Modern Nyelvoktatás 2021/3–4.
DOI: 10.51139/monye.2021.3-4.58.71

KOVÁCS TÍMEA

Névmási szerkezetek kvantitatív és kvalitatív elemzése forrásnyelvi és fordított büntetőjogi szövegekben

A jogi szaknyelv és a szaknyelvi fordítás – sajátos jellegénél fogva – a fordítástudományon és szaknyelveken belül is külön ágat képvisel. Több szerző foglalkozik a jogi szaknyelv jellemzőivel, valamint a nyelveken, kultúrákon és jogi intézményi rendszereken átívelő interlingvális jogi fordítás legfontosabb problémáival. Ugyanakkor a gépi fordítás az elmúlt pár évben megfigyelhető robbanásszerű fejlődése és az annak nyomán kialakult új és egyre nagyobb jelentőséggel bíró utószerkesztés (poszteditálás) újabb kihívások elé állítja a fordítástudományt és a gyakorló fordítókat. Ha egyre több szöveg és azok fordítása válik elektronikus formában szinte mindenki számára könnyen elérhetővé, az milyen új feladatok és elvárások elé állítja a fordítókat és utószerkesztőket (poszteditorokat)? Hogyan alakul át a fordítói minőségi sztenderd fogalma? Hogyan definiálhatjuk az ideális (jogi) szaknyelvi fordítást, mely mind a gépi, mind a humán fordítás minőségi szintje lehet?

Kulcsszavak: jogi szaknyelv, humán fordítás, gépi fordítás, utószerkesztés, minőségi sztenderd

Bevezetés

A jogi fordítás egyik legfontosabb jellemzője a forrásnyelvi tartalom lehető legpontosabb megőrzésére és a célnyelvi értelmezhetőségre való törekvés. A normatív vagy preskriptív jogi szövegek esetében a forrásnyelvi tartalom vagy a célnyelvi jól-megformáltság és kulturális beágyazottság kontinuuma mentén a forrásnyelvi tartalom megőrzése ugyanolyan fontos, ha nem fontosabb, mint a célnyelvi jól-megformáltság. Ugyanakkor ezt a fordítónak olyan célnyelvi eszközökkel kell elérnie, hogy a jogi szöveg legfontosabb jellemzője – a szöveg által érintett felek számára a lehető legpontosabb értelmezés lehetősége – ne sérüljön. Ebben a folyamatban a fordítónak figyelembe kell vennie számos olyan nyelvi, kulturális, pragmatikai és jogrendszerből adódó sajátosságot, amelyek a cél- és forrásnyelvi szövegek átjárhatóságát, fordíthatóságát és értelmezhetőségét befolyásolják. Az angol–magyar és magyar–angol jogi szaknyelvi fordítás műfaji sajátosságaival (Balogh, 2020) és a releváns stratégiákkal (Kovács, 2018) több szerző is foglalkozik. Mivel a jelen tanulmány szerzője nyelvész, a jogi szaknyelvi fordítás szempontjából jelentős forrás- és célnyelvi jellemzőket, azon belül is az összetett névmások angol forrásnyelvben és a gépi és humán magyar fordításban megjelenő használatát vizsgálja. Mivel a jogi szövegek fordításában, ahogy alapvetően az emberi kommunikációban, a referenciális funkció kulcsfontosságú, érdemes a jogi szövegekben megvizsgálni a névmási szerkezetek használatát és fordítá-

sát, hiszen azok jelentősen befolyásolják az adott elemek értelmezhetőségét és egyéb vonatkozásait.

Jelen tanulmány arra tesz kísérletet, hogy összehasonlítsa és elemezze az angol jogi szaknyelv egyik archaikus jellemzőjének – az összetett névmási szerkezetek – az angol forrásnyelvi normatív (preskriptív) szövegben és annak (nyers) gépi és humán fordításaiban való előfordulási gyakoriságát és használatát. A tanulmány arra keresi a választ, hogy a korpusznyelvészeti eszközökkel és az azokat kiegészítő kvalitatív módszerekkel vizsgált szövegek közül melyik felel meg jobban a forrásnyelvi angol szövegben megfigyelhető nyelvhasználati tendenciáknak. Melyik célnyelvi (a humán vagy a gépi fordított) szöveg „gravitál" (Halverson 2003, 2010a, 2017) jobban a forrásnyelvi szöveg felé.

Gépi fordítás, poszteditálás (utószerkesztés), korpusznyelvészet

A gépi fordítás a 2016 óta elterjedt neurális háló alapú alkalmazásoknak köszönhetően dinamikusan fejlődik. Ezzel párhuzamosan és részben ennek köszönhetően egyre több eredeti forrásnyelven megírt szöveg érhető el elektronikus formában. Ily módon az internet egy forrásnyelvi (és adott célnyelvre fordított) korpusznak is tekinthető (Seidl-Péch, 2020), amely szinte bárki számára elérhető. Mivel a gépi fordítás összehasonlíthatatlanul rövidebb idő alatt fordítja ugyanazt a forrásnyelvi szöveget a célnyelvre, a fordításszolgáltatók körében egyre inkább elterjed a használata. Olyannyira, hogy akár az ügyfél választhatja ki, hogy egy adott forrásnyelvi szöveg gépi (emberi közbelépés nélküli nyers) fordítását kéri-e, vagy annak (emberi közbelépéssel) utószerkesztett változatát. Az utószerkesztés (poszteditálás) definíciója az ISO 17100:2015 sztenderd szerint: „a gépi fordítás szerkesztése és javítása". Ugyanakkor az utószerkesztés során szem előtt tartott minőségi sztenderdek és referenciának tekintett szövegek megfogalmazása még meglehetősen pontatlan. Noha vannak erre vonatkozó számszerűsíthető mutatók (metrikák), valójában puszta alkalmazásuk még nem elegendő az adott célnyelvi szöveg minőségének pontos megítélésére. Általánosan elterjedt megközelítésmódok a BLEU (Papineni et al., 2002) és a (H)TER (Snover et al., 2006). Ezek a gépi fordítás és a referenciafordítás közötti átfedések és szerkesztés alapján (szókihagyás, csere stb.) számítanak ki egy értéket. Valójában a poszteditálásban nincsenek általánosan elfogadott vagy sztenderdnek tekinthető alapelvek (DePalma, 2013; TAUS, 2016). A fordításban résztvevő gyakorlati és elméleti szakamberek kidolgoztak több lehetséges minőségi szintet az utószerkesztés meghatározására. Ezek alapján megkülönböztetünk részleges (*partial*) és teljes (*full*) poszteditálást (Hu & Cadwell, 2016 nyomán). A jelenlegi tudásunk szerint azonban ez a felosztás nem tűnik elégségesnek a poszteditálás pontos minőségi meghatározásához.

A jelenleg a szemünk előtt zajló rohamos fejlődésnek köszönhetően a gépi fordítás eredményeképpen létrejött végtermék, a szöveg – természetesen a szöveg jellegétől, az érintett cél- és forrásnyelvektől, az alkalmazott gépi fordítótó programtól (Google, DeepL stb.) függően – egyre jobb minőségű. Egyes teoretikusok szerint akár el is érheti az emberi fordítási minőséghez (*near human*) közeli szintet (Lample et al., 2018). Kérdés, vajon – a folyamatos fejlődés eredményképpen – meg is haladhatja-e azt.

A fordítás minőségének megítélése és szükség szerinti javítása a lektori kompetenciába tartozik. A jogi szakszövegek esetében, mint láttuk, ez különösen összetett feladat, hiszen a célnyelvi szövegnek egyaránt (ha nem jobban) kell eleget tennie a forrásnyelvi pontos tartalom és a célnyelvi jól megformáltság időnként egymásnak feszülő elvárásainak. A gépi fordítás rohamos terjedése és fejlődése eredményeképpen a fordítói kompetencia egyre inkább átalakul poszteditálási kompetenciává (Pym, 2013). Ugyanakkor várat még magára a poszteditálással szemben támasztott elvárások meghatározása. A jelenlegi gyakorlat szerint a poszteditor az adott minőségi szinten (amely lehet „részleges" vagy „teljes") elvégzi az utószerkesztést egy referenciának tekinthető szöveg alapján. A referenciaszöveg definíciója sem egyértelmű, leginkább az adott műfajhoz tartozó regiszterben írt eredeti forrásnyelvi szövegek és azok elfogadottnak tekintett célnyelvi fordításai minősülnek annak. Éppen ezért érdemes megvizsgálni egyes nyelvi elemek használatát – referenciaszövegnek – tekinthető eredeti forrásnyelvi és célnyelvre fordított szövegekben.

A fordítástudomány módszertanában is használt számítógépes alkalmazások – az 1990-es évektől megfigyelhető – elterjedésének köszönhetően a korpusznyelvészet egyre nagyobb szerepet játszik a kutatásokban. Mivel a korpusznyelvészeti eszközök hatalmas terjedelmű elektronikusan elérhető szövegekben teszik lehetővé a kvantitatív kutatásokat, joggal népszerűek a kutatók körében. Ugyanakkor több elméleti szakember hívja fel a figyelmet a korpusznyelvészeti eszközök alkalmazásában rejlő veszélyekre. A kizárólag kvantitatív alapú és az egyes elemekre irányuló (*item-based*) kutatás egyoldalúvá teheti az eredmények értelmezését (Heltai, 2014). Éppen ezért a jelen tanulmányban megjelenő adatok gyűjtéséhez és elemzéséhez használt korpusznyelvészeti (*Sketch Engine*) kutatást kiegészíti az eredmények kvalitatív elemzése és értelmezése.

A kutatás célja és módszertana

Jelen tanulmány célja tehát referenciának tekinthető forrásnyelvi szövegekből és azok gépi és humán fordításából összegyűjtött nyelvi elemek korpusznyelvészeti eszközök segítségével történő kvantitatív vizsgálata és kvalitatív elemzése. Az elemzés elsősorban arra keresi a lehetséges választ, milyen tendenciák figyelhetők meg egy – a jogi szaknyelvi regiszterre jellemző – nyelvi elem (*thereof*) forrás- és célnyelvi (gépi és humán fordított) használatában, és ezek a tendenciák hogyan befolyásolhatják a fordítástól és a poszteditálás során elvárt minőséget és az arra ható tényezőket. Jelen tanulmány szerzője több korábbi tanulmányában vizsgálta angol–magyar és magyar–angol jogi szaknyelvi fordításban a forrás- és célnyelvi szövegek egymásra hatását humán (Kovács, 2020b) és gépi fordításokban (Kovács, 2020a, 2020c) is. Mivel fordítói és fordításoktatói gyakorlatom során sokszor szembesültem angol–magyar jogi fordítás során a *thereof* és egyéb összetett névmási szerkezetekkel, jelen kutatásom tárgyául ezek használatát választottam. A korpusznyelvészeti eszközök alkalmazását és a korpusz kiválasztását megelőzte az a feltételezés, miszerint érdemes a *thereof* és egyéb összetett névmások használatát összehasonlító elemzés alá vetni. Jelen kutatás tehát irányított korpusznyelvészeti (*corpus-based*), kutatás, melyet kiegészít a vizsgált nyelvi elemek forrás- és célnyelvi használatának kvalitatív és komparatív elemzése.

Az elemzéshez összesen négy egynyelvi korpuszt használtam. Angol forrásnyelvi – referencia- – korpusznak Kalifornia állam Büntetőjogi törvénykönyvét (*Penal Code of California*) választottam (a továbbiakban: PCC_EN_REF). A magyar forrásnyelvi referenciakorpusz a 2012. évi C. Törvény a Magyar Büntetőjogról (a továbbiakban: CC_HU_REF). A vizsgált célnyelvi angol fordítások pedig az interneten az Nemzeti Jogszabálytáron elérhető (humán fordítás), a továbbiakban: CC_HU_EN_HT és a Deep L Pro gépi (neurális háló alapon működő) alkalmazás fordítása (a továbbiakban: CC_HU_EN_MT). A kutatáshoz azért választottam a Büntetőjogi törvénykönyveket, melyek Šarčevič (1997) és más szakemberek (Kjaer, 2007; Ződi, 2017) felosztása alapján is egyértelműen normatív, preskriptív jellegű szövegek, mert különösen fontos mindenkire vonatkozó előíró jellegük. Természetesen az interlingvális jogi fordítás során előkerülnek olyan fontos szempontok, mint a két ország, kultúra, jogrendszer közötti nyelvi és nem-nyelvi tényezők. Tekintettel arra, hogy ez egy nyelvhasználatra irányuló elemzés, a vizsgálat kizárólag a konkrét nyelvi elem használatára terjed ki.

A fent említett négy szövegből korpuszt építettem. A négy korpuszra vonatkozó általános adatok az 1. táblázatban találhatók.

1. táblázat

A vizsgált korpuszok általános jellemzői

	PCC_EN_REF	CC_HU_EN_HT	CC_HU_EN_MT	CC_HU_REF
szavak száma	391257	66326	73872	53130
mondatok száma	9336	1430	1481	1531
lexikai sűrűség	41,9	46,4	49,9	34,7

Az 1. táblázat adatai alapján érdemes megjegyezni, hogy a kaliforniai büntető törvénykönyv (391257 szó) jóval terjedelmesebb, mint a magyarországi (53130 szó). Mivel a jelen elemzés nem terjed ki a két törvénykönyv felépítésében és tartalmában megfigyelhető hasonlóságok és eltérések vizsgálatára, ezzel ebben a tanulmányban nem foglalkozunk. Ugyanakkor ennek a vizsgálata egy jövőbeni kutatás célja lehet. Jelen tanulmány kizárólag az összetett névmások eredeti angol nyelvű referenciaszövegben megfigyelt használatára és azok célnyelvi fordításaira összpontosít.

A magyar forrásnyelvi referenciaszöveg és annak két fordítása kapcsán megfigyelhető, hogy mindkét fordított szöveg (humán fordítás: 66326 szó, gépi fordítás: 73872 szó) terjedelmesebb a magyar referenciaszövegnél (53130 szó). A három szöveg közül a legterjedelmesebb a HU–EN, tehát a gépi angol (nyers) fordítás. Ha megvizsgáljuk az adott szövegek lexikai sűrűségét, vagyis az egy mondatban felhasznált szavak átlagát, akkor azt látjuk, hogy a magyar törvénykönyv a legkevésbé sűrű (34,7), és a két fordítás lexikai sűrűségben meghaladja a forrásnyelvi szövegeket. Ugyanakkor a két angol célnyelvi (humán és gépi fordított) szöveg a lexikai sűrűség tekintetében közelebb áll a forrásnyelvi referenciaszövegnek tekinthető kaliforniai törvénykönyvben megfigyelt sűrűséghez. Ennek feltehetően az az oka, hogy a magyar szintetikus, agglutináló nyelv, az angol pedig analitikus.

Az összetett névmási szerkezetek használata a vizsgált korpuszokban

Az általános korpuszjellemzők után nézzük, hogy az összetett névmások használatában milyen eredményeket kapunk. Sok kritika éri az összetett névmások alkalmazását. Az angol jogi fogalmazást oktató tananyagok többsége ellenzi ezek használatát (például British Legal Centre), mivel szerintük nem elég egyértelmű, mire utalnak. Abban ugyanakkor egyetértés van, hogy fontos megérteni ezt az archaikusnak vélt nyelvhasználati formát, hiszen a jogi nyelv egyszerűsítésére irányuló törekvések ellenére a mai napig használatban vannak.

Pavlíčková (2008) szerint a *there*- és *here*-rel kezdődő elemek jogi deiktikus nyelvi eszközök, amelyek az adott beszédhelyzetre vagy azon kívülálló helyzetre, elemre, viszonyra utalnak. Valójában a *this* és *that* mutatónévmások helyett használatosak, csak a jogi szaknyelvi szövegben szükséges prepozíciókkal társítva. A leggyakrabban használt prepozíciók az -*after*, -*by*, -*upon*, -*in*, -*on*, -*to* stb. Ezekben a leginkább jogi szaknyelvben használt határozói szerkezetekben az első elem egy helyhatározó – a *there* vagy a *here* – amelyekhez egy prepozíció tapad – jelen esetben poszt-pozícióként funkcionálva. A fő különbség a *there*-rel és *here*-rel kezdődő szerkezetekben, hogy a *that*-et helyettesítő elem az adott szövegtől távolabbi dologra utal, míg a *this* közelebbi elemre. Egy másik értelmezés szerint a *there* előtag az adott szövegen kívüli másik dokumentumra vagy elemre utal, míg a *here*-rel kezdődő elemek az adott dokumentumra (Bázlik et al., 2010). Pavlíčková (2008) szerint ezek a nyelvi eszközök határozók, melyeket kötőszóként és melléknévként is lehet használni. A jogi kommunikációs helyzetben résztvevők közötti, valamint az adott szöveghez és a szövegben említett szövegekhez való viszonyt jelölik. Utalhatnak egy vagy több félre, időre, helyre, sőt ezek egyes részeire is. Ezen kívül bizonyos dolgokra is, például kijelentésekre, mondatokra vagy gondolatokra. Mivel a funkciójukat tekintve adott nyelvi elemek helyettesítésére és azokra való utalásra használjuk ezeket a kifejezéseket, jelen tanulmány szerzője is inkább összetett névmási szerkezetként (Balogh, 2020) utal rájuk.

A *here* és *there* szavakkal kezdődő és az ahhoz poszt-pozícióban csatolt prepozíciók által alkotott szerkezetek a jogi szaknyelvi használat sajátos példái. Éppen ezért érdemes megvizsgálni ezek használatát az angol forrásnyelvi referenciaszövegben és a magyar forrásnyelvi szöveg angol (humán és gépi) fordításaiban. A 2. táblázat a három korpuszban – PCC_EN_REF és annak emberi CC_HU_EN_HT és gépi CC_HU_EN_MT fordításában – leggyakrabban előforduló összetett névmási szerkezeteket és azok gyakoriságát szemlélteti. Értelemszerűen az összehasonlításból kimarad a magyar forrásnyelvi szöveg.

2. táblázat

Az összetett névmási szerkezetek használata a vizsgált korpuszokban

	PCC_EN_REF	Gyak.	CC_HU_EN_HT	Gyak.	CC_HU_EN_MT	Gyak.
thereof	296	0,0008	37	0,0006	20	0,0002
therein	66		29		23	
thereafter	39		0		0	

thereby	39		19		24	
thereto	34		0		6	
hereby	27		0		0	
herein	22		0		0	
therefrom	20		6		1	
therewith	13		3		2	
thereupon	10		0		0	
hereafter	8		0		0	
hereinafter	6		12		17	
thereunto	2		0		0	
hereinabove	1		0		0	
hereof	1		7		0	
hereto	1		0		0	
heretofore	1		0		0	
therethrough	1		0		0	
theretofore	1		0		0	

A 2. táblázat alapján jól látszik, hogy két szövegben a *thereof* a leggyakrabban használt szerkezet. A teljes szöveghez képest a leggyakrabban az angol forrásnyelvi referenciaszövegben fordul elő (0,0008%), ezt követi a humán fordított szöveg (0,0006), majd a gépi fordítás (0,0002%). Érdekes, hogy a két másik angol nyelvű szövegtől eltérően a gépi fordított szövegben a *thereby* a leggyakoribb, amit nem sokkal lemaradva a *therein* követ. Fontos megjegyezni, hogy a *hereto* az angol forrásnyelvi referenciaszövegben mindössze egy alkalommal szerepel, míg a humán fordításban hét alkalommal, a gépi fordított szövegben pedig egyszer sem. Ha megvizsgáljuk, hányféle összetett névmási szerkezet szerepel az egyes szövegekben, akkor azt láthatjuk, hogy a legnagyobb változatosság az angol forrásnyelvi referenciaszöveget jellemzi – összesen 19-féle elemet tartalmaz. A változatosság szempontjából azonban a két fordított szöveg azonos, igaz, az egyes elemek használata eltér. Mindkét szövegben előfordul a *thereof*, *therein*, *thereby*, *therefrom*, *therewith*, *hereinafter*, de a *hereof* csak a humán, a *thereto* pedig csak a gépi fordításban.

Ezt követően megvizsgáltam a *thereof* kifejezés mondatbeli elhelyezkedését a *Sketch Engine*-ben elérhető *N-grams* alkalmazás segítségével. Ez az alkalmazás összegyűjti a *thereof* használatát kettőtől hat elemig terjedő szókapcsolatokban. Ezt az alábbi három táblázatban foglaltam össze. A 3. számú táblázat tartalmazza azt a tíz leggyakoribb kettőtől hat elemig terjedő szókapcsolatot, melyben a *thereof* is előfordul.

3. táblázat

Az angol büntetőjogi szövegben (PCC_EN_REF) megjelenő N-grammok és azok gyakorisága

PCC_EN_REF	Gyakoriság
part thereof	36
or any part thereof	27
any part thereof	27
conviction thereof	26
upon conviction thereof	25
and upon conviction thereof	22
thereof shall	22
thereof to	19
upon conviction thereof shall	17
conviction thereof shall	17

A táblázatból egyértelműen látszik, hogy a *thereof* elem a leggyakrabban a *part* és *conviction* főnevekkel, illetve az azokra vonatkozó bővítményekkel, valamint a *shall* segédigével együtt jelenik meg. Érdekességképpen kigyűjtöttem a legtöbb, vagyis hat elemből álló szókapcsolatokat. Ezeket a szövegben való megjelenésük számával együtt a 4. táblázat foglalja össze:

4. táblázat

Hat elemből álló N-grammok a forrásnyelvi angol referenciaszövegben (PCC_EN_REF)

PCC_EN_REF	Gyakoriság
and upon conviction thereof shall be	13
conviction thereof shall be punished by	11
upon conviction thereof shall be punished	11
thereof shall be punished by a	8

Látható, hogy a *thereof* szókapcsolat központi elemét továbbra is a *conviction* és a *shall* alkotja, ezek egészülnek ki további bővítményekkel.

A következő, 5. táblázatban a magyar büntető törvénykönyv humán fordításában (CC_HU_EN_HT) előforduló N-grammokat gyűjtöttem össze.

5. táblázat

A magyar büntetőjogi szöveg humán angol fordításában (CC_HU_EN_HT) megjelenő N-grammok és azok gyakorisága

CC_HU_EN_HT	Gyakoriság
facilitating thereof	6
or facilitating thereof	6
part thereof	5

gaining knowledge thereof	4
after gaining knowledge thereof	4
promptly after gaining knowledge thereof	4
required therefor or facilitating thereof	4
offense promptly after gaining knowledge thereof	4
therefor or facilitating thereof	4
knowledge thereof	4

Az 5. táblázat alapján a vizsgált korpuszban kisebb a szórás a szókapcsolatok gyakoriságában. A legtöbb elem hat alkalommal fordul elő, de a tizedik leggyakoribb is négy alkalommal. A *thereof* érdekes módon a *facilitating* és *gaining knowledge* igeneves szerkezetekkel és azok bővítményeivel, vagy azok nélkül jelenik meg. A 6. táblázat a hat elemből álló szókapcsolatokat tartalmazza.

6. táblázat

A hat elemből álló N-grammok a humán fordított angol szövegben (CC_HU_EN_HT)

CC_HU_EN_HT	Gyakoriság
offense promptly after gaining knowledge thereof	4
impacts thereof increasing or decreasing the	2
and the impacts thereof increasing or	2
the impacts thereof increasing or decreasing	2
financial year and the impacts thereof	2
authorities before they become aware thereof	2
thereof increasing or decreasing the profit	2
programs required therefor or facilitating thereof	2
year and the impacts thereof increasing	2

A 6. táblázatból kitűnik, hogy a leggyakrabban használt hat elemből álló szókapcsolatokban már megjelenik az *impacts*, a *become aware* és az *increasing* elem is.

Ezt követte a magyar forrásnyelvi szöveg gépi fordításában előforduló N-grammok vizsgálata. A vizsgálat eredményeit a 7. táblázat foglalja össze.

7. táblázat

A magyar büntetőjogi szöveg gépi angol fordításában (CC_HU_EN_MT) megjelenő N-grammok és azok gyakorisága

CC_HU_EN_MT	Gyakoriság
part thereof	7
thereof shall	5
part thereof shall be	5
part thereof shall	5

thereof shall be	5
remaining part thereof shall	4
respect thereof	4
in respect thereof	4
the remaining part thereof	4
the remaining part thereof shall	4

A 7. táblázat adatai alapján megállapítható, hogy a gépi fordított szövegben található szókapcsolati elemek szintaktikai beágyazottsága jobban hasonlít az eredeti angol nyelvű referenciaszöveghez. Ebben a szövegben is a *part* és a *shall* a leggyakoribb központi elem a *thereof*-val közösen alkotott szókapcsolatokban. Ezeken kívül megjelenik még a *respect* elem. A hat elemből álló szerkezeteket a 8. táblázat foglalja össze.

8. táblázat

A hat elemből álló N-grammok a gépi fordított angol szövegben (CC_HU_EN_MT)

CC_HU_EN_MT	Gyakoriság
the remaining part thereof shall be	4
part thereof shall be converted into	4
or the remaining part thereof shall	4
thereof shall be converted into imprisonment	4
remaining part thereof shall be converted	3
service or the remaining part thereof	3
a third person in respect thereof	3
unlawful advantage or a promise thereof	3

A 8. táblázat alapján megállapítható, hogy továbbra is a *part* és a *shall* elemek a leggyakoribbak a *thereof*-os szókapcsolatokban, kiegészülve a *respect* és *promise* elemekkel.

A 8. táblázat adatai ismeretében összegzésképpen elmondható, hogy a *thereof* elem szövegekben előforduló gyakorisága alapján a két (humán és gépi) angolra fordított szöveg közül a humán fordítás közelít jobban az eredeti angol – referenciának tekintett – szöveghez. Ha azonban a szövegekben betöltött szerepét és a leggyakoribb szókapcsolatokban való előfordulását tekintjük, akkor egyértelműen a gépi fordítás által létrejött szöveg hasonlít jobban az angol nyelvű referenciaszöveghez.

A fordításokban használt összetett névmási szerkezetek kvalitatív elemzése

A kvantitatív elemzést követte a fordítások kvalitatív elemzése. Az elemzés fókuszában a két fordításban kigyűjtött *thereof* elemek összehasonlítása állt. A fenti elemzés adatai alapján a *thereof* a humán fordításban 37, míg a gépi fordításban 20 alkalommal jelent meg. A humán fordításban megjelent 37 *thereof* szóelem más elemekkel való használata a gépi fordításban mindössze négy alkalommal egyezett. A közös ezekben az elemekben, hogy minden estben a *part* elemmel együttes szókapcsolatban jelent meg a *thereof*.

1. példa:

CC_HU_REF	CC_HU_EN_HT	CC_HU_EN_MT
Ha a fiatalkorú a munkakötelezettségének önként nem tesz eleget, a közérdekű munkát vagy ***annak hátralévő részét*** szabadságvesztésre kell átváltoztatni.	If a juvenile offender fails to voluntarily comply with his work obligation, the community service work, or the ***remaining part thereof***, ***shall*** be substituted to imprisonment.	If the juvenile does not voluntarily fulfil his/her work obligation, the community service or the ***remaining part thereof shall*** be converted into imprisonment.

Jól látszik, hogy a forrásnyelvi magyar szövegben *az annak hátralevő részét* fordításaként jelenik meg a *the remaining part thereof* mindkét fordításban. Ugyanakkor mindössze négy esetben volt átfedés a humán és gépi fordított szövegben megjelenő *thereof* használatában: a humán fordításban használt 33 esetben a gépi fordításban a *thereof* használatától eltérő megoldás található. Ha megnézzük, hogy a magyar forrásnyelvi szöveg mely elemei fordításában jelent meg a *thereof* kifejezés, akkor azt látjuk, hogy a legtöbb esetben az *ez* vagy *az* mutatónévmás ragozott alakjai (*ezek*, *azok*, *ennek*, *annak*, *ezt*, *azt* stb.) fordításaként (összesen 20 elem). Ezt a következő példában is láthatjuk:

2. példa:

CC_HU_REF	CC_HU_EN_HT	CC_HU_EN_MT
[...] a csoportot olyan életfeltételek közé kényszeríti, amelyek azt vagy ***annak egyes tagjait*** pusztulással fenyegetik,	[...] constrains the group into living conditions threatening the demise of the group on the whole or ***certain members thereof,***	[...] forces the group into living conditions that threaten to destroy it or ***some of its members***,

A *thereof* elem az angol célnyelvi humán fordításban a magyar forrásnyelvi szövegben szereplő főnév birtokos ragozott alakjaként (összesen hat esetben) is megjelent. Ezt a következő – 3. példa – szemlélteti:

3. példa:

CC_HU_REF	CC_HU_EN_HT	CC_HU_EN_MT
egyes, a halálbüntetés, a kínzás vagy más kegyetlen, embertelen vagy megalázó bánásmód vagy büntetés során alkalmazható áruk kereskedelméről szóló, 2019. január 16-i (EU) 2019/125 európai parlamenti és tanácsi rendelet – ***II. mellékletében*** meghatározott áruk tekintetében – 33. cikkének,	Article 17 of Council Regulation (EC) No. 1236/2005 of 27 June 2005 concerning trade in certain goods which could be used for capital punishment, torture or other cruel, inhuman or degrading treatment or punishment, as regards the goods specified in ***Annex II thereof***;	17. Article 33 of Regulation (EU) 2019/125 of the European Parliament and of the Council of 16 January 2019 concerning trade in certain goods which could be used for capital punishment, torture or other cruel, inhuman or degrading treatment or punishment, as regards the goods defined ***in Annex II,***

Érdekes módon akkor is a thereof szerkezet jelent meg a fordításban, amikor a magyar forrásnyelvi szövegben egy egész tagmondat szerepel (összesen négy elem), lásd 4. példa:

4. példa:

CC_HU_REF	CC_HU_EN_HT	CC_HU_EN_MT
[...] nem tett meg a hatáskörében álló minden szükséges és indokolt intézkedést annak érdekében, hogy megakadályozza elkövetését, vagy haladéktalanul nem tett feljelentést azt követően, ***hogy a bűncselekmény elkövetéséről tudomást szerzett.***	[...] the public executive did not take the measures deemed necessary and justified within his power to prevent the criminal offense, or did not report the criminal offense promptly ***after gaining knowledge thereof***.	[...] failed to take all necessary and reasonable measures within his or her power to prevent the commission of the offence or to report the offence promptly ***after becoming aware that it had been committed.***

Három esetben a *thereof* nem feleltethető meg a magyar forrásnyelvi szöveg egyetlen elemének sem, lásd 5. példa:

5. példa:

CC_HU_REF	CC_HU_EN_HT	CC_HU_EN_MT
[...] az ehhez szükséges vagy ***ezt könnyítő*** gazdasági, műszaki vagy szervezési ismeretet másnak a rendelkezésére bocsátja.	[...] conveys economic, technical and/or organizational expertise required therefor or ***facilitating thereof***.	[...]makes available to another person the economic, technical or organisational knowledge necessary or ***useful for that purpose***.

Látható, hogy többségében a mutatónévmások, például *ezek*, *annak*, de akár a tagmondatok rövidítéseként is megjelenik a *thereof*, főleg, ha a fordító a nominális szerkezeteket akarja megőrizni. Ugyanakkor azt látjuk, hogy a gépi fordított szövegben tendenciózus a vonatkoztatott elem megismétlése, a birtokviszony birtokos melléknévvel (*its*, *their*) jelölése vagy a teljes tagmondat használata.

Érdemes megvizsgálni azokat a szövegrészleteket, ahol a gépi fordításban megjelenik a *thereof* kifejezés (13 alkalommal), de a humán fordításban nem. A *thereof* a leggyakrabban a *part* elemmel alkotott szókapcsolatban (összesen három alkalommal) jelenik meg. Erre láthatunk példát az alábbiakban:

6. példa:

CC_HU_REF	CC_HU_EN_HT	CC_HU_EN_MT
Ha az elítélt a pénzbüntetést nem fizeti meg, illetve részletfizetés engedélyezése esetén egyhavi részlet megfizetését elmulasztja, a pénzbüntetést vagy ***annak meg nem fizetett részét*** szabadságvesztésre kell átváltoztatni.	If the defendant did not pay the fine, or if failed to pay a monthly installment where payment by installment had been authorized, ***the fine or the unpaid portion*** shall be substituted by the appropriate term of imprisonment.	If the convicted person has ceased to serve his sentence, the degree of execution of the sentence or ***the remaining part thereof*** shall be imprisonment.

Ezen kívül megjelenik a *thereof* még a *possession*, *promise*, *respect*, *members* elemekkel alkotott szókapcsolatokban is. Erre az alábbiakban láthatunk példát (7. példa):

7. példa:

CC_HU_REF	CC_HU_EN_HT	CC_HU_EN_MT
Aki gazdálkodó szervezet részére vagy érdekében végzett tevékenységével kapcsolatban jogtalan előnyt kér, avagy a jogtalan előnyt vagy ***ennek ígéretét*** elfogadja,	Any person who requests or receives an unlawful advantage in connection with his activities performed for or on behalf of an economic operator, for himself or for a third party, or accepts a ***promise of such an advantage,***	Any person who, in connection with the activities of an economic entity for or on behalf of the entity, requests an unlawful advantage or accepts an unlawful advantage or ***a promise thereof,***

Ha összevetjük a forrásnyelvi magyar szöveget a gépi fordítással, akkor azt tapasztaljuk, hogy a *thereof* elem az *az* és *ez* mutató névmások ragozott alakjaival (*annak*, *azzal*, *ezzel*, *azok*, *erre*) alkotott főnévi szókapcsolataiban jelenik meg.

Összegzés

A fenti adatok alapján jól látszik, hogy a célnyelvi humán fordításban a *thereof* elem gyakorisága jobban megfelel az angol eredeti referenciaszövegben forrásnyelvében kirajzolódó gyakorisági mintának. Ugyanakkor a *thereof* elem szintaktikai beágyazottsága a kettőtől hat elemből álló szókapcsolatokban az angol eredeti referencia- és a magyarról angolra gépi fordított szöveg között feleltethető meg jobban. A forrásnyelvi magyar szöveggel való összevetésben az látszik, hogy a gépi fordításban a *thereof* kizárólag a mutatónévmások ragozott alakjai és főnevek által alkotott szókapcsolatok fordításaként van jelen. A humán fordításban nagyobb a szórás: a mutatónévmások ragozott alakjai és főnevek által alkotott szókapcsolatokon kívül ragozott főnevekre, sőt akár tagmondatokra referáló funkcióban is előfordul. Sőt arra is van példa, hogy ekvivalensnek tekinthető forrásnyelvi elem hiányában is megjelenik, vagyis redundáns használatban.

Konklúzió

A *thereof* használatának gyakoriságát illetően tehát a humán fordított szöveg jobban követi a forrásnyelvi referenciaszövegben is megfigyelhető gyakoriságot, de a szintaktikai szinteken, a szókapcsolatokban és mondatokban való elhelyezkedésének mintája inkább a gépi és az angol nyelvű referenciaszöveg viszonylatában feltűnő. Ez megerősítheti a Halverson (2003, 2010, 2017) által a fordított szövegben megfigyelhető *gravitációs* hatást, miszerint az adott regiszterre jellemző lexikai vagy grammatikai elemeket nagyobb valószínűséggel választják a fordítók, így azok használata felülreprezentált a fordításokban, sőt akár redundáns is lehet. A célnyelvi szövegben tehát a humán fordító törekszik a jogi szaknyelvi használat jellegzetes eleme, a *thereof* gyakori használatára, ez azonban nem egyezik a referenciának tekintett forrásnyelvi szintaktikai gyakorlattal. A gépi fordított szöveg tehát jobban „gravitál" az angol nyelvű referenciaszövegre jellemző szintaktikai használathoz, a humán fordításban pedig a jogi szaknyelvi regiszter jellegzetes elemének tekintett *thereof* gyakori használata követi jobban a referenciaszövegben megfigyelt mintázatot. Fenti állítás megerősítése

további adatokat és elemzést igényel. Ugyanakkor a jogi szövegek poszteditálása során a *thereof* használatát illetően érdemes a fentiekben leírtakat figyelembe venni.

IRODALOM

Balogh, D. (2020). *Műfajtudatosság a jogi szakfordításban és szakfordítóképzésben*. ELTE Doktori Iskola, kiadatlan doktori disszertáció. https://doi.org/10.15476/ELTE.2020.035

Bázlik, M. & Ambrus, P. & Beclawski, M. (2010). *The Grammatical Structure of Legal English*. Translegis.

DePalma, D. (2013). *Post-editing in practice*. http://www.tcworld.info/e-magazine/translation-and-localization/article/post-editing-inpractice/

Halverson, S. (2003). The cognitive basis of translation universals, *Target*, *15*(2), 197–241. https://doi.org/10.1075/target.15.2.02hal

Halverson, S. (2010). Cognitive Translation Studies and the merging of empirical paradigms: The case of 'literal translation', *Translation Spaces, 4*(2), 310–340, https://doi.org/10.1075/ts.4.2.07hal

Haverson, S. (2017). Gravitational pull in translation. Testing a revised model. In De Sutter, G. & Lefer, M. A. & Delaere, I. (Eds.), *Empirical Translation Studies. New Methodological and Theoretical Traditions* (pp. 9–47). De Gruyter Mouton. https://doi.org/10.1515/9783110459586

Heltai, P. (2014). Pszicholingvisztikai és neurolingvisztikai eljárások a fordítástudományban, *Fordítástudomány*, *16*(2), 5–23.

Hu, K. & Cadwell, P. (2016). A Comparative Study of Post-editing Guidelines, *Baltic J. Modern Computing*, *4*(2), 346–353. https://aclanthology.org/W16-3420.pdf

ISO, ISO 17100:2015: Translation services – Requirements for translation services, http://www.iso.org/iso/catalogue_detail.htm?csnumber=59149

Kjaer, A. L. (2007). Legal Translation in the European Union. A Research Field in Need of a New Approach. In Kredens, K. & Goźdź-Roszkowski, S. (Eds.) *Language and the Law: International Outlooks* (pp. 69–95). Frankfurt am Main: Peter Lang.

Kovács, T. (2018). Átváltási műveletek (megoldások) az angol–magyar jogi szaknyelvi fordításban. *Modern nyelvoktatás. Fordítástudomány különszám*, 25(2–3), 40–53.

Kovács, T. (2020a). Humans, machines, and texts: The implications of the rise of neural machine translation for the educators of future translators. In Besznyák, R., Fischer, M. & Szabó, Cs. (Eds.) *Fit-For-Market Translator and Interpreter Training in a Digital Age (*pp. 135–151*).* Vernon Press.

Kovács, T. (2020b). Interferenciajelenségek összehasonlító vizsgálata nem fordított magyar és angolról magyarra fordított jogi szaknyelvi szövegekben. In Szabóné Papp, J. (Ed.). *Alkalmazott Nyelvészeti Közlemények, Interdiszciplináris tanulmányok, A Miskolci Egyetem Közleményei. Publications of the University of Miskolc*, *13*(2), 150–165. Miskolci Egyetemi Kiadó.

Kovács, T. (2020c). Jogi szaknyelvi és fordítási feladatok. In Dróth, J. (Ed.). *Fordítási és szaknyelvi gyakorlatok távoktatásban*. (pp. 122–179). KRE–L'Harmattan, Károli Könyvek.

Lample, G., Ott, M., Conneau, A., Denoyer, L. & Ranzato, M.A. (2018). *Phrase-based and Neural Unsupervised Machine Translation. Cornell University*, https://doi.org/10.18653/v1/D18-1549

Pavlíčková, E. (2008). Ways of addressing in legal discourse. *Discourse and Interaction*, *1*(1), 87–98. https://journals.muni.cz/discourse-and-interaction/article/view/6901/8267

Papineni, K., Roukos, S., Ward, T. &Zhu W.J. (2002). BLEU: A Method for Automatic Evaluation of Machine Translation. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL*), Philadelphia, July 2002, pp. 311–318. https://doi.org/10.3115/1073083.1073135

Pym, A. (2013). Translation Skill-Sets in a Machine-Translation Age. *Meta: Translators' Journal*, *58*(3), 487–503. https://doi.org/10.7202/1025047ar

Šarčevič, S. (1997). *New Approach to Legal Translation*. Hague: Kluwer Law International.

TAUS. (2016). TAUS Post-Editing Guidelines, https://www.taus.net/think-tank/articles/postedit-articles/taus-post-editing-guidelines

Seidl-Péch, O. (2020). Az internet mint korpusz a szakfordításban. In Adorján, M. & Kovács Tímea (Eds.), *Korpusznyelvészet és nyelvi közvetítés* (pp. 109–125). KRE–L'Harmattan, Károli Könyvek.

Snover, M.,Dorr, B., Schwartz, R., Micciulla, L. & Makhoul, J. (2006). A Study of Translation Edit Rate with Targeted Human Annotation,

Ződi, Zs. (2017). Jogi szövegtípusok. *Magyar Jogi Nyelv*, (2), 20–30. https://joginyelv.hu/jogi-szovegtipusok/

https://www.british-legal-centre.com/source/GLEC_Presentaions/GLEC_Lesson-5.pdf

Felhasznált források

California állam Büntetőjogi Törvénykönyve (Penal Code of California) https://leginfo.legislature.ca.gov/faces/codesTOCSelected.xhtml?tocCode=PEN

A 2012. évi C. törvény a Büntető Törvénykönyvről (hivatalosnak tekinthető) angol fordítása, elérhető: https://njt.hu/translation/J2012T0100P_20210708_FIN.pdf

A 2012. évi C. törvény a Büntető Törvénykönyv gépi (Deep L Pro) angol fordítása

Quantitative and qualitative analysis of compound pronouns in source language and translated criminal law texts

Arising out of their very nature , legal language and translation constitute separate disciplines within the field of translation and languages for special purposes. Several authors have addressed the characteristics of legal language and the main problems of interlingual legal translation that crosses the borderlines between languages, cultures and legal institutional systems. At the same time, the explosion in machine translation in the last few years, and the new and increasingly important field of post-editing to which it has given rise, are posing new challenges to translation studies and to translators in practice. If an ever-increasing number of texts and their translations are becoming easily accessible to almost everyone in electronic format, what new challenges and expectations does this pose for translators and post-editors? How does the concept of a quality standard for translators evolve? How can we define the ideal (legal) translation in a specialised language, which can offer a standard qualitative level for both machine and human translation?

Keywords: legal language use, human translation, machine translation, post-editing, quality standards