://doi.org/10.1075/target.24.1.04shl

Sinclair, J. H. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Spinolo, N. 2017. Studying figurative languages in simultaneous interpreting: the IMITES corpus. In: Russo, M., Bendazzoli, C., Defranq, B. (eds) *Making way in corpus-based interpreting studies.* Singapore: Springer. https://doi.org/10.1007/978-981-10-6199-8_8

Szegh H. 2017. Harmadik kód a tolmácsolásban: létezik tolmácsolási szöveg? *Fordítástudomány* 19. évf. 1. szám. 60–74.

Tohyama, H., Matsubara, Sh. 2006. Collection of Simultaneous Interpreting Patterns by Using Bilingual Spoken Monologue Corpus. In: Calzolari, N., Choukri, K., Gangemi, Aldo., Maegaard, B., Mariani, J., Odijk, J., Tapias, D. (eds) *Proceedings*

of the Fifth Annual Conference on Language Resources and Evaluation. Genoa: European Language Resources Association (ELRA). 2564–2569.

Tohyama, H., Ryu, K., Matsubara, Sh., Kawaguchi, N., Inagaki, Y. 2004. CIAIR Simultaneous Corpus. https://pdfs.semanticscholar.org/d5f6/9e1db393df443b6cc2f754a4488526378cde.pdf?_ga=2.15170165.233430643.1553866200-1540436159.1553613303&_gac=1.226022888.1553866200.Cj0KCQjwhPfkBRD0ARIsAAcYycEuue-xdO8bLup8J6ApW-IObc00CrfAMmo9YZnsTuKadXO-n2M8QGgaAh4tEALw_wcB

Toury, G. 1980. *In Search of a Theory of Translation*. Tel-Aviv: Tel-Aviv University

Toury, G. 1995. *Descriptive Translation Studies and Beyond*. Amsterdam: John Benjamins. https://doi.org/10.1075/btl.4

Wang, B. 2012. Interpreting Strategies in Real-life Interpreting. *Translation Journal* Vol. 16. No. 2. http://translationjournal.net/journal/60interpreting.htm

Internetes hivatkozások

Magyar Nemzeti Szövegtár. Elérhető: http://mnsz.nytud.hu/

Wordsmith Tools. Elérhető:

https://www.lexically.net/wordsmith/

Diskurzusjelölők és kötőelemek a magyar szinkrontolmácsolt és eredeti diskurzusban[1]

Götz Andrea

gotz.andrea@kre.hu

Károli Gáspár Református Egyetem

Kivonat: A jelen feltáró jellegű kutatás a magyar szinkrontolmácsolt diskurzus diskurzusjelölő és pragmatikai elemkészletét vizsgálja angol–magyar párhuzamos korpusz és egy megegyező nagyságú összehasonlítható, nem tolmácsolt magyar korpusz segítségével. A korpuszok európai parlamenti felszólalásokat tartalmaznak. A kutatás alapvetően arra keresi a választ, mennyiben motiválják a forrásszövegek a magyar tolmácsolás pragmatikai és diskurzusjelölőit, gyakoribbak-e ezek az elemek a tolmácsolt, mint a nem tolmácsolt szövegekben, továbbá ez a szignifikáns eltérő-e. A vizsgálat eredményei szerint a tolmácsolt szövegekben található diskurzusjelölő és pragmatikai elemek többsége hozzáadás eredménye, mivel az elemek 53,41%-a hozzáadás eredményeként került a magyar tolmácsolt diskurzusba. Továbbá a jelen kutatás eredményei szerint a tolmácsolt szövegek szignifikánsan kevesebb (7,78%).

Kulcsszavak: diskurzusjelölő, európai parlamenti felszólalások, párhuzamos korpusz, pragmatikai elemek, tolmácsolt szöveg

1. Bevezetés

Ellentétben a fordított szövegekkel, a tolmácsolt szövegek diskurzusszervezésének, kohéziós elemeinek és diskurzusjelölő-használata kevésbé kutatott. A diskur-

* Az Emberi Erőforrások Minisztériuma ÚNKP-18-3 kódszámú Új Nemzeti Kiválóság Programjának támogatásával készült

Götz Andrea: Diskurzusjelölők és kötőelemek a magyar szinkrontolmácsolt és eredeti diskurzusban: feltáró korpuszkutatás. In: Robin E., Seidl-Péch O. (szerk.) 2020. *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordításkutatás Magyarországon.* Segédkönyvek a nyelvi közvetítésről I. Budapest: ELTE BTK Fordítástudományi Doktori Program, MANYE Fordítástudományi Szakosztály. DOI: https://doi.org/10.36252/Nyelvikozvsegedkonyv1.12

zusjelölők és kötőelemek használatát vagy az elemek gyakoriságának vagy a hozzáadott elemek arányának, motiváltságának (azonosítható-e forrásnyelvi párhuzamos elem a forrásszövegben), vagy a betoldott diskurzusjelölők céljának, illetve funkciójának szempontjából vizsgálják. A kutatások azonban egyes nyelvpárokra korlátozódnak, és bizonyos tolmácsolási események esetében sem nyújtanak rendszeres leírást, ugyan egyes diskurzusjelölők használatát már vizsgálták bizonyos tolmácsolási környezetben széles körűen (l. Defrancq 2016: a *well* az eredeti és tolmácsolt európai parlamenti szövegekben).

Ennek ellenére bizonyos tendenciák kirajzolódnak az eddigi kutatásokból. A hozzáadott diskurzusjelölők, vagyis azok a célnyelvben megjelenő elemek, amelyek nem rendelkeznek forrásnyelvi párhuzammal az eredeti megnyilatkozásban, a tolmács feldolgozási folyamatára utalnak, ugyanakkor a hallgatóság az eredeti beszélőnek tulajdonítja őket (Blakemore és Gallai 2014). Bizonyos tolmácsolási események szempontjából a diskurzusjelölők szerepe különösen fontos, mivel negatív módon is hathatnak vagy reflektálhatnak az eredeti beszélőre. Így például a bírósági tolmácsolás esetében a tolmács hozzáadott – vagy az eredeti megnyilatkozásból kihagyott – diskurzusjelölői tévesen közvetíthetik a megnyilatkozást vagy elbizonytalaníthatják a beszélőket, akik „az érvelés logikáját körültekintően megválasztott diskurzusjelölőkkel segítik” (Hale 2010: 70). Továbbá a tolmácsolt szövegek, a forrásszövegekkel vagy összehasonlítható szövegekkel egybevetve, gyakrabban tartalmazhatnak kötőelemeket (pl. Defrancq, Plevoets és Magnifico 2015, Defrancq 2016).

A jelen feltáró jellegű kutatás a magyar szinkrontolmácsolt diskurzus diskurzusjelölő- és kötőelemkészletét tárja fel egy angol–magyar párhuzamos korpusz és egy megegyező nagyságú összehasonlítható, nem tolmácsolt magyar korpusz segítségével. A korpuszok európai parlamenti felszólalásokat tartalmaznak. A kutatás alapvetően arra keresi a választ, mennyiben motiválják a forrásszövegek a magyar tolmácsolás diskurzusjelölőit és kötőelemeit, valamint gyakoribbak-e ezek az elemek a tolmácsolt, mint a nem tolmácsolt szövegekben, és amennyiben tapasztalható eltérés, vajon ez szignifikáns-e.

2. Diskurzusjelölők

Fraser (1999) meghatározásában a diskurzusjelölők azt mutatják, hogy a beszélő szándéka szerint a jelen alapüzenet hogyan kapcsolódik a megelőző üzenethez vagy annak egy részéhez (Fraser 1990: 387, Fraser 1999: 938). Ez a definíció Fraser (1990: 386) mondatjelentés-felfogásából ered. Ennek értelmében a mondatjelentést két összetevőre, tartalmas és pragmatikai jelentésre lehet osztani. A pragmatikai jelentés a beszélő kommunikatív szándékát jeleníti meg, és három jelölőosztály fejezheti ki, amelyek az imént említett kommunikációs szándékot jelölik. E jelölőosztályok a következők: (a) alapvető pragmatikai jelölők, (b) kommentáló pragmatikai jelölők, valamint (c) párhuzamos pragmatikai jelölők. E három kategóriát további alkategóriákba lehet sorolni (Fraser 1996). Fraser (1999) a diskurzusjelölőket a pragmatikai jelölők kommentáló alkategóriájába tartozó elemeknek tekinti, a „diskurzusszegmens", majd pedig a „megnyilatkozás" (Fraser 2015) kifejezést is használja üzenet helyett.

Ugyan a diskurzusjelölők Fraser felfogásában szemantikai viszonyt jeleznek két szegmens között, tipikusan a második szegmens kezdőpozíciójában jelenve meg, a diskurzusjelölő elemnek magának nincs szemantikai tartalma (Fraser 2015: 48). A diskurzusjelölők tovább csoportosíthatók (a) kontrasztív (pl. *de*, *ám*), (b) elaboratív (*és*, *továbbá*), valamint (c) implikatív (*így*, *tehát*) elemek kategóriájába. Ez a megközelítés közel áll Halliday és Matthiessen (2014) kötőszó fogalmához, amely szintén szemantikai kapcsolatot jelenít meg és kohézív viszonyt fejez ki. Ezt a nézetet rokoníthatjuk a pragmatikai kötőszó fogalmával (Németh T. 1998), amelynek értelmében olyan kötőszavak, mint a *hát*, *így*, *tehát*, *mert* szintaktikai funkciójukon túl pragmatikai tulajdonsággal is bírnak (illokúciós, interperszonális, attitudinális) az adott használatban, így gyakran a kötőszó és diskurzusjelölő fogalma nehezen választható el.

A magyar irodalomban a diskurzusjelölőket tágabb jelentésben és ebből következően funkciókörben találjuk meg, a diskurzusjelölő kategóriáját funkcionális szóosztálynak tartva (Schirm 2009) (függetlenül attól, hogy minden elem szónak tekinthető-e), vagyis a pragmatikai funkció, nem pedig formális jegyek jelölik ki. Ez a felfogás túlnyomóan jellemző a magyar diskurzusjelölő- vagy pragmatikaijelölő-kutatásra. Empirikus

kutatások ugyanakkor arra is rámutattak, milyen nehézséggel jár ezen elemek funkcióinak elkülönítése, vagy éppenséggel az elemek szintaktikai vagy pragmatikai használatának azonosítása (Dér és Markó 2010). Dér és Markó (2010: 150) spontán beszéden végzett vizsgálata ezen túlmenően azt is megállapította, hogy sem pozicionális (pl. szegmenskezdő), sem fonológiai jegyek alapján sem lehet egymástól minden esetben megkülönböztetni a pragmatikai és szintaktikai szereppel is rendelkező elemek aktuális használatát. Ezenkívül az adatközlők egyéni nyelvhasználatában is akadtak különbségek, nemcsak a szerzők által diskurzusjelölőnek tartott elemek gyakoriságában, hanem az azok betöltötte funkciókban is (Dér és Markó 2010: 151).

A funkcióalapú definíció túl tágan meghatározott kategóriához is vezethet, amit egyes kutatók formális megkötések bevezetésével szűkítenek – a diskurzusjelölők szegmenskezdő helyet foglalnak el (Fraser 1999), a modális partikulák mediális helyzetben fordulnak elő (Aijmer 2007), míg mások nem tekintik e jegyet döntőnek (Traugott 2007: 141). A konszenzus hiányának következtében ugyanazok az elemek más-más kategória besorolást kaphatnak különböző kutatóknál, így például az angol *you know* (pragmatikai használatban megközelítőleg 'érted') leírható diskurzusjelölőként (Schiffrin 1987), vagy párhuzamos pragmatikai jelölőként (Fraser 1990: 392). A különbségtétel alapja Fraser (1990) szemléletében, hogy a *you know* inkább beszélői attitűdöt, mint szekvenciális diskurzusviszonyt közvetít. Fraser (1990) megközelítése a diskurzuskohézióhoz köti ezen elemeket, míg Schiffrin (1987) leírása a diskurzusjelölők ennél összetettebb diskurzusfunkciójára épül, magában foglalva a kohézió- és koherencia teremtő szerepet. Schiffrin felfogásában a „diskurzusjelölők elemzése a diskurzuskoherencia tágabb elemzésének részét alkotja" (Schiffrin 1987: 49).

A diskurzuskoherencia annak együttese, hogy hogyan építik a résztvevők az általuk kifejezett jelentést a diskurzus szerkezetébe, beszédcselekvéseket valósítva meg, amelynek létrehozásában a beszélők együttesen vesznek részt (Schiffrin 1987: 49). A diskurzusjelölők funkciói a következők: 1. kontextuális koordinátorként helyezik el a megnyilatkozásokat a diskurzus valamely szintjén, 2. a megnyilatkozásokat a beszélőhöz, a

hallgatóhoz, vagy mindkét résztvevőhöz kötik, 3. a megnyilatkozásokat korábbi vagy későbbi diskurzushoz rögzítik. A diskurzusjelölők egyszerre több szinten is működhetnek.

3. Diskurzusjelölők és kötőelemek tolmácsolt szövegekben

Blakemore és Gallai (2014) angol–olasz és portugál–olasz rendőrségi tolmácsolásban vizsgálta a diskurzusjelölők használatát. Megfigyelésük szerint a tolmácsok diskurzusjelölőket adnak hozzá a tolmácsolt megnyilatkozáshoz annak céljából, hogy az eredeti beszélő nézőpontját jelenítsék meg. Ugyanezen hozzáadott elemek a tolmács megnyilatkozás-feldolgozásából erednek, ám nem a tolmács, hanem az eredeti beszélőhöz kötődnek, azt az illúziót kelteve, hogy a hallgatóság a beszélő, nem a tolmács „hangját" hallja (Blakemore és Gallai 2014: 118).

Defrancq, Plevoets és Magnifico (2015) intermodális kutatása eredeti francia európai parlamenti felszólalások írott és beszélt nyelvi változatának angol és holland fordításának és szinkrontolmácsolásának kötőelemeit vizsgálta. Az eredmények szerint a tolmácsolt szövegek megnövelik a kötőelemek gyakoriságát. A normalizált szövegekben többféle megoldás található, több kötőelemet törölnek, ugyanakkor több elemet is adnak hozzá a szöveghez, mint a fordítások. Ezek a hozzáadások a tolmácsolt szöveg kohézióját többféle módon optimalizálják: a tagmondatok közötti kapcsolatokat teszik explicitté, a tolmácsolásból kihagyott szegmensek után stabilizálják a szöveg kohézióját kötőelemmel jelölve a kihagyott szegmens előtt és után közvetlenül elhelyezkedő szegmensek kapcsolatát, optimalizálva a tolmácsolt szöveg feldolgozhatóságát (Defrancq, Plevoets és Magnifico 2015). Az angol és a holland szinkrontolmácsolás megnövelte a kötőelemek gyakoriságát, azonban a holland tolmácsolás jobban, mint az angol. Vanhauwaert (2016) német–holland tolmácsolási irányban elemezte a kötőelemek használatát. Hasonló módon azt találta, hogy a tolmácsok a nyelvi problémamegoldás részeként toldanak be kötőelemeket a tolmácsolt szövegbe.

Ezek az eredmények arra utalnak, hogy a tolmácsolt szövegek eltérően jelölhetik logikai szerkezetük explicitségét kötőelemek segítségével, mind az összehasonlítható szö-

vegekhez, mind a forrásszövegekhez viszonyítva. Azonban ez az eltérés függ a kötőelemek típusától, és minden valószínűség szerint a tolmácsolási nyelvpártól is. A franciáról hollandra szinkrontolmácsolt szövegek kötőelemeinek 33,98%-a hozzáadás, a franciáról angolra tolmácsolt szövegekben ez 32,07% (Defrancq, Plevoets és Magnifico 2015). A németről hollandra tolmácsolt szövegek esetében ez az arány 21%, amelyeknek többségét additív elemek alkotják (pl. a holland *ook* = 'is') (Vanhauwaert 2016).

Defrancq (2016) a *well* diskurzusjelölő használatát vizsgálta eredeti és (olasz, spanyol és francia forrásnyelvekről) tolmácsolt európai parlamenti felszólalásokban. Az eredmények szerint a tolmácsok nem használják ritkábban e diskurzusjelölőt, és a *well* tolmácsolt előfordulásainak többségét (85%) nem motiválta párhuzamos forrásnyelvi elem.

Desmarets (2017) angol diskurzusjelölők (*well*, *now*, *so*) használatát hasonlította össze brit parlamenti, valamint eredeti és tolmácsolt angol nyelvű európai parlamenti felszólalásokban. Az elemek gyakoriságára és funkcionális spektrumára kiterjedő eredményei szerint a három korpusz viszonylatában a *now* használata tért el a leginkább az eredeti és a tolmácsolt szövegek között, azonban összességében az elemek használta magas fokú átfedést mutatott. Vagyis egyes elemek bizonyos funkciói (pl. a *now* témaváltó és saját meglátást bevezető használata) prominensebbek lehetnek tolmácsolt vagy nem tolmácsolt beszédekben egymáshoz viszonyítva, ezek a különbségek azonban többnyire nem számottevők.

Götz (2017a) európai parlamenti felszólalásokban vizsgálta a *vajon* fordítását az angol–magyar nyelvpárban. Az eredmények arra utalnak, hogy a *vajon* ritkább a magyarra tolmácsolt, mint a fordított szövegekben. A 75 megvizsgált felszólalásból, amely tartalmazza fordított változatában a *vajon*-t, az elem csupán hét előfordulással található a felszólalások magyar tolmácsolásában. Megjegyzendő, hogy a vizsgált angol felszólalások közül tízben a magyar tolmácsolás kihagyta az adott szegmenst, amelyben a magyar fordított szöveg használta a *vajon*-t. Ez arra utalhat, hogy a tolmácsolásban bizonyos elemek ritkábbak lehetnek azonos – vagy közel azonos – forrásszövegek tolmácsolt és fordított változataiban.

4. A kutatás bemutatása

A jelen tanulmány két korpuszt használ: 1. párhuzamos angol–magyar tolmácsolt, és 2. eredeti magyar beszédekből álló összehasonlítható korpuszt. Mindkét korpusz szövegei a Pannónia korpusz (l. Robin et al. 2017) intermodális alkorpuszának (l. Götz 2017b; Varga 2018) állományából kerültek ki és európai parlamenti felszólalások foglalnak magukban.

4.1 Korpuszok

A korpuszok adatait az *1. táblázat* tartalmazza. A tolmácsolt magyar korpusz (TMK) és az eredeti magyar korpusz (EMK) hosszúsága 15 másodperc híján megegyezik, hezitációk nélkül mért szószáma viszont eltér, aminek oka a felszólalók és tolmácsok eltérő szövegprodukciója. Továbbá a felszólalásokat egyenlő arányban tolmácsolták női és férfi szinkrontolmácsok, az eredeti magyar felszólalások korpusza szintén egyenlő arányban tartalmazza a női és férfi felszólalók szövegeit.

1. táblázat

A korpuszok

	Angol eredeti	Tolmácsolt magyar korpusz (TMK)	Eredeti magyar korpusz (EMK)
Szószám	7632 szó (hezitáció nélkül)	5335 szó (hezitáció nélkül)	6287 szó (hezitáció nélkül)
Hossz	1’ 4” 30”’		1’ 4” 45”’
Felszólalások száma	32		28

4.2 Kontrasztív és gyakorisági elemzés

Az angol és magyar szövegekből párhuzamosított, szegmensekre bontott korpusz készült. Minden olyan elemet motiválatlannak értékelve, amelynek használatára az angol szöveg párhuzamos forrásnyelvi elemei (l. részletesen 4.1) nem adnak direkt impulzust. Az elemek azonosítására a tolmácsolt és nem tolmácsolt szövegek alapján került sor, figyelembe

véve a Dér és Markó (2010) által listázott elemeket. Az ezek alapján létrehozott lista (l. Függelék) elemeinek gyakoriságát automatikus keresés mérte a tolmácsolt és nem tolmácsolt magyar szövegeket tartalmazó fájlokon. A gyakoriságot az abszolút, vagyis a teljes tokenszámot mutató számadat mellett normalizált, jelen esetben tíz percre átlagolt gyakoriság is szemlélteti. Az elemzések a *Wordsmith Tools 7* programmal készültek.

4.3 Hipotézisek

Számos kutatás (pl. Defrancq, Plevoets és Magnifico 2015, Defrancq 2016) figyelte meg eltérő nyelvek közötti közvetítés esetében a diskurzusjelölők és kötőelemek gyakori betoldását. Ennek alapján a jelen kutatás az alábbi hipotéziseket állította fel:

1. A tolmácsolt célnyelvi szövegekben található diskurzusjelölő és kötőelemek többsége betoldás eredménye.
2. A tolmácsolt célnyelvi szövegek több diskurzusjelölőt és kötőelemet tartalmaznak a nem tolmácsoltak forrásnyelvi szövegeknél.
3. A tolmácsolt célnyelvi és a nem tolmácsolt forrásnyelvi szövegek diskurzusjelölőinek és kötőelemeinek gyakorisága szignifikánsan eltér.

4.4 Hozzáadott elemek azonosítása

A jelen elemzés betoldott elemnek tekint minden olyan, a tolmácsolt szövegekben megjelenő diskurzusjelölőt és kötőelemet, amelyet a forrásszöveg nem motivál, a párhuzamos elemmel rendelkező jelölők használatát viszont motiváltnak tekinti (vö. Defrancq 2016 „triggered” és „untriggered” fogalmaival). A motiváció hiányát okozhatja a célnyelvi elemmel párhuzamos forrásnyelvi elem hiánya, vagy az is, hogy ugyan a forrásnyelvi szöveg rendelkezik az adott szegmensek viszonyát jelző elemmel, a magyar tolmácsolt

szövegben megjelenő elem ezt a viszonyt módosítja, és új, illetve a forrásnyelvitől eltérő kapcsolatot hoz létre a tolmácsolt magyar szövegben.

Érdemes tehát különbséget tenni kétféle betoldás között: míg bizonyos betoldások módosítják – például explicitálják vagy egyszerűsítik –, a forrásnyelvi szövegben már meglévő, a forrásnyelvi tagmondatok közötti logikai kapcsolatokat, addig mások olyan szerkezeteket hoznak létre, amelyeket a forrásszöveg nem tartalmaz, illetve nem is motivál sem saját szerkezeteivel, sem diskurzusviszonyaival. Az alábbi példák a *Magyar Intermodális Korpusz* átírási konvenciói (Götz 2017b) alapján jelennek meg.

Az (1) példában a tolmácsolt szövegbe betoldott *és* kötőszó egyszerűsíti a forrásnyelvi angol tagmondatok közt fennálló viszonyt, azzal nem párhuzamos. Ennek okán az *és*-t az elemzés hozzáadott kötőszónak tekinti, mivel a forrásszöveg tartalmaz ugyan diskurzusjelölő elemet, amely a magyar tolmácsolt szövegben érvényesülő viszonyt nem motiválta.

(1) So we expect the industry to take the prime responsibility for learning lessons rather than relying on the regulator to tell it what to do.

(1a) Azt várjuk tehát az érintett iparágaktól, hogy vonják le a tanulságokat, vállaljanak felelősséget *és* ne várják a szabályozótól, hogy mondja meg, hogy mit kell tennie.

Ebből a szempontból az (1) és (2) és (3) példában található betoldás eltérő természetű. A (2) mondatban a *hogy*-gyal jelölt alárendelt tagmondati kapcsolat jelen van az angol forrásnyelvi tagmondatok között is. A *hogy* ezekben a tolmácsolt tagmondatokban betoldásnak tekinthető, mivel forrásnyelvi elem nem motiválta a használatát.

(2) Therefore as a Parliament we need to be clear on what our priorities are and be prepared to fight for them.

(2a) ezért az Európai Parlamentnek tisztáznia kell, *hogy* mik a prioritásaink és hogy ezekért hajlandóak vagyunk-e harcolni.

(3) She thought nobody could be so stupid *as* to do what I've done

(3a) azt gondolta, hogy hogy lehettem ilyen ö hülye, *hogy* ezt tettem

A (2) példában a *hogy* kötőszó használata opcionális. Ugyanakkor a jelen vizsgálat olyan eseteket is a betoldások közé sorol kontrasztív megfontolás alapján, amelyeket a forrásnyelv nem motivál, a magyarban azonban a betoldott elem nem elhagyható. Ezt mutatja a (4) példában látható eset.

(4) spend 30 million pounds to comply with EU regulations

(4a) 30 milliárd (sic) fontot adnak Nagy-Britanniában, *hogy* betartsák az EU szabályozást

(4b) *30 milliárd (sic) fontot adnak Nagy-Britanniában betartsák az EU szabályozást

A *hogy* betoldásával a tolmács lehetővé teszi a már elhangzott, külön szegmenst alkotó részek beszövését a magyar diskurzusba, valamint a még nem hallott szegmensek nyelvtani beszövését is a már megkezdett tagmondati kapcsolatokba. Az (5) példában a tolmács *hogy*-os tagmondatszerkezettel alárendelt tagmondatként fűzte be a magyar diskurzusba az (5/3) forrásnyelvi mondat tartalmát, átugorva az (5/2)-t.

(5/1) Rather than focusing on all the other stuff we focused on this week.

(5/2) Let's make sure the EU works better.

(5/2a) Annyi mindenre koncentráltunk ezen a héten biztosítsuk sokkal inkább,

(5/3) Let's focus on reducing the administrative burden for small and medium-sized businesses

(5/3a) *hogy* csökkenjen az adminisztretív>adminisztratív< teher a kis- és középvállalatok vállán

Ez a fajta betoldás tehát egy még korábbi tolmácsolt szegmensben megkezdett vagy hozzáadott tagmondati, illetve logikai kapcsolatot folytat. További különbséget jelent, hogy ezek a motiválatlan hozzáadások beleillenek-e a tolmácsolt célszöveg szerkezetébe – és így annak kohéziós folytonosságát segítik-e, vagy éppenséggel megtörik-e azt. A (6) példában az *ugyanis* betoldása lehetővé teszi a magyar (6/1a) és (6/2a) szegmens összekapcsolását, azonban a logikai kapcsolat a két tagmondat között nem optimális.

(6/1) This damages the European Union when we don't take firm action to help.

(6/1a) Egész egyszerűen arról van szó, hogy mi magunk sem lépünk fel kellő eréllyel ezekben az ügyben [ezekben az ügyekben],

(6/2) It is the mantra of the European Commission *after all* to think small first.

(6/2a) *ugyanis* ö az Európai Bizottságnak ö lenne itt feladata.

Az *ugyanis* kötőszó használatát az *after all* is motiválhatta, azonban annak nem felel meg teljesen, az ugyanis a két összekötött szegmens közötti kontrasztív viszony helyett kiegészítő kapcsolatot létesít.

5. Eredmények

A jelen elemzés megvizsgálta az eredeti és tolmácsolt szövegek diskurzusjelölő- és kötőelemkészletét, valamint azt, hogy a tolmácsolt szövegek esetében ezek használatát az

angol eredeti felszólalások mennyiben motiválták. A *2. táblázat* az elemek abszolút gyakoriságát mutatja, valamint azt, hogy a tolmácsolt elemek mekkora hányadát motiválták forrásnyelvi elemek. Az elemek teljes listája a Függelékben található.

2. táblázat

A korpuszok leggyakoribb elemei és a tolmácsolt korpusz elemeinek forrásnyelvi motiváltsága

	TMK	EMK	Hozzáadott elemek a TMK-ban	
			No.	%
Összes elem	927	997	497	53,61

A vizsgált elemek abszolút gyakoriságát tekintve a nem tolmácsolt szövegek több diskurzusjelölőt és kötőelemet használnak. A tolmácsolt szövegek elemeit összességében 53,61%-át nem motiválták a forrásnyelvi szövegekben megtalálható párhuzamos elemek, vagyis a tolmácsolt szövegekben a legtöbb elem hozzáadás eredménye. Összességében 430 elem használatát motiválták angol forrásnyelvi elemek és 497 elemet toldottak be a tolmácsok forrásnyelvi megfelelő nélkül.

Mivel a hozzáadott elemek aránya nem ismert más angol–magyar európai parlamenti szinkrontolmácsolást elemző kutatások, vagy egyéb tolmácsolási esemény szövegprodukciójának esetében, az angol–magyar tolmácsolt szövegek elemhasználatát illetően nem lehet általánosítani a jelen eredményeket. Továbbá megjegyzendő, hogy a két korpusz között az egyes elemek gyakorisága, valamint a magyar tolmácsolt szövegek elemeinek forrásnyelvi motiváltsága is igen változatos (l. Függelék).

A *3. táblázat* a motivált és nem motivált tolmácsolt elemek eltérésén végzett t-próbát mutatja. A motivált és nem motivált tolmácsolt elemek mennyiségi eltérése nem szignifikáns. Az abszolút gyakoriság nem veszi figyelembe a korpusz eltérő szószámát, vagy ad képet az elemek átlagolt használatáról, így nem megfelelő alap a gyakoriság összevetéséhez.

3. táblázat

A motivált és nem motivált elemek abszolút gyakoriságának t-próbája a tolmácsolt korpuszban ($p \leq 0,05$)

	t-érték	p-érték
Összes elem	-0,377	0,353

Ezért a *4. táblázat* a diskurzusjelölők és kötőelemek normalizált gyakoriságát mutatja a hezitáció nélkül számított szószám és a korpuszok időbeli hosszának viszonylatában.

4. táblázat

A diskurzusjelölők és kötőelemek normalizált gyakorisága

	Gyakoriság	
	/1000 szó	/10 perc
TMK	173,76	143,72
EMK	158,58	153,98

Amíg a normalizált szószám arányában mért gyakoriság magasabb a tolmácsolt korpuszban, addig a beszédidőre normalizált gyakoriság magasabb az eredeti szövegekben. Ez az eredmény hangsúlyozza a gyakoriság többféle mérésnek szükségességét az adatok megfelelő értelmezésének érdekében. Ezek a gyakoriságok utalhatnak arra, hogy amíg a tolmácsok rövidebb szövegeket produkálnak, ezek kötőelem- és diskurzusjelölő-tartalma magasabb, azonban a beszédidő viszonylatában a különbség csökken, és a jelen esetben az így mért gyakoriság magasabb az eredeti, mint a tolmácsolt szövegekben. Ezen feltételezések bizonyítása azonban további, nagyobb mintán végzett kutatást igényel.

A gyakoriságbeli eltérések nem szignifikánsak a vizsgált korpuszokban. Az 1. ábra a korpuszok 15 leggyakoribb elemének tíz percre számolt gyakoriságát szemlélteti. Ebben a mintában a tolmácsolt szövegek az elemek többségének esetében (*hogy*, *azt*, *ezt*, *ha*, *úgy*, *annak*, *mert*, *hogyha*) nagyobb gyakoriságot mutatnak, mint a nem tolmácsolt szövegek.

5. táblázat

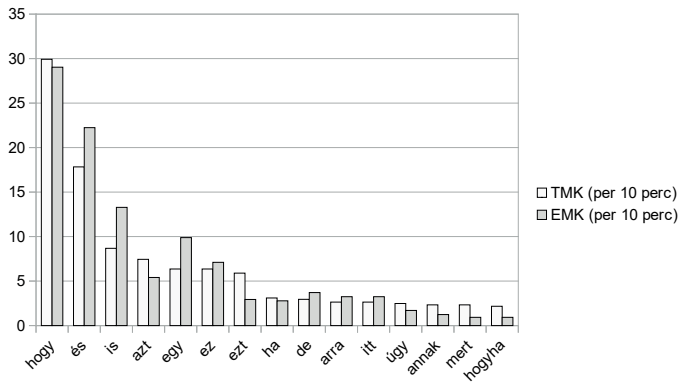
A diskurzusjelölők és kötőelemek gyakoriságának eltérésén végzett t-próba a tolmácsolt és nem tolmácsolt korpuszban ($p \leq 0,05$)

	t-érték	p-érték
Gyakoriság per 1000 szó	-0,265	0,396
Gyakoriság per 10 perc	-0,191	0,424

Mindkét korpusz esetében a leggyakoribb 15 elem a kohézív kötőelemek csoportjából került ki. Ezek a diskurzuskohéziót a szegmensek közötti folytonosság biztosításával szolgálják.

1. ábra

A tolmácsolt magyar korpusz (TMK) és eredeti magyar korpusz (EMK) leggyakoribb 15 elemének normalizált gyakorisága (per 10 perc)



6. Következtetések és összegzés

A jelen kutatás a szinkrontolmácsolt magyar diskurzus diskurzusjelölőit és kötőelemeit kutatta, gyakorisági és kontrasztív elemzéseket végezve. Specifikusan arra kereste a választ, hogy a forrásnyelvi felszólalások mennyiben motiválják a tolmácsolt magyar szövegek elemhasználatát, valamint, hogy mutatkozik-e különbség a tolmácsolt és nem tolmácsolt szövegek elemszáma között, és amennyiben igen, ez a különbség szignifikáns-e.

A vizsgálat eredményei szerint a tolmácsolt szövegekben található diskurzusjelölők és kötőelemek többsége hozzáadás eredménye (1. hipotézis), mivel az elemek 53,61%-a hozzáadás eredménye a magyar tolmácsolt diskurzusban. Továbbá a jelen kutatás eredményei szerint a tolmácsolt szövegek ritkábban használnak diskurzusjelölőket és kötőelemeket a nem tolmácsoltaknál a tíz percre normalizált gyakoriság szerint, azonban a szószámra normalizált gyakoriság az ellenkezőjét mutatja, vagyis a tolmácsolt szövegek 1000 szóra átlagolva több elemet tartalmaznak, mint a nem tolmácsoltak. Így azt nem lehet kijelenteni, hogy a tolmácsolt szövegekben ennek az elemnek a használata gyakoribb (2. hipotézis). A gyakoriságban mért eltérések nem szignifikánsak (3. hipotézis). Ezzel igazolódott az 1. hipotézis, a 2 és a 3. hipotézis ellenben nem.

A tolmácsok beszédprodukciója eltér a felszólalókétól, bizonyos különbségekre tehát emiatt számítani lehet, azonban a gyakorisági méréseknek éppen ezért figyelembe kell venni az idő és a szószám viszonylatában mért adatok eltéréseit. Amennyiben a szinkrontolmácsolt európai parlamenti diskurzus tulajdonságaira vagyunk kíváncsiak, félrevezető lenne a felszólalások számát és hosszúságát ignorálni, és csak a szószám alapján határozni meg a használt korpuszok nagyságát. Amennyiben így tennénk, jogosan merülne fel a kérdés, hogyan lehetne például másfél órányi tolmácsolást csupán negyven percnyi eredeti felszólalással egybevetni. Az ilyen módszertani problémák kiküszöbölésére hasznos lehet többféle normalizált gyakorisági mérést végezni.

A tolmácsolt diskurzus diskurzusjelölő- és kötőelemkészletének valamivel több, mint felében a vizsgált korpuszban a forrásnyelvi párhuzamos elemek nem motiválják a célnyelvi előfordulást. Ugyan ez az elemek többségét teszi ki, a forrásnyelvi motiváció nem elhanyagolható. Más forrásnyelv magyar tolmácsolásának vizsgálatával megválaszolható lenne a kérdés, a magyar tolmácsolt célnyelvi szövegek bizonyos elemhasználati szintre törekszenek, tehát nyelvpártól függetlenül hasonló elemgyakoriság jellemzi a tolmácsolt magyar célnyelvi szövegeket, vagy ehelyett a forrásnyelvi elemek bizonyos esetekben a forrásnyelvi kapcsolatok bizonyos százalékát explicitté vagy implicitté teszik a tolmácsolt szövegekben. Feltehető, hogy a kohéziós és logikai explicitséget magasabb szinten jelölő nyelv magyarra tolmácsolása is magasabb gyakoriságot mutat ezen elemek

tekintetében, mint egy kevésbé explicit viszonyokat preferáló forrásnyelv. Valószínűleg a hozzáadott elemek aránya, vagy a megfeleltetett elemek aránya ezen nyelvpárok esetében is hasonló lenne, ha a magyar hivatalos nyelvhasználati norma irányítaná a tolmácsolási beszédprodukciót, nem pedig egyéb tényezők (pl. a tolmácsolási beszédprodukció kognitív folyamatai, kontrasztív nyelvi eltérések, tolmácsolási univerzálék, egyéni tolmácsolási preferencia, adott tolmácsolási esemény jellemzői stb.). A jelen eredmények alapján ezekre a kérdésekre csak korlátozottan lehet reflektálni.

A két korpusz elemkészletének és elemgyakoriságának hasonlósága alól csupán néhány elem képez kivételt. A nem tolmácsolt szövegek például 20%-kal több *és*-t, valamint 35%-kal több *is*-t használnak. Ezt részben lehet magyarázni a forrásnyelvi diskurzusszerkezet módosításával, szegmensek kihagyásával, valamint az *is* esetében az angol–magyar kontrasztív különbségekkel. Az *és* esetében a motiválatlan elemek száma a tolmácsolt szövegekben 37,39%, az *is*-nél 85,71%. Ez arra utalhat, hogy az *is* használatára kevesebb impulzust kap a tolmács, ennek ellenére él vele. Mindemellett, ugyan a tolmácsolt és nem tolmácsolt szövegek elemkészlete nagy átfedést mutat, néhány elem kizárólag egyik vagy másik korpuszban jelentkezik. Egyes elemek hiánya a korpuszok kis méretének tudható be, mások azonban a két szóbeli diskurzus eltérő jegyeire vezethetők vissza. A *nemtom* elem például csupán a tolmácsolt korpuszban fordul elő, az *amely* vonatkozó névmás pedig megközelítőleg négyszer olyan gyakori a nem tolmácsolt, mint a tolmácsolt korpuszban. Ezek a különbségek a tolmácsolt diskurzus kevésbé formális, inkább beszélt jellegére utalhatnak. Ez azonban további kutatást igényel, mivel a jelen vizsgálat eredményei alapján annak korlátozott keretei miatt messzemenő következtetések nem vonhatók le.

Irodalom

Aijmer, K. 2007. The meaning and functions of the Swedish discourse marker alltså – Evidence from translation corpora. *Catalan Journal of Linguistics* Vol. 6. 31–59. https://doi.org/10.5565/rev/catjl.123

Blakemore, D., Gallai, F. 2014. Discourse markers in free indirect style and interpreting. *Journal of Pragmatics* Vol. 60. 106–120. https://doi.org/10.1016/j.pragma.2013.11.003

Defrancq, B., Plevoets, K., Magnifico, C. 2015. Connective Items in Interpreting and Translation: Where Do They Come From? In: Romero-Trillo, J. (ed.) *Yearbook of Corpus Linguistics and Pragmatics 2015, Vol. 3*. Cham: Springer. 195–222. https://doi.org/10.1007/978-3-319-17948-3_9

Defrancq, B., 2016. Well, interpreters… a corpus-based study of a pragmatic particle used by simultaneous interpreters. In: Corpas Pastor, G., Seghiri, M. (eds) *Corpus-Based Approaches to Translation and Interpreting*. Peter Lang D, Bern. 105–128. https://doi.org/10.3726/b10354/16

Dér, Cs. I., Markó, A. 2010. A Pilot Study of Hungarian Discourse Markers. *Language and Speech* Vol. 53. No. 2. 135–180. https://doi.org/10.1177/0023830909357162

Desmarets, S. 2017. *Simultaneous Interpreting as a Situated Speech Act.* Ghent: Ghent University. Elérhető: https://lib.ugent.be/fulltxt/RUG01/002/348/997/RUG01-002348997_2017_0001_AC.pdf.

Fraser, B. 1990. An approach to discourse markers. *Journal of Pragmatics* Vol. 14. No. 3. 383–398. https://doi.org/10.1016/0378-2166(90)90096-v

Fraser, B. 1996. Pragmatic markers. *Pragmatics* Vol. 6. No. 2. 167–190. https://doi.org/10.1075/prag.6.2.03fra

Fraser, B. 1999. What are discourse markers? *Journal of Pragmatics* Vol. 31. No. 7. 931–952. https://doi.org/10.1016/s0378-2166(98)00101-5

Fraser, B. 2015. The combining of Discourse Markers – A beginning. *Journal of Pragmatics* Vol. 86. 48–53. https://doi.org/10.1016/j.pragma.2015.06.007

Götz A. 2017a. Az első magyar intermodális korpusz bemutatása. Kutatási lehetőségek. *Argumentum* Vol. 13. 126–139.

Götz, A. 2017b. Translating doubt: the case of the Hungarian discourse marker *vajon*. In: Wachowski, W., Kövecses, Z., Borodo, M. (eds) *Micro-Scale Perspectives on Cognition, Translation and Cross-Cultural Communication.* Oxford: Peter Lang. 125–146.

Hale, S. B. 2010. *The Discourse of Court Interpreting: Discourse Practices of the Law, the Witness, and the Interpreter*. Amsterdam: John Benjamins. https://doi.org/10.1075/btl.52

Halliday, M. A. K., Matthiessen, C. M. I. M. 2014. *Halliday's Introduction to Functional Grammar*. Milton Park: Routledge. https://doi.org/10.4324/9780203431269

Németh T. E. 1998. A hát, így, tehát, mert kötőszók pragmatikai funkciójának vizsgálata. *Magyar Nyelv* Vol. 94. No. 3. 324–331.

Robin E., Götz A., Pataky É., Szegh H. 2017. Translation Studies and Corpus Linguistics: Introducing the Pannonia Corpus. *Acta Universitatis Sapientiae, Philologica* Vol. 9. No. 3. 99–116. https://doi.org/10.1515/ausp-2017-0032

Schiffrin, D. 1987. *Discourse Markers*. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511611841

Schirm A. 2009. Partikula és/vagy diskurzusjelölő? In: *Diskurzus a grammatikában – grammatika a diskurzusban. Segédkönyvek a nyelvészet tanulmányozásához 88.* Budapest: Tinta Könyvkiadó. 304–311.

Traugott, E. C. 2007. Discussion article: Discourse markers, modal particles, and contrastive analysis, synchronic and diachronic. *Catalan Journal of Linguistics* No. 6. 139–157. https://doi.org/10.5565/rev/catjl.128

Vanhauwaert, G. 2016. *Die* Übertragung *von Kohäsiven Junktionen beim Simultandolmetschen – Ein Korpusbasierter Vergleich Deutscher Quelltexte mit Niederländischen Zieltexten*. MA-szakdolgozat. Ghent: Ghent University.

Varga D. Á. 2018. *Egy intermodális korpusz építésének gyakorlati tapasztalatai*. Előadás. Elhangzott: Fordításkutatás – Fordítóképzés. Budapest: Pázmány Péter Katolikus Egyetem. 2018. május 24.

Internetes hivatkozások

Lexical Analysis Software 2018. Wordsmith Tools 7. Oxford University Press. elérhető: https://www.lexically.net/wordsmith/

Függelék: Elemkészlet

	TMK	EMK	Hozzáadott elemek a TMK-ban		Gyakoriság			
					TMK		EMK	
			No.	(%)	/1000 szó	/10 perc	/1000 szó	/10 perc
(1) hogy	193	188	135	69,95%	36,18	29,92	29,9	29,03
(2) és	115	144	43	37,39%	21,56	17,83	22,9	22,24
(3) is	56	86	48	85,71%	10,5	8,68	13,68	13,28
(4) azt	48	35	20	41,67%	9	7,44	5,57	5,41
(5) egy	41	64	8	19,51%	7,69	6,36	10,18	9,88
(6) ez	41	46	19	46,34%	7,69	6,36	7,32	7,1
(7) ezt	38	19	17	44,74%	7,12	5,89	3,02	2,93
(8) ha	20	18	10	55,00%	3,75	3,1	2,86	2,78
(9) de	19	24	1	5,26%	3,56	2,95	3,82	3,71
(10) arra	17	21	5	29,41%	3,19	2,64	3,34	3,24
(11) itt	17	21	13	76,47%	3,19	2,64	3,34	3,24
(12) úgy	16	11	1	6,25%	3	2,48	1,75	1,7

	TMK	EMK	Hozzáadott elemek a TMK-ban		Gyakoriság			
					TMK		EMK	
			No.	(%)	/1000 szó	/10 perc	/1000 szó	/10 perc
(13) annak	15	8	6	40,00%	2,81	2,33	1,27	1,24
(14) mert	15	6	8	33,33%	2,81	2,33	0,95	0,93
(15) hogyha	14	6	4	35,71%	2,62	2,17	0,95	0,93
(16) ugyanakkor	13	2	5	38,46%	2,44	2,02	0,32	0,31
(17) ezeket	12	9	9	75,00%	2,25	1,86	1,43	1,39
(18) őket	12	0	4	33,33%	2,25	1,86	0	0
(19) ennek	11	20	6	54,55%	2,06	1,71	3,18	3,09
(20) már	11	17	8	72,73%	2,06	1,71	2,7	2,63
(21) akkor	12	15	8	66,67%	2,25	1,86	2,39	2,32
(22) csak	10	20	5	50,00%	1,87	1,55	3,18	3,09
(23) azért	10	4	9	90,00%	1,87	1,55	0,64	0,62
(24) most	9	15	5	55,56%	1,69	1,4	2,39	2,32
(25) tehát	9	9	6	66,67%	1,69	1,4	1,43	1,39
(26) vagy	9	5	5	55,56%	1,69	1,4	0,8	0,77
(27) erre	8	3	5	62,50%	1,5	1,24	0,48	0,46
(28) szerintem	8	0	1	12,50%	1,5	1,24	0	0
(29) ezért	7	13	4	57,14%	1,31	1,09	2,07	2,01
(30) ami	7	9	4	57,14%	1,31	1,09	1,43	1,39
(31) azonban	7	9	6	85,71%	1,31	1,09	1,43	1,39
(32) ezzel	7	5	4	57,14%	1,31	1,09	0,8	0,77
(33) még	6	21	4	66,67%	1,12	0,93	3,34	3,24
(34) pedig	6	11	6	100,00%	1,12	0,93	1,75	1,7
(35) -e	6	5	1	16,67%	1,12	0,93	0,8	0,77
(36) mint	5	15	1	20,00%	0,94	0,78	2,39	2,32
(37) hát	5	4	5	100,00%	0,94	0,78	0,64	0,62
(38) ugyanis	5	3	5	100,00%	0,94	0,78	0,48	0,46
(39) hiszem	5	2	3	60,00%	0,94	0,78	0,32	0,31
(40) viszont	5	1	4	80,00%	0,94	0,78	0,16	0,15
(41) aztán	5	0	1	80,00%	0,94	0,78	0	0
(42) amely	4	17	2	75,00%	0,75	0,62	2,7	2,63

	TMK	EMK	Hozzáadott elemek a TMK-ban		Gyakoriság			
					TMK		EMK	
			No.	(%)	/1000 szó	/10 perc	/1000 szó	/10 perc
(43) illetve	4	7	1	25,00%	0,75	0,62	1,11	1,08
(44) így	4	5	0	0,00%	0,75	0,62	0,8	0,77
(45) éppen	4	3	4	100,00%	0,75	0,62	0,48	0,46
(46) ugye	4	3	3	75,00%	0,75	0,62	0,48	0,46
(47) hiszen	3	8	2	66,67%	0,56	0,47	1,27	1,24
(48) azokat	3	4	3	100,00%	0,56	0,47	0,64	0,62
(49) éppen ezért	3	1	3	100,00%	0,56	0,47	0,16	0,15
(50) annak ellenére	3	0	0	0,00%	0,56	0,47	0	0
(51) talán	2	8	0	0,00%	0,37	0,31	1,27	1,24
(52) bár	2	3	1	50,00%	0,37	0,31	0,48	0,46
(53) nos	2	3	2	100,00%	0,37	0,31	0,48	0,46
(54) noha	2	1	1	50,00%	0,37	0,31	0,16	0,15
(55) ahhoz	1	4	1	100,00%	0,19	0,16	0,64	0,62
(56) ehhez	1	3	1	100,00%	0,19	0,16	0,48	0,46
(57) mégis	1	3	1	100,00%	0,19	0,16	0,48	0,46
(58) s	1	2	0	0,00%	0,19	0,16	0,32	0,31
(59) sőt	1	2	1	100,00%	0,19	0,16	0,32	0,31
(60) vajon	1	1	1	100,00%	0,19	0,16	0,16	0,15
(61) tényleg	1	0	1	100,00%	0,19	0,16	0	0
(62) ennek következtében	1	1	1	100,00%	0,19	0,16	0,16	0,15
(63) nemtom	1	0	1	100,00%	0,19	0,16	0	0
(64) tudniillik	1	0	1	100,00%	0,19	0,16	0	0
(65) ennek ellenére	2	0	1	50,00%	0,37	0,31	0	0
(66) egyébként	0	3	0	–	0	0	0,48	0,46
(67) vagyis	0	1	0	–	0	0	0,16	0,15

Egyetlen előfordulással szerepel kizárólag a TMK-ban: *ugyan*, *gyakorlatilag*.

Mariachiara Russo, Claudio Bendazzoli, Bart Defrancq

Making Way in Corpus-based Interpreting Studies

(Szingapúr: Springer, 2018. 215 pp. ISBN: 978-981-10-6198-1)

Benedek Enikő

benedek.eniko@btk.elte.hu

Eötvös Loránd Tudományegyetem

Fordítástudományi Doktori Program

A Defeng Li, Tolmács- és Fordítóképző Világszövetség (World Interpreter and Translator Training Association) elnöke szerkesztette *New Frontiers in Translation Studies* sorozat 2015-ös indulásától kezdve számos kötetet szentelt a korpuszalapú kutatások ismertetésének, ezen belül is több teljes kiadás a hazánkban kevésbé kutatott kínai–angol nyelvpárral foglalkozik. A 2018-as megjelenésű *Making Way in Corpus-based Interpreting Studies* szerkesztői – Russo, Bendazzoli és Defrancq – valamennyien korpuszalapú tolmácsoláskutatással foglalkoznak, és mindannyian tevékeny részt vállaltak az EPIC (az Európai Parlament Tolmácsolási Korpusza), illetve az EPICG (az Európai Parlament Tolmácsolási Korpusza – Gent) megalakulásában. A kiadvány megismerteti az olvasót a szakképzett tolmácsok nagy mennyiségű tolmácsolását tartalmazó adatbázissal, azaz a korpusznyelvészeti módszer szerint csoportosított korpuszokkal, amelyek segítségével mind a tolmácsolás folyamatát, mind eredményét vizsgálni lehet. (Részletesen bemutatja például az EPIC és a fordításokat is tartalmazó EPTIC – az Európai Parlament Fordítási és Tolmácsolási Korpuszát, illetve a NAIST-Korpuszt – a japán Nara Tudományos és Technológiai Intézet összeállítását.) A kötet ezen kívül Miriam Shlesinger 1998-as felhívását értelmezi újra és

ösztönzi a kutatói közösséget a korpuszalapú tolmácsolástudomány bővítésére, az épülő korpuszokhoz pedig részletes, gyakorlati segítséget nyújt. A kötet tizenegy tanulmányból áll. Az első két publikáció a korpuszalapú tolmácsolástudomány elméleti keretrendszerét adja meg, a továbbiak pedig a szakképzett tolmácsok produkciójának korpuszalapú elemzéséből levont következtetéseket mutatják be.

A kötet kezdő tanulmánya (Claudio Bendazzoli: *Corpus-based Interpreting Studies: Past, Present and Future Developments*) áttekinti a korpuszalapú kutatások történetét, a huszadik század közepétől Miriam Shlesinger 1998-as javaslatáig. Mérföldkőnek tekinti az izraeli kutató felhívását a korpuszalapú fordítástudomány alapján a korpuszalapú tolmácsolástudomány megalapítására. A felhívás a szerző szerint ismét aktuális, mivel az elmúlt évtizedek információrobbanása nyomán a mai adatmennyiséget kézzel már lehetetlen lenne vizsgálni, és a rendelkezésre álló tolmácsolási korpuszok még mindig kis méretűek olyan nagy referenciakorpuszokhoz viszonyítva, mint a Brit Nemzet Korpusz (Dembry és Love 2015). Éppen ezért fontos, hogy a gyakorló tolmácsok megfelelő nyelvészeti számítástechnikai eszközökkel felvértezve vállaljanak részt a korpuszépítésben.

A tanulmánykötet következő részében (Silvia Bernardini, Adriano Ferraresi, Mariachiara Russo, Camille Collard és Bart Defrancq: *Building Interpreting and Intermodal Corpora: A How-to for a Formidable Task*) a tolmácsolási és intermodális korpuszok építésének mikéntjével ismerkedhet meg az olvasó. Megtudhatjuk, hogy milyen nehézségekkel kell szembenézni a szinkrontolmácsolási és intermodális korpuszok építésekor, ugyanakkor a szerzők megoldásokat is kínálnak ezekre, amelyeket az Európai Parlament plenáris ülésein keletkezett szövegekből összeállított korpuszokban is alkalmaztak (EPIC, EPICG, EPTIC). A kutatás lépésről lépésre bemutatja a hangfelvételek transzkripcióját, a metaadatok felvételét, a szöveg egyes részeinek illusztrálását hangfelvételekkel és lemmainformációkkal, a szöveg-szöveg, illetve szöveg-audió/videó tartalom hozzárendelést, valamint a korpusz megfelelő indexálását is a keresőeszközök számára.

A harmadik tanulmány (Bart Defrancq, Koen Plevoets: *Over-uh-Load, Filled Pauses in Compounds as a Signal of Cognitive Load*) a szinkrontolmácsok kognitív terhelését vizsgálja összetett lexémák összetételénél fellépő *„ő-zéssel”* kitöltött szünetek

esetében. A korpusz az EPICG tolmácsolt és a CGN (Corpus Gespoken Nederlands) nem tolmácsolt holland szövegek gyűjteményéből áll, ezeknek a *bakeri* módszerekkel történő összehasonlításával a két kutató kiszűrte a tolmácsolás befolyását a kitöltött szünetekre. Az eredményeket a kognitív terhelés különböző szempontjai szerint elemezték. A tolmácsolt holland szöveg számottevően több kitöltött szünetet tartalmaz, mint a nem tolmácsolt szöveg, továbbá a szünetkitöltések sokkal gyakrabban fordulnak elő tolmácsolt szövegek esetében. Ez utóbbi azt jelenti, hogy a főként a szóösszetételek vezetnek a tolmács kognitív leterheltségéhez, a magasabb kognitív terhelés pedig a tolmácsok lexikai hozzáférésére is kihat. Feltételezhető valamilyen viszony továbbá a kognitív terhelés és a következtében létrejövő szünetkitöltések, valamint az összetett szavak összetételének sorrendjében található két nyelv közti különbségek között is.

Wang és Zou tanulmánya (*Exploring Language Specificity as a Variable in Chinese-English Interpreting. A Corpus-based Investigation*) az angol és kínai nyelvpárra jellemző tolmácsolási kutatások számát kívánja gyarapítani. Vizsgálatuk tárgya a kínairól angolra tolmácsolt szövegek szintaktikai aszimmetriája és a tolmácsok megoldásai az aszimmetria kiküszöbölésére. Vizsgálatuk középpontjában tehát a kínai nyelvben általában a szavak előtt álló bonyolult és hosszú jelzős szerkezetek angol nyelvű átváltásának módjai állnak. A vizsgált tolmácsolási szövegekben az elöl álló jelzős szerkezetek több mint 80%-a hátravetett jelzős szerkezetté, vagy vegyesen elöl és hátul álló jelzős összetétellé válik angolra tolmácsoláskor. A tolmács tehát különösen nagy kognitív terhelésnek van kitéve, ha két ennyire különböző szerkezetű nyelv között kell közvetítenie. A kutatás igazolhatóságát az összehasonlítható CEIPCC (Kínai Miniszterelnöki Sajtótájékoztatók) korpusz, és az Egyesült Államok Kormányának napi sajtótájékoztatói biztosították. A tanulmány felhívja a figyelmet a vizsgált nyelvpár jelentőségére, amely nem elhanyagolható tényező a tolmácsok beszédviselkedésének további kutatása során.

Guy Aston tanulmánya (*Acquiring the Language of Interpreters: A Corpus-based Approach*) a tolmácsolt diskurzust vizsgálja, anyagát az Európai Parlament ügymenete közben keletkezett korpusz szolgáltatja. Aston vizsgálata alapján a folyékonyan tolmácsoló szakemberek gyakran használnak visszatérő frazeologizmusokat. Mivel ezeket a

memória egyedi lexikai elemekként tárolja, amelyekhez bizonyos prozódikus elemek kapcsolódnak, a frazeologizmusok használatával a tolmács kognitív terhelése csökken, a formulák előhívása sokkal kevesebb erőfeszítést igényel, így szerepet játszanak a folyékony beszédprodukcióban. A szakirodalom alapján azonban a nem anyanyelvi beszélő frazeológiai repertoárja jóval kisebb, mint anyanyelvű kollégájáé, ezért már a tolmácsolás oktatásától kezdve nagy hangsúlyt kellene fektetni az idegen nyelvű frazeologizmusok elsajátítására és helyes használatuk megtanítására.

A fordításnyelvűség *(translationese)* mintájára megalkotott tolmácsnyelvűség *(interpretese)* témáját vizsgálja Kajzer-Wietrzny kutatása (*Interpretese vs. Non-native Language Use: The Case of Optional* That*)*. A fordított és a tolmácsolt szövegek esetében egyaránt nehéz általánosítható hasonlóságokat kimutatni. Bár több kísérlet is történt a tolmácsok szövegegyszerűsítésének tudományos feltérképezésére (Sandrelli és Bendazzoli 2005, Kajzer-Wietrzny 2012), ezek nem jártak teljes sikerrel, mivel a hasonlóságok csak bizonyos nyelvpárok esetében voltak jellemzőek. A *that* alárendelő, az explicitációt előrejelző kötőszó vizsgálatával azonban biztonsággal megállapítható, hogy eredeti vagy tolmácsolt szöveggel állunk szemben. Korábbi kutatások (Kajzer-Wietrzny 2012) már kimutatták, hogy az angolra tolmácsolt szövegekben gyakrabban fordul elő a *that*, mint az angolul keletkezett szövegekben, azonban Halverson (2003) kutatása alapján bizonyos, a fordításra vonatkozó jellegzetességek nemcsak tolmácsolt, hanem egyéb, többnyelvű környezetben létrejött szövegprodukcióra is jellemzőek lehetnek. E tanulmány a *that* kötőszó opcionális használatát vizsgálja az angolra szinkrontolmácsolt szövegekben, összehasonlítva az Európai Parlament angol anyanyelvű képviselőinek beszédével és az idegen anyanyelvű, de angol nyelven felszólaló képviselők beszédprodukciójával.

Nemtől függő beszédprodukciós jellegzetességeket vett górcső alá Russo az EPIC korpuszon (*Speaking Patterns and Gender in the European Parliament Interpreting Corpus: A Quantitative Study as a Premise for Qualitative Investigations*), a tolmácsok célnyelvi produkciójának hosszúságát vizsgálta a beszélő beszédmódja, a beszéd sebessége, a nyelvpár, és a beszédtéma tükrében. A szinkrontolmácsok angolról spanyolra és olaszra, valamint az olasz és spanyol nyelv között közvetítettek. A vizsgálat eredményei a követ-

kezők: a felolvasott angol nyelvű beszédeket a hölgy tolmácsok gyorsabban tolmácsolták spanyolra (143 szó/perc), mint a férfiak (124 szó/perc), a célnyelvi beszédprodukció pedig (16%-kal) rövidebb a férfiak esetében a forrásnyelvi beszédhez képest, mint a nők esetében (akik átlagosan 8%-ot rövidítettek az eredeti szöveghez képest). A célnyelvi produkció a politika, valamint eljárások és formalitások témakörben a férfiak körében 18%-kal, illetve 4%-kal rövidebb lett, mint a női tolmácsok esetében (ők 21% és 0,3%-kal rövidebben tolmácsoltak). Fordított lineáris tendencia volt megfigyelhető továbbá az előadó beszédsebessége és a tolmács produkciója tekintetében, amely főként a női tolmácsok esetében angolról spanyolra és olaszra tolmácsoláskor figyelhető meg. E kvantitatív tanulmány megfelelő alapot szolgáltat további, kvalitatív vizsgálatokhoz, amelyek segítségével a tolmács nemének szemantikai különbségekre gyakorolt hatását is vizsgálni lehetne.

A képes beszéd fordítása gyakran okoz gondot a fordítónak, a szinkrontolmácsnak pedig az azonnali időkeret miatt a fordítónál is bonyolultabb feladattal kell megbirkóznia a célnyelvi megfelelő megtalálásakor. A spanyol IMITES projekt 1135 metaforikus kifejezést és azok tolmácsolt változatát vizsgálja huszonhárom eredeti, valamint tolmácsolt olasz és spanyol nyelvű európai bizottsági ülésen elhangzott beszéd tükrében. Spinolo tanulmánya (*Studying Figurative Language in Simultaneous Interpreting: The IMITES* [Interpretación de la Metáfora Entre ITaliano y ESpanol] *Corpus*) a tolmácsok stratégiáját vizsgálja, valamint feltérképezi, hogy mely metaforikus kifejezések okozzák a legtöbb nehézséget a tolmácsolás során.

Dal Fovo vizsgálata (*European Union Politics Interpreted on Screen: A Corpus-based Investigation of the Third 2014 EU Presidential Debate*) az Európai Bizottság 2014-eselnökjelölti vitája két különböző televízióban elhangzott tolmácsolását hasonlítja össze: az Európai Bizottság tolmácsainak olasz célnyelvi produkcióját az Eurovision csatorna adásában, valamint az olasz Rainews24 csatorna közvetítését, amely a saját tolmácsait alkalmazta a feladatra. A kutatás feltérképezi, hogy a forrásnyelvi diskurzus mely elemei tekinthetők információs szórakoztatásnak *(infotainment)*, azaz mely elemek ötvözik a politika és televíziós nyelvhasználat sajátosságait, és ezen elemeket hogyan tolmácsolják a két különböző televíziós adásban. További kutatási kérdéseket vet fel a tolmá-

csolt szöveg minőségének és megfelelőségének megítélése, amely a televíziós elnökjelölti vitához hasonló hibrid típusú diskurzusban nem feltétlenül áll teljes összhangban az uniós tolmácsokkal szemben támasztott szakmai elvárásokkal.

Sandrelli tanulmánya (*Interpreter-Mediated Football Press Conferences: A Study on the Questioning and Answering Strategies*) a 2008-as Labdarúgó Európabajnokság FOOTIE korpuszán alapul, amely szinkrontolmácsolt sajtótájékoztatókon készült felvételeket tartalmaz. A kutató a riporter és az interjúalany kérdés–válasz stratégiáit, például a beszédstílust és a kihagyásokat, és azok tükröződését vizsgálja a tolmácsolt szövegekben.

Az utolsó cikk (Graham Neubig, Hiroaki Shimizu, Sakriani Sakti, Satoshi Nakamura, Tomoki Oda: *The NAIST Simultaneous Interpretation Corpus*) az angol–japán/japán–angol nyelvpárral foglalkozik. Az angol–japán/japán–angol szinkrontolmácsolási korpusz a NAIST gyűjteménye, amely két szempontból is különbözik a többi hasonló korpusztól: mivel pályakezdő és gyakorlott tolmácsok beszédprodukcióját is tartalmazza, a tolmács szakmai tapasztalata nem befolyásolja a vizsgálat eredményét. Másrészt pedig a korpusz egy része fordított szövegeket is tartalmaz, így segítségével ugyanazon forrásnyelvi szöveg tágabb időkorlát mellett született, azaz fordított, és szűkebb időkorláthoz kötött, azaz tolmácsolt változatát is vizsgálni lehet. A teljes korpusz mintegy 387 000 szövegszónyi adatot tartalmaz, megfelelő lehet különböző tolmácsolási stílusok vizsgálatára, ezen kívül pedig referenciaként használható új szinkrontolmácsolási korpuszok megalkotására is.

A *New Frontiers in Translation Studies* sorozat többi kötetéhez hasonlóan a *Making Way in Corpus-based Interpreting Studies* tanulmánykötet is a fordítás- és tolmácsolástudomány legújabb kutatási eredményeit mutatja be, a kitöltendő kutatási űrt ezúttal a korpuszalapú tolmácsolástudományhoz kapcsolódó kutatási kérdések és lehetséges válaszaik alkotják. A gépi eszközökkel vizsgálható korpuszalapú tolmácsoláskutatás további feltárást igénylő terület, ezért a szerzők a kötetben megjelent eredmények bővítésére és további részterületek feltérképezésére buzdítják az olvasót.

Irodalom

Dembry, C., Love, R. 2015. *Collecting the Spoken BNC2014 – overview of methodology.* Korpusznyelvészeti Konferencia, Egyesült Királyság, Lancaster University, 2015. július 21–24.

Halverson, S. 2003. The cognitive basis of translation universals. *Target*. Vol. 15. No. 2. 197–241.

Kajzer-Wietrzny, M. 2012. *Interpreting universals and interpreting style*. PhD értekezés, Adam Mickiewicz University.

Sandrelli, A., Bendazzoli, C. 2005. *Lexical patterns in simultaneous interpreting: A preliminary investigation of EPIC (European Parliament Interpreting Corpus).* Korpusznyelvészeti konferencia-sorozat 1(1).

http://www.birmingham.ac.uk/research/activity/corpus/publications/conference-archives/2005-conf-ejournal.aspx

Claudio Fantinuoli és Federico Zanettin (eds)

New directions in corpus-based translation studies

(Berlin: Language Science Press. 2015. 175 pp. ISBN: 978-3-944675-83-1)

Klenk Márk

klenky1990@gmail.com

Eötvös Loránd Tudományegyetem, Bölcsészettudományi Kar

Fordító- és Tolmácsképző Tanszék

A Claudio Fantinuoli és Federico Zanettin által szerkesztett *New directions in corpus-based translation studies* című tanulmánykötet a berlini Language Science Press *Translation and Multilingual Natural Language Processing* elnevezésű könyvsorozatában jelent meg. A sorozatban megjelent kötetek az emberi és a gépi fordítás témakörét járják körül, nagy hangsúlyt fektetve az empirikus kutatások bemutatására, valamint a kapcsolódó tudományterületek (számítógépes nyelvészet, korpusznyelvészet és kognitív nyelvészet) fordítással kapcsolatos szempontjainak ismertetésére. A sorozat keretében eddig hét tanulmánykötet jelent meg, és a közeljövőben két újabb megjelenése várható. A kiadványok bárki számára elérhetőek és ingyenesen letölthetőek a kiadó honlapjáról.[1]

A kötet egyik szerkesztője, Claudio Fantinuoli a Mainzi Egyetem adjunktusa, kutatási területe a korpuszalapú fordítástudomány és a tolmácsoláselmélet, továbbá a fordítók és a tolmácsok információkezelése. Szerkesztőtársa, Federico Zanettin a Perugiai

1 http://langsci-press.org/catalog/series/tmnlp

Egyetem Politikatudományi tanszékének angol nyelvi és fordítástudományi docense, kutatási területe a korpusznyelvészet és a fordítástudomány kapcsolata, mind elméleti, mind pedig a fordításoktatásban való alkalmazhatóság szempontjából.

A korpuszalapú megközelítés az elmúlt két évtizedben a fordítástudomány egyik fő kutatási módszertanává vált, tulajdonképpen egy paradigmaváltásnak lehettünk szemtanúi. A szerzők tanulmánykötetükben gyakorlati példákon keresztül mutatják be a különféle korpuszok fordítástudományban való alkalmazásának lehetőségeit. A kötetben hét tanulmány szerepel, így a szerkesztők ezek mentén hét fejezetre osztották könyvüket.

A kötet nem rendelkezik előszóval, ezért az első tanulmány, amelynek címe *Creating and using multilingual corpora in translation studies,* valójában a szerkesztőpáros előszavának is tekinthető. Röviden bemutatják a tanulmánykötet fő témáit, amelyek hét európai nyelvre koncentrálnak (baszk, holland, német, görög, olasz, spanyol és angol). A tanulmányokat a European Society of Translation Studies 2013 júliusában tartott germersheimi konferenciájára készítették. A kiválasztott tanulmányok szigorú, kettős vak lektoráláson estek át. A kötet legtöbb tanulmánya olyan korpuszok létrehozásának bemutatására koncentrál, amelyek korábban még nem álltak rendelkezésre, így most akár új kutatások alapjait is képezhetik. A szerkesztőpáros ezután röviden bemutatja a korpuszépítés, az annotáció és a párhuzamosítás (*alignment*), valamint a korpusz adatainak elemzésével kapcsolatos különböző megoldásokat, amelyek a kötet tanulmányaiban is felbukkannak. Kitérnek arra, hogy a párhuzamos korpusz és összehasonlítható korpusz elnevezés néha nem egészen helytálló, valamint a két korpuszfajta elkülönítése sem mindig világos. A párhuzamos korpusz ugyanis nem minden esetben tartalmaz fordításokat (például az Europarl és az Acquis Communautaire), hiszen ezek a szövegek az EU nyelvpolitikájának mentén mind eredeti szövegnek számítanak, és nem tekinthetők egymás fordításainak. Másodsorban az összehasonlítható korpuszok is tartalmazhatnak fordításokat, egyes hibrid szövegek pedig nehezen besorolhatók, mert egyaránt tekinthetők fordításnak és autentikus szövegnek (például a Wikipédia). Fontos tényező még a korpuszok annotálása is. Nagymértékben befolyásolja a korpusz későbbi felhasználását, hogy milyen nyelvi és nyelven kívüli információt tárolunk a korpuszunk elemeiről. Végezetül érdekes tényként

kitérnek arra, hogy a *Translation Studies Abstracts* adatbázisában megjelent absztraktok kulcsszavai alapján tízből egy publikáció kapcsolódik korpuszokhoz, vagyis a korpuszalapú fordítástudományhoz.

A kötet második tanulmánya a *Development of a keystroke logged translation corpus* címet viseli, írói Tatiana Serbina, Paula Niemietz és Stella Neumann. A szerzők ismertetik egy többnyelvű párhuzamos korpusz létrehozásának folyamatát. A forrásnyelvi és a célnyelvi alkorpusz mellett egy billentyűleütéses alkorpuszt is létrehoztak, amelyet az előző kettővel összehangolva vizsgálni tudják a fordítók munka közben végrehajtott javításait, törléseit, változtatásait, valamint egérmozgásukat is. Az adatokat a PROBRAL projekt keretében gyűjtötték tizenhat résztvevőtől a Saarlandi Egyetemen, valamint a Minas Gerais-i Szövetségi Egyetemen. A résztvevők feladata egy angol szöveg német nyelvre történő fordítása volt, időkeret nélkül. Fordítás közben csak egy online szótárat használhattak, billentyűleütéseiket, kurzormozgásukat és szüneteiket a Translog nevű programmal rögzítették a fordítás során. Ezen kívül szemmozgásukról és pupillaátmérőjükről is adatokat gyűjtöttek, azonban a korpusz jelenleg csak a billentyűleütéses adatokat tartalmazza. A tudománynépszerűsítő szövegekből álló forrásnyelvi és célnyelvi korpusz összesen körülbelül 3650 szót tartalmaz. A szövegek annotálása és párhuzamosítása automatikusan zajlott, azonban a billentyűleütésekből származó adatok hozzáadása feldolgozást igényelt. Mivel egy nem szokványos korpuszfajtáról van szó, a szerzők több lehetséges lekérdezési lehetőséget is felvázoltak. A billentyűleütések rögzítésének köszönhetően adatokat nyerhetünk a fordítás folyamatáról, alternatív fordítási változatokat, befejezetlen struktúrákat is kereshetünk, vagyis bepillanthatunk abba, hogy a fordító hogyan alkotja meg a szöveget. A változtatásokat vizsgálva a fordítók alternatív célnyelvi hipotéziseit is megfigyelhetjük (például a grammatikai nem változtatása, egyes szám és többes szám változtatás), ezen felül adatokat nyerhetünk a szófajváltással, valamint a lexikai behelyettesítésekkel kapcsolatban.

A harmadik tanulmány a *Racism goes to the movies: A corpus-driven study of cross-linguistic racist discourse annotation and translation analysis* címet viseli, Effie Mouka, Ioannis E. Saridakis és Angeliki Fotopoulou nevéhez fűződik. A szerzők egy há-

romnyelvű párhuzamos korpuszt hoztak létre angol, görög és spanyol nyelvű filmfeliratokból, kutatásuk a rasszista diskurzus vizsgálatára fókuszált fordítási perspektívából. A forrásnyelv minden esetben az angol volt. A kiválogatott öt film, amely a korpuszt képezi, műfajilag a dráma kategóriájába sorolandó, 1989 és 2006 között forgatták, történetük a rasszizmus és a faji kapcsolatok körül forog, valamint tartalmaz verbálisan kifejtett rasszizmust párbeszéd vagy monológ formájában. A korpusz 51000 angol, 29700 görög és 35900 spanyol szót tartalmaz. A szerzők azért választották éppen a görög és a spanyol nyelvet, mert a két kultúra közös tapasztalatokkal rendelkezik az irreguláris migrációval kapcsolatban. A kilencórányi filmanyag feliratait ELAN és GATE annotációs programmal dolgozták fel. Az angol forrásnyelvi felirat a hangzó szöveg átirata volt, a két célnyelvi felirat ennek a fordítása. A feliratkorpuszon kívül a szerzők referenciaként felhasználták az enTenTen12, GkWaC és az esTenTen11 korpuszt is. Korpuszuk segítségével azt vizsgálták, milyen módon kerül átültetésre a rasszista diskurzus, amikor egy filmet egy szociológiai és kognitív szempontból is eltérő és kissé távolabb álló társadalom nyelvére fordítanak le. A szerzők arra a megállapításra jutottak, hogy minden nyelvben/kultúrában létrejönnek és rögzülnek olyan (rasszista) jelentéstartalmak, amelyek más nyelvekben és kultúrákban ismeretlenek vagy legalábbis aszimmetrikusan vannak jelen.

A negyedik tanulmány a *Building a trilingual parallel corpus to analyse literary translations from German into Basque* címet viseli, szerzői Naroa Zubillaga, Zuriñe Sanz és Ibon Uribarri. A szerzők tanulmányukban egy többnyelvű, párhuzamos korpusz építésének lépéseit mutatják be, amelyet annak érdekében hoztak létre, hogy a németről baszk nyelvre történő közvetlen fordítást, valamint a spanyol nyelv beiktatásával a közvetett fordítást vizsgálják. Az Aleuska korpusz három alkorpuszból áll: egy gyermekirodalmi alkorpuszból, egy narratív szövegeket tartalmazó alkorpuszból, valamint egy filozófiai szövegeket tartalmazó alkorpuszból. A baszk nyelv kissebségi helyzete miatt nehéz olyan fordítást találni, amely a spanyol mint közvetítőnyelv beiktatása nélkül, közvetlenül készült. Továbbá Baszkföld lakosságának jelentős része sem beszéli a nyelvet, így a téma nyelvtervezési kérdéseket is érint. Egészen 2006-ig még német szótárral sem rendelkeztek. Korpuszuk tervezéséhez felhasználták az *Aleuska katalógust*, amely az összes németről baszk nyelvre fordított könyv bibliográfiai adatait tartalmazta. Az adatok begyűjtése

és az OCR -programmal beolvasott fájlok javítása után a korpusz mérete 5 551 204 szó lett (ebből forrásnyelvi német 2 722 000 szó, a fordított baszk 2 298 472 szó és a spanyol nyelv közvetítésével fordított baszk 490 732 szó). Az adatok címkézésére és párhuzamosítására a résztvevők által fejlesztett TRACE-Aligner számítógépes programot használták. A szerzők jövőbeni terve a három alkorpusz egyesítése egy nagy, német–baszk párhuzamos korpuszba, amelyet kutatási célokra szeretnének felhasználni. Hozzáteszik továbbá, hogy a korpusz részleteinek vagy egészének jelenlegi formában történő publikussá tétele a szerzői jogi vonatkozások miatt nehézségekbe ütközik.

Az ötödik tanulmány a *Variation in translation: Evidence from corpora* címet viseli, szerzője Ekaterina Lapshinova-Koltunski. A cikk egy angolról németre fordított szövegekből álló korpusz vizsgálatát írja le. A vizsgálat célja az emberi fordítás és a gépi fordítás eredményeképpen kapott szövegek összehasonlítása. A fordításnyelvet (*translationese*) három fordítási művelet megvalósulásának tükrében vizsgálja: az egyszerűsítés (kevesebb tartalmas szó a fordításban, sok ismétlődő szó), az explicitáció (a forrásnyelvi implicit információ kibontása), a normalizálás (célnyelvre jellemző minták használata), valamint ezen jegyek egymáshoz való viszonyának, vagyis konvergenciájának tekintetében. A vizsgálat alapjául szolgáló szövegek különböző fordítási módszerekkel keletkeztek, szerepelnek közöttük gyakorlott fordítók fordításai, kezdő fordítók CAT-eszközzel készített fordításai, egy szabályalapú gépi fordítórendszer és két statisztikai gépi fordítórendszer fordításai; mindegyik szöveg a VARTRA-SMALL korpusz részét képezi. A gyakorlott fordítók fordításai a CroCo korpuszból származnak, valamint az összehasonlítás alapjául szolgáló eredeti német és angol nyelvű szövegek is innen kerültek kiválasztásra. A VARTA korpuszból válogatott adatok mennyisége körülbelül hatszázezer szövegszó, az összehasonlítható alkorpusz mérete pedig kétszázötvenezer szövegszó. A korpusz elemeit tokenizálták, lemmatizálták, szófaji információkkal címkézték fel, szintaktikai egységekre és mondatokra bontották. A korpuszt vizsgálva a szerző egyedül a konvergencia jellemzőit tudta bizonyítani, arra a megállapításra jutott, hogy az elemzett fordítások konvergálnak a különböző fordítási módszerektől függetlenül. A zárszóban megemlíti, hogy a későbbiek során mondat és szószinten is párhuzamosítani tervezik a szövegeket annak érdekében,

hogy a kétértelmű részek fordítását, a közvetlen fordítói megoldásokat, valamint azok eltérő változatait is vizsgálni tudják.

A hatodik tanulmány a *Non-human agents in subject position: Translation from English into Dutch: A corpus-based translation study of "give" and "show"* címet viseli, szerzője Steven Doms. Az angol nyelvben a *give* és a *show* igéhez hasonlatos cselekvő igék alanyai lehetnek élettelen dolgok, amelyek így a mondatban ágensi funkciót is betölthetnek. A holland nyelv esetében élettelen dolgok ágens szerepben sokkal ritkábban fordulnak elő. A szerző hipotézise az volt, hogy a fordítások során a fordítók elvonatkoztatnak az angol nyelvi rendszertől, ilyen módon nem vagy csak ritkán fordítják az angol szintaktika szerint a mondatokat, vagyis az angol nyelvű élettelen ágensek mint a *give* és a *show* ige alanyai a holland változatokban más szerkezettel kerülnek megfeleltetésre. A vizsgálathoz felhasznált szövegek a Dutch Parallel Corpusból származnak. A szerző összesen 1986 találatot kapott a két kifejezésre keresve, majd ezeket kutatási kritériumainak megfelelően szelektálva és ellenőrizve szűkítette le a korpusza méretét 388 párhuzamos mondatra. Arra a megállapításra jutott, hogy azokban az esetekben, amikor a holland fordításból kikerültek az élettelen dolgok az ágensi szerepből, a mondatban inkább szemantikai változtatások történtek, mintsem explicitáció vagy implicitáció. A szemantikai változtatások a célnyelvi szövegben gyakran vezetnek eltérésekhez, amelyek eredményeképpen az ige valenciája megváltozik, más bővítményeket követel meg, mint a forrásnyelvi szövegben, azonban ezek a folyamatok lexikai információveszteséggel nem járnak. A kutatás eredményei azt mutatták, hogy a fordítók az esetek közel 60 százalékában meghagyták az élettelen ágenseket a hollandra fordított szövegekben, ez a megállapítás pedig az angol nyelv interferenciájára szolgáltat bizonyítékot.

Az utolsó tanulmány az *Investigating judicial phraseology with COSPE: A contrastive corpus-based study* címet viseli, szerzője Gianluca Pontrandolfo. Tanulmányában a büntetőjog tárgykörébe tartozó speciális frazeológiát vizsgálja meg, korpuszát a büntetőítéletek képezik. Kutatásának célja, hogy olyan multifunkcionális forrást biztosítson a jogi fordítók számára, amely segíti a fordítás folyamatát, valamint pozitív hatást gyakorol a fordítások minőségére is. Továbbá lehetőséget kíván biztosítani fordítóknak és jogi

szakértőknek egyaránt arra, hogy autentikus és élőnyelvi szövegek segítségével fejlesszék frazeológiai kompetenciájukat a frazeologizmusok szövegkörnyezetben való elhelyezésén keresztül. Jogi szakfordítás esetén nemcsak a lexikai és a terminológiai ekvivalencia megteremtése okoz nehézséget a fordítók számára, hanem a két eltérő jogi rendszer megfeleltetése is, valamint az iratok ebből következő eltérő felépítése és szövegezése. A szerző a háromnyelvű, hatmillió szövegszót tartalmazó (spanyol, olasz és angol) Corpus of Criminal Judgements (Corpus de Sentencias Penales, COSPE) szövegeit használta fel kutatásához, amely két alkorpuszt tartalmaz. Az egyik 2005 és 2012 között született legfelsőbb bírósági ítéletekből áll, a másik különböző ítélőtáblák ugyanebben az időszakban született ítéleteit tartalmazza. Arra a megállapításra jutott, hogy a büntetőítéletek nagyszámú frazeológiai egységet tartalmaznak mindhárom nyelv esetében. A vizsgált frazeologizmusok alacsony számú előfordulása a referenciakorpuszokban (CREA a spanyol esetében, CORIS/CODIS az olasz esetében és BNC az angol nyelv esetében) azt bizonyította, hogy ezek a vizsgált műfaj kulcsfontosságú lexiko-szintaktikai elemei, és a bírák jellegzetes fogalmazási gyakorlatát tükrözik.

A fordítások kutatásának manapság elengedhetetlen eszközei lettek a korpuszok. A tanulmányok ismertetéséből is kitűnik, hogy a korpuszok többféle megközelítési módot nyújthatnak a fordítások vizsgálatával kapcsolatban. Ha specifikus információkkal bővítjük, a fordítás folyamatával kapcsolatban is releváns megfigyeléseket tehetünk. Párhuzamos korpuszunk segítségével az átváltási műveletekre kaphatunk példákat, vizsgálhatjuk a fordításnyelvet, valamint a fordítási interferencia jelenségét. A kötet hasznosnak bizonyulhat a fordítástudomány korpusznyelvészeti megközelítése iránt érdeklődők számára, mivel rálátást nyújt a korpuszok sokrétű felhasználhatóságára. Ugyanakkor a terület ismerőinek is betekintést biztosíthat a jelenleg zajló kutatásokba.

A korpuszalapú fordításkutatás az ezredforduló talán legjelentősebb és legtermékenyebb kutatási paradigmájává vált. A technika fejlődése és a modern eszközök újabb lendületet adtak a korpuszalapú vizsgálatoknak mind a nemzetközi, mind a hazai fordításkutatók körében. A fordítási folyamat és a fordítási szöveg sajátosságainak leírását célzó kutatások ma már nem nélkülözhetik sem a korpusznyelvészeti eszközöket, sem a nyelvi közvetítés eredményeként létrejött célnyelvi szövegeket tartalmazó, nagyméretű digitális korpuszokat. A szövegállományok gépi elemzése statisztikai adatokkal támasztja alá a kutatási eredményeket.

A jelen tanulmánykötet a magyar fordítástudomány berkeiben folyó korpuszalapú kutatásokba kíván bepillantást nyújtani elméleti kérdések tárgyalásával és empirikus kutatások bemutatásával. A könyv tanulmányai áttekintik a korpusznyelvészeti vizsgálatok fő elméleti és módszertani dilemmáit, rávilágítanak a feltárt normák és univerzális fordítási sajátosságok különbségeire, valamint érdekes ismeretekkel és eredményekkel gazdagítják a korpuszalapú fordításkutatás iránt érdeklődők tudását a fordítás és tolmácsolás során keletkezett szövegek általános nyelvi jellemzőiről és a nyelvi közvetítők stratégiáiról.