

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 tolmácsolatát tartalmazó korpusssal dolgoztam (Szegh 2017). Második lépésként intermodális elemzést végeztem, megvizsgáltam ugyanazon forrásnyelvi szöveg tolmácsolatát és fordított változatát is. Mindkét esetben azonos módszertan szerint haladtam: először statisztikai vizsgálatnak vettem alá a szövegeket, majd a kvalitatív elemzés során igyekeztem a számadatok mögé nézni, és felfedni, mi áll a felszíni szövegjellemzők és a tolmácsok döntéseinek háttérében. Az eredmények az első és a második szakaszban is megerősítették előzetes hipotézisemet. A kutatás folytatásaként autentikus szövegekkel tervezem összehasonlítani a tolmácsolat 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-Pécs O. (szerk.) 2020. *Fókuszban a fordított és a tolmácsolat 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: Recent 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 egyrésről, hogy hatalmas adatbázist biztosítanak az empirikus alapokon nyugvó nyelvészeti elemzésekhez, másrésről 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, mintázatokot 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 mérföldkönek 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 egyértelműsége, a nagyobb explicitésre 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 explicitég 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 kiegyenlítődé (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 Español*) (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, kollokvialis 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 szövegprodukción. Ennek megfelelően a korpusz három al-korpuszból 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 szorítkoznánk 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 egyrészt a tolmácsok szeretik kihangsúlyozni az új információkat, amire leginkább a szenvedő szerkezet alkalmas, másrészt, hogy a tolmácsolásban a *that* kötelelemet 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 memóriagigény 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 szerkezetűvé 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 egyszerűsítanak 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, alapszókinsz 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, kollokvialis 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 szöveg-típus, 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 tolmácsolat verzióját vettem 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 tolmácsolat szövegek nemcsak a forrásnyelvi eredetihez, de a fordításokhoz képest is egyszerűbbek, illetve hogy természetük sokkal inkább beszéltnyelvi, mint fordítási sajátosságokkal rendelkeznek.

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 beszédprodukciónkat 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 szövegprodukción 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 mulékony, 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 szuprasegmentális 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 vettem 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 6.0² 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 szövegprodukciónban, 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 alapján 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 egyrészt, hogy a tolmács folyamatosan az idő nyomása alatt dolgozik, ügyelnie kell arra, nehogy túlságosan lemaradjon a beszélő mögött, másrésztől 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 szövegprodukciónak 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 szöveggörnyezetből 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 megtiltana 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, elnémitó, 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 szöveké. 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 trilógusra, 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 Committee* 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 trilogue 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 tolmácsolásakor 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 mellett 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 céloom kettős volt. Elsősorban arra voltam kíváncsi, vannak-e a tolmácsolt szövegeknek általános, univerzális tulajdonságai, felfedezhetők-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

Statistikai 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 találhatók. 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 rendelkezésre 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 implicitabbak, 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 tény: 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 mondathosszság, 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 átíratához 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 Cognition in 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 trascrizioni di eventi di interpretazione in televisione*. Unpublished dissertation. Advanced School for Translators and Interpreters (SSL-MIT). University of Bologna at Forlì.

- Cencini, M. 2002. On the importance of an encoding standard for corpus-based interpreting studies. Extending the TEI scheme. in *TRAlinea Special Issue* <http://www.intralinea.org/specials/article/1678>
- Chernov, G.V. 1978. *Teoriya i praktika 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?* Unpublished 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'interprete. 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 Translation 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 Università 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*

Interpreting 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/d5f6/9e1db393df443b6cc-2f754a4488526378cde.pdf?_ga=2.15170165.233430643.1553866200-1540436159.1553613303&_gac=1.226022888.1553866200.Cj0KCQjwhPfkBRD0ARIsAAcYycEuue-xdO8bLup8J6ApW-IObc00CrfAM-mo9YZnsTuKadXO-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/>