

KÓBOR MÁRTA

Gépi fordítás (nem csak) a fordítóképzésekben

Módszertani adalék a tudatos MT-használat oktatásához

A neurális gépi fordítás (NMT) egyre hatékonyabban működik az olyan „ritka” nyelvek esetében is, mint a magyar, ám a továbbra is előforduló hibák miatt az MT-kimenetek utószerkesztésre (PE) szorulnak. Az MT/PE egyre szervezettebb részét képezi a professzionális fordítási munkafolyamatoknak, amihez a fordítóképzéseknek is igazodniuk kell. Bár sokan tartanak attól, hogy a túl korai MT-használat veszélyezteti a fordítási kompetencia kialakulását, a képzéseknek nem szabad figyelmen kívül hagyniuk, hogy a diákok ma már a nyelvtanulás minden szintjén – tehát már a fordítóképzés megkezdése előtt is – használnak gépi fordítóeszközöket. Ezért tiltás helyett érdemesebb az MT korlátainak és lehetőségeinek tudatosítására törekedni, és megfelelő MT-használati kultúrát (MT literacy) kialakítani nem csak a fordítóképzésben, hanem a nyelvoktatásban is. Tapasztalataink szerint több gépi fordítás párhuzamos használata és egybevetése bizonyítható mértékben segíti a kritikus szemlélet kialakítását, és ezáltal a megfelelő kompetenciák fejlesztését.

Kulcsszavak: gépi fordítás, NMT, MT-használat, tudatosítás, fordítóképzés

Bevezető

A gépi fordítóeszközök (köztük a legismertebb Google Fordító és DeepL) az utóbbi években a laikus felhasználók és a magyarhoz hasonló „ritka” nyelvek képviselői számára is egyre meggyőzőbb eredményeket produkálnak. Ma már nem csak a professzionális nyelvi közvetítők (ELIS, 2021, 2022, 2023; Hunnect, 2021), hanem egyéb szakterületek képviselői is előszeretettel használják őket (Kasperé et al., 2021; Seresi et al., 2022). Oktatási intézményekben végzett felmérések egész sora bizonyítja azt is, hogy a gépi fordítás használata mára jócskán túlmutat a fordítóképzések keretein: az egyetemeken nyelvi és egyéb képzéseinek hallgatói mellett már az általános és középiskolai nyelvtanulók is rutinszerűen alkalmaznak gépi fordítóeszközöket, sokszor anélkül, hogy akár ők, akár tanáraik tisztában lennének azok lehetőségeivel és korlátaival (Tsai, 2019; Niño, 2020; Liubinien et al., 2022). Az alábbiakban – a MANYE Fordítástudományi Szakosztálya által 2023 januárjában a gépi fordítás oktatásáról szervezett workshop I. részéhez¹ kapcsolódóan – azokat az ismereteket igyekszem egybegyűjteni, amelyek elsősorban fordítóképzéseinkben, de az idegennyelv-oktatás különféle szintjein is segíthetnek a gépi fordítást a kompetenciafejlesztés szolgálatába állítani.

¹ A MANYE Fordítástudományi Szakosztálya 2023 januárjában két részes online képzést hirdetett a gépi fordításról és az utószerkesztésről fordításoktatóknak. Ez a tanulmány a 2023. január 19-én tartott első workshop anyagának kibővített változata.

A neurális gépi fordítás jellemzői és korlátjai

A mesterséges intelligencián (MI) alapuló ún. neurális gépi fordítás (*neural machine translation* – NMT) 2016-tól váltotta fel a korábban használatos szabályalapú (*rule-based*, RBMT), illetve statisztikai (*statistical*, SMT) fordítórendszereket (Yang, 2018; Prószéky, 2021). Az MI-alapú chatbotok (pl. ChatGPT) fordítási célú használatának – a közeljövőben várható, de még felmérhetetlen hatású – elterjedéséig az NMT-t tekinthetjük a legkorszerűbb MT-megoldásnak. A korábbi technológiákkal ellentétben, amelyek fókuszában előbb a szavak és a nyelvtani szabályok (RBMT), később pedig a korábbi fordításokra épülő párhuzamos korpuszok (SMT) álltak, a hatalmas adatállományokon tanított és önálló mélytanulásra is képes neurális háló-alapú NMT-rendszerek a közvetlen kontextus figyelembevételével készítik el – alapvetően mondat szinten – a célnyelvi változatot. Az NMT ma bizonyos nyelvirányok és szövegtípusok esetében olyan minőségű kimenetet produkál, amely vetekszik a humán fordításokéval, de a ritkább nyelvek esetén is meggyőző eredményeket kínál, amelyek megfelelő utószerkesztés (*post-editing* – PE) esetén akár professzionális célokra is hasznosíthatók.

1. táblázat

Példák NMT-vel fordított rövid szövegrészekre

EN forrásszöveg	MT1_HU: Google Fordító (2023.04.)	MT2_HU: DeepL (2023.04.)
It is still relatively early days when it comes to assessing the true impact of ChatGPT, but AI Chatbot technology is progressing at pace. And with the recent release of GPT-4, eyebrows are being raised in the translation industry. Ask ChatGPT itself and it says it is likely it will have a "significant impact" on the industry and that "it could help to reduce the time and cost of translation services". Great – but the devil is in the detail!	A ChatGPT valódi hatásának felmérése még viszonylag korai szakaszban van, de az AI Chatbot technológia ütemben fejlődik. A GPT-4 közelmúltbeli megjelenésével pedig felvonják a szemöldököt a fordítóiparban. Kérdezd meg magát a ChatGPT-t, és azt mondja, hogy valószínűleg „jelentős hatással lesz” az iparágra, és „segíthet csökkenteni a fordítási szolgáltatások idejét és költségeit”. Nagyszerű – de az ördög a részletekben rejlik!	A ChatGPT valódi hatásának felmérése még viszonylag korai, de a mesterséges intelligencia chatbot-technológia gyors ütemben fejlődik. A GPT-4 közelmúltbeli megjelenésével pedig a fordítóiparban is felvonul a szemöldök. Kérdezzük meg magát a ChatGPT-t, és azt mondja, hogy valószínűleg "jelentős hatással" lesz az iparágra, és hogy "segíthet csökkenteni a fordítási szolgáltatások idejét és költségeit". Nagyszerű - de az ördög a részletekben rejlik!

Forrás: <https://www.rws.com/language-weaver/events/2023/how-will-chatgpt-impact-the-translation-industry/>

A neurális gépi fordítás legfőbb erényeként a gördülékenységet (*fluency*) szokás emlegetni, amelynek azonban sokszor a pontosság (*accuracy*) látja kárát. A neurális technológia térnyerése után megjelenő tanulmányok a korábbi statisztikai rendszerekéhez hasonlították az NMT-kimeneteket, és egyes nyelvirányokban akár 60%-kal kevesebb hibát találtak (Wu et al., 2016). Összességében természetesebbnek ítélték a nyelvezetüket, viszont felhívták a figyelmet az NMT elsőre szinte észrevehetetlen

hibáira – köztük akár félrefordításokra –, valamint a gyakori kihagyásokra, felesleges betoldásokra és ismétlésekre (Yamada, 2019; Loock, 2020; Teixeira, 2020). A kutatók az NMT szintjén máig hiányolják a szövegszintű konzisztenciát (ami logikusan következik a mondat szintű megfeleltetésből), és gyakran kifogásolják a szövegekben előforduló szakkifejezések kezelését: az NMT ugyanis hajlamos téves terminusokat illeszteni a szövegbe és következetlenül használni őket (terminológiai inkonzisztencia) (Lesznyák, 2019).

Az NMT minősége kapcsán mindenképpen fel kell hívni a figyelmet egyrészt a nyelvpárok, illetve szakterületek, másrészt a fordítómotorok² típusának jelentőségére. A motorok teljesítményét ugyanis nagyban befolyásolja a betanításukhoz – és később önálló tanulásukhoz – használt tanítóanyag, ezért egészen más minőségű kimenetek várhatók egy-egy nyelvpárra és/vagy (szak)területre specializált (ún. *custom*) motoroktól, mint a Google Fordító, DeepL típusú általános, univerzális rendszerektől (Teixeira, 2020). Számolni kell ugyanakkor a gépi fordítórendszerek folyamatos fejlődésével is (Szlávik, 2022b), ami megnehezíti a motorok hatékonyságának megítélését vagy rangsorolását. A Google pl. 2017-ben bevezette a Transformer nevű neurális architektúrát³, amely 2020-ra látványos minőségi javulást hozott gépi fordításaiban. 2021-ben a hazai felhasználók mégis a DeepL interaktív, alternatív megoldásokat is kínáló fordításait ítélték használhatóbbnak (Hunnect, 2021), és a két óriás rendszer azóta is fej fej mellett fejlődik, messze felülmúlva pl. a Facebook (Meta) sokszor mósolyt fakasztó gépi fordításait – de egy-egy mondat szintjén akár a specifikus (*custom*) motorok teljesítményét is.

A kifejezetten EU-s szövegek fordításához fejlesztett eTranslation fordítórendszer⁴ gépi kimeneteit vizsgáló *Human in the Translation Loop* kutatócsoport (Robin et al., 2023) első eredményei szintén fejlődésről tanúskodnak: 2022-ben már viszonylag kevés pontatlanság, kihagyás, „dadogás” stb. jellemezte a magyar fordításokat. A terminológia terén azonban még a kifejezetten uniós szövegeken és az IATE⁵ uniós terminológiai adatbázison „edzett” motorokkal sem sikerült igazi áttörést elérni. Fontos megjegyezni, hogy az NMT-technológia különösen alkalmatlan kontextusból kiragadott terminusok (pl. egy autós tartozéklista – lásd a 2. táblázatot) fordítására, mivel a neurálisháló-alapú rendszerek ilyenkor nem tudnak a szavak környezetére támaszkodni, ami nélkülözhetetlen a hatékony működésükhöz.

² A gépi fordítást végző számítógépes eszközöket – a köznyelvben tévesen elterjedt fordítóprogram helyett – gépi fordítóeszközöknek vagy -motoroknak nevezzük, és ebben a cikkben csak az írásbeli nyelvi közvetítés szempontjából vizsgáljuk. Fordítórendszer alatt több fordítómotor összességét értjük (pl. az uniós eTranslation fordítórendszer ma 10-féle motort foglal magában).

³ A transformer modell szolgál azóta számos „okos” alkalmazás, köztük a várhatóan a fordítás és a nyelvtanítás világát gyökeresen átalakító ChatGPT alapjául is (amellyel ebben a tanulmányban csak érintőlegesen foglalkozunk, mivel – várható hatásaira tekintettel – ez a téma külön tanulmányt érdemel).

⁴ Oktatási intézmények számára regisztráció után ingyenesen használható fordítórendszer: <http://webgate.ec.europa.eu/etranslation>

⁵ Az Európai Unió interaktív terminológiai adatbázisa: <https://iate.europa.eu/home>

2. táblázat

Példák egy kontextus nélküli terminuslista humán és gyakorlatilag használhatatlan gépi fordításaira

EN	FR	HU	DeepL - EN>HU	DeepL - FR >HU	eTranslation - EN>HU	eTranslation - FR>HU
PROTECTIVE SUN SHADES - SIZE XL	Chaussettes de protection solaire - Taille XL	Napfényvédő huzat - XL méret	VÉDŐ NAPSZEMÜVEG - XL MÉRET	Napvédő zokni - XL méret	VÉDŐ NAPERNYŐK - XL MÉRET	Napvédő zokni - XL méret
Rigid boot tray	Bac de coffre rigide	Merev csomagtértálca	Merev csomagtartó tálcá	Merev csomagtartó tálcá	Merev indítótálca	Merev mellkas tálcá
Boot mat with bumper protection	Tapis de coffre avec protection bouclier	Csomagtérszőnyeg lökhárítószerűvel	Csomagtartószőnyeg lökhárítóval	Pajzsvédővel ellátott csomagtérszőnyeg	Cipőszőnyeg lökhárító védelemmel	Mellszőnyeg pajzs védelemmel
Luggage compartment net	Filet de coffre	Csomagrögzítő háló	Csomagtérháló	Törzsháló	Csomagter háló	Csomagtartó háló
Thermoformed boot chocks	Cales de coffre	Csomagrögzítő ékek	Hőformázott bakancsrögzítők	Trunk ékek	Hőformázott csizmák	A melltartók
Boot organizer bag	Sac de coffre	Csomagtéri tárolózsák	Csizma szervező táská	Csomagtartó táská	Boot szervező táská	Mellkasi táská

A nyelvirányok szerepére és az NMT minőségi változásaira visszatérve megállapíthatjuk, hogy a domináns – főként indoeurópai – nyelvpárok esetén méltán dicsért gördülékenység angol–magyar viszonylatban már nem mindig meggyőző, a stílus ugyanis a ritkább nyelvirányokban kifejezetten sok (és egyre több) kívánnivalót hagy maga után (Szlávik, 2022b). E tendencia vélhetőleg az utószerkesztés nélküli gépi fordítások mennyiségének növekedésével és tanulóanyagba visszakerülésével magyarázható, és a jövőben határt szabhat a gépi fordítás minőségi fejlődésének. A magyar nyelvű NMT-kimenetekben emellett továbbra is előfordulnak központozási és tipográfiai hibák (a DeepL pl. 2023-ban is felülre teszi a nyitó idézőjeleket), és egyik motor kimenete sem mentes a nyelvtani (pl. ragozási, egyeztetési) stb. problémáktól (további példákért lásd Szlávik, 2023).

A neurális gépi fordításról összességében elmondható, hogy felülmúlja a korábbi technológiák hatékonyságát és folyamatosan fejlődik. Az NMT minősége – különösen néhány domináns nyelvirány és speciális (*custom*) motor esetében – egyre inkább közelít a humán fordításokéhoz, és hibái is sokban hasonlítanak a humán fordítóktól megszokottakhoz. Az ingyenesen elérhető ún. általános motorokkal magyarról, illetve magyarra fordított szövegeket azonban továbbra is erős fenntartásokkal kell kezelni. A professzionális felhasználók számára egyértelmű, hogy az ún. nyers MT-kimenet feltétlenül utószerkesztésre szorul, az NMT korlátait azonban a laikus felhasználókban és a nyelvtanulóknak is tudatosítani kell, akik hajlamosak a gépi fordítást egy számológép megkérdőjelezhetetlen eredményéhez hasonlítani (Loock & Léchaugette, 2021).

Gépi fordítás a professzionális fordítói gyakorlatban

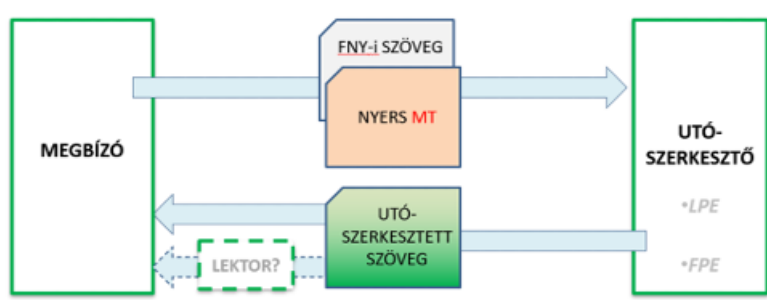
Az elmúlt években készült tanulmányok és felmérések egész sora bizonyítja, hogy az NMT-t széles körben használják a hivatásos nyelvi közvetítők: a szabadúszó fordítók éppúgy, mint az irodák és a nagy fordítószervek alkalmazottai (ELIS 2021, 2022, 2023; Hunnect, 2021). A nyelvi közvetítők mellett egyre gyakrabban élnek a gépi fordítás lehetőségével más szakterületek képviselői is, a közigazgatástól az üzleti élet számos területén át az orvosi, mérnöki stb. tudományokig (Kasperé et al., 2021). A laikus felhasználók általában az ingyenesen elérhető online motorok szolgáltatásait használják, jellemzően böngészőjükben (ritkábban okoseszközökre telepített alkalmazással), a megfelelő szövegdobozba beírva vagy bemásolva a fordítandó szöveget. A

professzionális nyelvi közvetítésben a gépi fordítás (MT) és a tőle gyakorlatilag elválaszthatatlan utószerkesztés (PE) (a szakmai zsargonban sokszor összevontan: MTPE) ma egyre gyakrabban és egyre változatosabb formákban jelenik meg, alaposan felforgatva a korábbi munkafolyamatokat és szerepeket.

Az ISO 18578:2017 szabvány által szabályozott, ám a becsléseknél valójában kisebb arányban jelen lévő (ELIS, 2023) „tisztán” MTPE munkafolyamatban már nem is jelennek meg a hagyományos értelemben vett fordítási folyamatnak sem a szereplői (fordító, lektor), sem pedig olyan meghatározó dokumentumai, mint a (lektorálatlan, illetve lektorált) fordítás. A forrásnyelvi szöveg mellett itt egy nyers gépileg fordított fájl kap központi szerepet, melynek részleges (*light post-editing* – LPE), illetve teljes (*full post editing* – FPE) utószerkesztése eredményeképpen születik meg a célnyelvi szöveg. Ennek minőségellenőrzéséről a – jelenleg felülvizsgálat alatt álló – szabvány eredeti kiadása nem rendelkezik (Szlávik, 2022a).

1. ábra

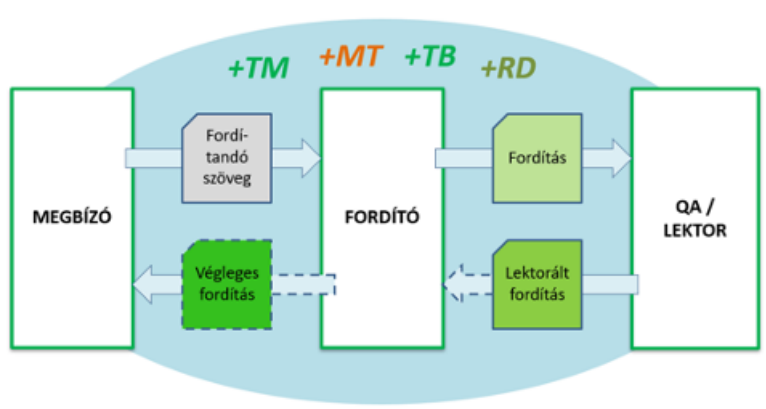
Az ISO 18578:2017 szabvány által szabályozott „tisztán” MTPE



A nyelvi közvetítés szakemberei azonban tudják és nap mint nap tapasztalják, hogy a gyakorlatban a gépi fordítás ma még (vagy talán egyre) inkább az ISO 17100:2015 szabvány által meghatározott fordítási folyamatba integráltan, a fordító rendelkezésére álló segédeszközök egyikeként jelenik meg. A mai fordító fordítástámogató eszközeinek (CAT – *computer aided/assisted translation*, pl. Trados, MemoQ, Phrase stb.) jellemzően integrált felületén napjainkban már nem csak a megbízótól vagy korábbi megbízásokból származó fordítómemóriák (*translation memory* – TM), terminológiai adatbázisok (*term base* – TB), illetve egyéb kapcsolódó dokumentumok (*reference document* – RD) kapnak helyet. Egyre gyakrabban megjelenik mellettük a gépi fordítás, amelyet adott esetben a megbízó bocsát a fordító rendelkezésére. A fordítók azonban maguk is gondoskodhatnak gépi támogatásról: megfelelő API vagy plug-in segítségével az integrált CAT-felületen, vagy akár böngészőjükön keresztül is segítségül hívhatják az MT-t – hosszabb távon pedig várhatóan az MI-alapú chatbotokat is – „fordításuk” elkészítéséhez. Az idézőjel használata itt azért indokolt, mert a mai fordító ilyen körülmények között akár egyetlen mondat hagyományos értelemben vett fordítása nélkül is teljesíthet egy megbízást: vagy a fordítómemóriából érkező találatok szerkesztésével, vagy – használható TM-találat híján – a gépi kimenet utószerkesztésével.

2. ábra

Az ISO 17100:2015 szabvány által szabályozott fordítási folyamatba integrált MTPE



A helyzetet tovább bonyolítja, hogy ebben – a szakirodalomban MT-vel (és/vagy MI-vel) támogatott vagy „gyorsított” fordításként (*MT-enhanced translation* – EMT, 2022) is emlegetett – integrált folyamatban az utószerkesztett gépi fordítások (az esetek többségében lektorálás után) bekerülnek a fordítás során folyamatosan bővülő fordítótömörébe, és egy következő projektben ezek már nem MT-, hanem TM-találatként jelennek meg, és adott esetben a gépi motor tanítókorpuszába is visszakerülhetnek (ez történik pl. az uniós intézményeknél, ahol rendszeres időközönként „újrataníják” az eTranslation rendszert, de léteznek már valós időben tanuló, ún. adaptív gépi fordító-rendszerek is).

A professzionális nyelvi közvetítésben tehát a gépi fordítás ma már a munkafolyamat és a végeredmény szintjén is egyre inkább összemosódik a humán fordítással. Mielőtt megnéznénk, milyen következményekkel jár ez a fordítóképzésekre nézve, mindenképp említést kell tennünk a gépi fordítás használatának etikai vonatkozásairól. Ezek professzionális kontextusban leginkább adatvédelmi jellegűek: az ingyenes gépi fordítóeszközök használata során ugyanis megbízónk szövegeit – így a bennük szereplő adatokat – óhatatlanul megosztjuk az eszközök fejlesztőivel, pedig a fordítókra általános titoktartási kötelezettség vonatkozik. Érzékeny információkat és adatokat tartalmazó szövegek esetén ez komoly kockázatot jelenthet a megbízóra nézve (Nitzke et al., 2019). Az árázással kapcsolatban is sokféle kérdés felvetődik: vajon indokolt-e kedvezményt, és ha igen, mekkorát, ha egy fordító ingyenes gépi támogatással „gyorsítja” munkáját? Kötelező-e erről tájékoztatni a megbízóját? Rengeteg olyan kérdés vár tehát még megválaszolásra, amelynek a fordítók nagy része nincs is tudatában, amikor gépi támogatást használ.

Gépi fordítás a professzionális fordítók képzésében

A fentiek ismeretében aligha meglepő, hogy a gépi fordítás használatának és oktatásának kérdései a világon mindenütt foglalkoztatják a fordítóképző intézmények oktatóit és a már gyakorló fordítókat tömörítő szakmai szervezetek munkatársait is. Logikus-

nak tűnne, hogy a képzések gyorsan reagáljanak a piaci fejleményekre, az MT/PE oktatásával kapcsolatban mégis bizonytalanság érzékelhető.

Az Európai Bizottság Fordítási Főigazgatósága 2007-ben létrehozott egy nemzetközi szakértői testületet a mesterfokú fordítóképzés egységes rendszerének kialakítása, illetve az alapjául szolgáló kompetenciamodell kidolgozása céljából. Az EMT (*European Masters of Translation*) mára világszerte referenciának tekintett kompetenciamodelljének legfrissebb változata (EMT, 2022) elismeri, hogy a gépi fordítás egyre nagyobb részt képvisel a fordítási munkafolyamatokban, mégis csupán néhány általános és óvatos utalást tesz az MT/PE-re. A GALA⁶ nevű nemzetközi szakmai szervezet is a tervezettnél később hozta nyilvánosságra MTPE-oktatási ajánlásait (Gene et al., 2023), melyben problémásnak ítélik az MT-oktatás jelenlegi helyzetét.

Magyarországon évek óta élénk párbeszéd folyik a fordítóképző intézmények között: már több műhely- és kerekasztal-beszélgetésre, széles körű egyeztetésre került sor e témáról. A BME-n 2022 októberében rendezett tanár–tanár kerekasztal eredményeképpen közös ajánlások is napvilágot láttak (Robin & Eszenyi, 2022) arról, hogy mikor és milyen formában javasolt a gépi fordítás használatát és az utószerkesztés oktatását bevezetni fordítóképzéseinkbe. Az ajánlások összhangban állnak azon vélekedésekkel (Liu et al., 2022), melyek szerint a gépi fordítóeszközök túl korai használata gátolhatja a tanulási folyamatot, nevezetesen a fordítási kompetencia elsajátítását, ezért bevezetésüket célszerűbb a fordítástechnikai alapok megszerzése utánra (a hazai ajánlások szerint a képzés második felére) időzíteni, a „minősítő” feladatok (pl. képesítő fordítások) esetében pedig később is korlátozni. A KRE 2023. februári kerekasztal-beszélgetésén⁷ a képzőintézmények képviselői mégis nagyon változatos gyakorlatokról számoltak be, melyek némelyike az ajánlásoktól eltérve korábbi MT-használatot is enged, sőt tiltás helyett egyenesen a képesítő fordítások részévé teszi a gépi fordítás használatát (Veresné Valentinyi, 2022).

Az eltérő gyakorlatok megítéléséhez fontos szem előtt tartanunk, hogy – a CAT-eszközökkel szemben, amelyekkel jellemzően csak a képzés második felében, egyetemi licenceknek köszönhetően ismerkednek meg a hallgatók – az ingyenesen és online elérhető gépi motorokat általában már a fordítóképzés megkezdése előtt, illetve az esetleges tiltások ellenére is használják (O’Neill, 2019; Fáy & Yang, 2023). A folytatásban látni fogjuk, hogy ez sok esetben helytelenül és ellenőrizetlenül történik, ami szintén hátráltathatja a megfelelő kompetenciák kialakulását (Bowker, 2020).

Gépi fordítás az idegennyelv-oktatásban

Az MT-használat világszerte egyre népszerűbb és gyakoribb az idegennyelv-szakos diákok (Tsai, 2019; Loock & Léchaugette, 2021), egyéb szakterületek hallgatói (Dorst et al., 2022; Bin Dahmash, 2020) és a fiatalabb nyelvtanulók körében is (Liubinien et al., 2022). Az utóbbi években számos tanulmány és felmérés vizsgálta a gépi fordításnak az idegennyelv-tanulásra gyakorolt hatásait (Ducar & Shocket, 2018; Tsai, 2019;

⁶ Globalisation and Localisation Association: <https://www.gala-global.org/>

⁷ Tiltjuk, türjük vagy támogatjuk? A MANYE Fordítástudományi Szakosztályának kerekasztala, 2023. február 16.

Niño, 2020; Resende & Way, 2021), a Google Fordító nyelvoktatásban játszott szerepéről külön konferenciát is szerveztek⁸. Az eredmények között pozitív és negatív hatások is fellelhetők: egyes kutatók a gépi fordítások hibás megoldásainak rögzülésében rejlő veszélyekre (Resende & Way, 2021), mások a potenciális tanulási lehetőségekre (Niño, 2020) és a megfelelő tanári hozzáállás jelentőségére (Lee, 2019) hívják fel a figyelmet, míg láthatunk példát rövid távon áldásos, de hosszú távon nem hasznosuló hatásokra is pl. a gépi fordítással „előállított” L2-fogalmazások kapcsán, melynek ismeretlen lexikai elemei nem épülnek be az MT-használó diák szókincsébe (Liubinien et al., 2022). Érdeemes megjegyezni, hogy a nyelvtanulók többnyire L1–L2 nyelvirányban használják az MT-kimeneteket, leginkább szövegalkotási feladataikhoz – az L2–L1 irányú gépi fordításokat többnyire csak olvassák (szövegértési célból).

Alaposan megvizsgálva a kutatások eredményeit, egyértelmű tendenciákat és mintázatokat figyelhetünk meg, ha a különböző csoportokat elhelyezzük egy képzeletbeli egyenes mentén, melynek egyik végére a fiatal, kezdő nyelvtanulókat, tőlük jobbra a középhaladó szintű idegennyelvűtudással rendelkező, nem nyelvszakos diákokat, másik vége felé pedig a nyelvi, illetve fordítói szakképzésben is részesülő „haladókat” pozícionáljuk. A motorválasztás terén pl. megfigyelhetjük, hogy a fiatal és kezdő nyelvtanulók szinte kizárólag a Google Fordítót használják, egyéb motorokat inkább csak haladó szinten, illetve bizonyos országokban preferálnak (pl. Franciaországban a DeepLt, Kínában a Baidut). A motorok használatában is jelentős különbségek mutatkoznak: a kezdő nyelvtanulók és gyakran még a nyelvszakos diákok is előszeretettel használják a Google Fordítót – jellemzően okostelefonjukon – az általuk nem ismert szavak és kifejezések megfelelőinek kereséséhez. Erre a jelenségre érdemes figyelni, hiszen az NMT jellemzőinek bemutatásakor láthattuk, hogy a neurálháló-alapú rendszerek az egyes szavak megfelelőire mindig a mondaton belüli környezetük alapján tesznek javaslatot, ezért nem igazán alkalmasak kontextus nélküli szavak fordítására, vagyis szótárszerű használatra. Ezért kiemelten fontos annak tudatosítása, hogy az ismeretlen szavak megértését az ún. korpuszalapú szótárak (Glosbe, Linguee) sokkal hatékonyabban segítik, hiszen azok a keresett szó(kapcsolato)t egy párhuzamos szövegeket – nagyrészt lektorált, pl. uniós fordításokat – tartalmazó fordítómemóriában, autentikus környezetben, valós használatban tárják elénk.

⁸ Google Translate and Modern Languages in Education: <https://www.nottingham.ac.uk/conference/fac-arts/clas/google-translate-and-modern-languages-education/index.aspx>

3. ábra

Példa egy FR terminus (code rural) MT-, illetve korpuszalapú szótár által javasolt fordítására

The screenshot shows the Glosbe website interface for translating 'code rural' from French to Hungarian. The search bar contains 'Fordítás francia nyelvről'. The results section shows 'code rural' with a warning icon and text: 'Jelenleg nincs fordítás az **code rural** szóra a szótárban, esetleg hozzáadhat egyet? Ügyeljen arra, hogy ellenőrizze az automatikus fordítást, a fordítási memóriát vagy a közvetett fordításokat.' Below this, there are search results for 'code rural' with references to French legal articles, such as '17 L'article R*553-10 du code rural prévoit : 17 A mezőgazdasági törvénykönyv R*553-10 cikke kimondja:' and '80 Anheuser-Busch maintient, par ailleurs, que l'article L. 641-2, quatrième alinéa, du code rural est applicable. 80 Az Anheuser-Busch másfelől azon az állásponton van, hogy a mezőgazdasági törvény L. 641-2. cikkének negyedik bekezdése alkalmazandó.'

Szintén egyértelmű összefüggés figyelhető meg a nyelvtanulás és/vagy képzettség szintje, valamint a fordítómotorok iránti bizalom között. A kezdő és kevésbé autonóm nyelvhasználók hajlamosak teljesen megbízni a gépi fordításban (Rossetti et al., 2020), de még a magukat kritikusként tartó idegennyelvszakos diákok is csak a hibák kevesebb mint 40%-át azonosították egy franciáról angolra fordított szövegben (Loock & Léchauguet, 2021). A kutatások szerint magas szintű, magabiztos nyelvtudásra van szükség ahhoz, hogy az MT-használók kritikusan viszonyuljanak, és hozzá merjenek nyúlni a kezdők által „készpénznek vett” MT-kimenethez. Yamada (2020) e jelenséget azzal magyarázza, hogy amíg a tanult idegen nyelvre géppel fordított szöveg (pl. egy fogalmazási feladat) nyelvi szintje magasabb az MT-használóénál, addig nyilván nem lesz képes annak megítélésére vagy korrigálására. Professzionális értelemben vett utószerkesztésről (PE) – amelyet a szakemberek jellemzően anyanyelvükön végeznek – a nyelvtanulás kontextusában nyilván nem lehet beszélni, de a kutatások egyértelműen arra utalnak, hogy a gépi fordítás (MT) ma már megkerülhetetlen része az idegennyelv-oktatásnak.

Magyarországon egyelőre kevés adat áll rendelkezésre a nyelvtanulók és a gépi fordítás viszonyáról, csupán néhány kutatás (Lesznyák et al., 2022) enged betekintést a diákok eszközhasználati szokásaiba. A Pécsi Tudományegyetem Fordítástudományi Kutatóközpontjának munkatársai jelenleg dolgoznak egy általános és középiskolás diákok, illetve tanáraik, valamint nyelv-, fordító-, tanár- és egyéb szakos egyetemisták MT/(MI)-használati szokásaira is kiterjedő felmérésen. Ennek pilotjai és első eredményei a nemzetközi szakirodalmi eredményekből kirajzolódó tendenciákat látszanak igazolni: a magyar általános és középiskolások, valamint a fordítói képzettségben nem részesülő egyetemisták is leginkább a Google Fordítót használják, sok esetben ismeretlen szavak kereséséhez (ennek korlátaira

már utaltunk), valamint L2 szövegalkotási feladataik elkészítéséhez (ami ellenőrizetlen körülmények között nem segíti a fejlődésüket, hiszen általában megelégszenek a gépi fordítás kritika nélküli átvételével). Az első eredményekből arra is következtethetünk, hogy a nyelvtanárok sem rendelkeznek mélyreható ismeretekkel a gépi fordításban rejlő lehetőségekről és kockázatokról – az esetek többségében nem is sejtik, vagy nem veszik észre, hogy diákjaik ilyen eszközök segítségével készítik el feladataikat (és a helyzet várhatóan tovább fog bonyolódni a ChatGPT használatának rohamos terjedésével).

A gépi fordítás nyelv tanulási célokra történő használata is felvet etikai kérdéseket: elfogadható-e a nyelv tanulók részéről gépi segítséget igénybe venni feladataik elkészítéséhez? Ugyanakkor a kérdés más oldalról is megközelíthető: vajon szabad-e a nyelv tanároknak magukra hagyniuk diákjaikat ezekkel az eszközökkel, melyek ellenőrizetlen használata akár hátráltatja is nyelvi kompetenciáik fejlődését?

Módszertani javaslatok a tudatos MT-használathoz

A fentiekre tekintettel megállapíthatjuk, hogy a gépi fordítás (és általában véve az MI) szerepét nagyon alaposan át kell gondolni – nem csak a fordítóképzésekben, hanem ma már az idegennyelv-oktatás különböző szintjein és a tanárképzésben is. A „tiltjuk, tűrjük vagy támogatjuk?” dilemma megoldását e sorok szerzője egy negyedik „t-ben”, a tudatosításban látja, amelyre az EMT-modell is felhívja a figyelmet (EMT, 2022): az MT-ben rejlő lehetőségeket és kockázatokat az oktatókban és a diákokban is tudatosítani kell a gépi fordítás megfelelő használatához.

Az utóbbi években egyre szélesebb körben hivatkoznak (Bowker & Buitrago Ciro, 2019; O'Brien & Ehrensberger-Dow, 2020; Bowker, 2021; Liu et al., 2022; EMT, 2022) az *MT literacy* fogalmára, amely a tudatos MT-használathoz szükséges ismereteket, készségeket és attitűdöket foglalja magában. Bowker & Buitrago Ciro (2019, p. 88) definíciója szerint az *MT literacy* elemei a következő három kérdés körül artikulálódnak:

- Hogyan működnek a gépi fordítórendszerek?
- Hogyan használhatók a gépi fordítórendszerek? (Hogyan választhatók ki a gépi fordításra alkalmas szövegek és motorok?)
- Mikor és hogyan lehet, illetve kell módosítani a gépi fordítás kimenetét?

Az alábbiakban összefoglaljuk azokat a – nagyrészt e cikkben és/vagy Szlávik (2023) kapcsolódó tanulmányában körüljárt – témákat, amelyekről a szakirodalmi adatok és személyes tapasztalataink szerint ma a hazai nyelvoktatás és fordítóképzés különböző szintjein a tanároknak és a diákoknak is fontos volna ismeretekkel rendelkezniük (a cikkekben nem említett témákhoz hivatkozásokat fűzünk, illetve felhívjuk a figyelmet arra, hogy a lista a nyelvtechnológia és a mesterséges intelligencia gyors fejlődése miatt folyamatosan változhat). A listán dőlt betűvel jelöljük azokat az elemeket, amelyek hallgatói tudatosítására inkább csak a professzionális fordítóképzésben van szükség. Természetesen a többi ismeret és készség közül is a mindenkori pedagógiai célok és körülmények függvényében célszerű válogatni. Az is nyilvánvaló, hogy a motorok működésére vonatkozó ismeretek átadása leghatékonyabban gyakorlat, vagyis az MT használata közben történhet, és a mindig a diákjaink nyelvi szintjéhez, tanulási igényeihez és lehetőségeihez kell igazítani.

Hogyan működnek a gépi fordítórendszerek? (az MT-re vonatkozó általános ismeretek)

- *(a gépi fordítás történetének fő fejezetei: RBMT, SMT, NMT – és a várható MI-fejlemények)*
- a mai (neurális) MT-technológia lényege; mondatszintű megfeleltetés (NMT vs online/korpuszalapú szótárak / *terminológiai adatbázisok* [Seidl-Péché, 2018]); a ChatGPT fordításpecifikus lehetőségei
- a neurális motorok típusai: ingyenes, online elérhető, általános eszközök (Google Fordító és DeepL) vs *fizetős vagy regisztrációval elérhető specifikus (custom) rendszerek* (pl. *eTranslation, Globalese*); telepíthető alkalmazások; *CAT-eszközbe integrálható szolgáltatások, API-k és plug-inok, GT4T típusú „többkimenetes” eszközök*
- a gépi kimenetek általános jellemzői (erősségek és gyengeségek), jellemző NMT-hibák
- a különböző eszközök kimenetei közötti különbségek, folyamatosan változó, fejlődő minőség, *gépi metrikák* (Yang, 2018)
- adatvédelmi és etikai kérdések és kockázatok

Hogyan használhatók a gépi fordítórendszerek? (MT-használati készségek)

- ingyenes, online elérhető általános gépi fordítóeszközök (Google Fordító és DeepL) irányított használata (egész mondatok fordítása) böngészőben, *a kapcsolódó alkalmazások telepítése és használata*
- egyszerű szövegbevitel és a gépi kimenet kimásolása billentyűkombinációk vagy a vágólap(ikon) segítségével
- a DeepL rendszer interaktivitásának kipróbálása: az alternatív javaslatokban rejlő (lexikai és szintaktikai / *fordítási*) kompetenciafejlesztési lehetőségek hasznosítása
- *a professzionális munkafolyamatok igényeinek megfelelő teljesfájl-fordítási lehetőségek kipróbálása*
- *CAT-kompatibilis kimenetek generálása; a CAT-eszközbe integrált MT-használatot lehetővé tévő API-k és plug-inok aktiválása egyetemi licencekkel / ingyenes próbaverziókkal; CAT-en belüli gépi fordítás; MT (/MI) és TM kombinált használata*
- a motorok teljesítményének tesztelése, összehasonlítása, értékelése
- *fokozott adatbiztonságot jelentő fizetős szolgáltatások (próbaverzióinak) kipróbálása*

A gépi fordításra terjedelmi okokból (pl. rövid szavak), vagy jellegük miatt (kreatív, kulturálisan beágyazott stb.) alkalmatlan szövegekre az oktatás minden szintjén találhatunk példákat. A hatékony motorválasztás gyakorlásához tudatosítani kell a diákokban az eszközök felhasználói felülete és lehetőségei, valamint a kimeneteik közötti különbségeket, amihez szintén bármilyen – akár 1-2 mondatos, de szemléletes – szöveg használható. A lényeg az, hogy a diákok minél több eszköz működésével és kimeneteivel megismerkedjenek, és megtapasztalják, hogy egyszer (vagy akár csak egy szövegrész esetében az) egyik, másszor egy másik motor „sikeresebb”. Az összehasonlító gyakorlatokat a nyelvoktatásban is érdemes L2–L1 nyelvirányban végezni, mert a diákok az anyanyelvi kimenetekben könnyebben észreveszik az MT hibáit és korlátait, ezáltal ki tud alakulni bennük egyfajta

„egészséges szkepticizmus” (O’Brien & Ehrensberger-Dow, 2020) a gépi fordítással kapcsolatban. Így könnyebben megértik, hogy az L2-kimeneteket sem kezelhetik gondolkodás nélkül, amikor pl. szövegalkotási feladathoz használnak gépi „támogatást”, és pl. a DeepL alternatív javaslatait L2-eszköztáruk bővítésére használhatják.

Személyes fordítói és oktatói tapasztalataim szerint több motor kimenetének egymás mellé állítása sokat segít a kritikus szemlélet kialakításában. A módszer hatékonyságát egy – egyelőre kis mintás – felmérés friss eredményei is igazolják (Kóbor, 2023b). A módszer tesztelése érdekében a PTE és az ELTE 10-10 fordítóképzős (utószerveztési ismeretekkel és tapasztalattal nem rendelkező) hallgatóját arra kértük, keressék meg és jelöljék a hibákat a Politico magazin egy rövid (190 szavas) cikkének DeepL-lal (DL – lásd a 3. táblázat MT1-oszlopában), illetve Google Fordítóval (GF – MT2-oszlop) készített gépi fordításaiban.

3. táblázat

„Hibavadászat” egy vs több gépi kimenetben (2023. februári felmérés)

EN forrásszöveg	MT1_HU	MT2_HU
<p>EU vows more cash for frontier policing as border fence debate revives</p> <p>The EU that once largely dismissed border walls as a crude Trumpian solution is vowing to channel “substantial” funds into border guards and frontier surveillance equipment, as countries ramp up calls for help paying for their border fences.</p> <p>For several hours at an EU Council summit on Thursday, the only show in town was Volodymyr Zelenskyy. Then the Ukrainian leader left, and a well-worn clash over migration instantly returned. There were stark signs that the bloc’s stance is hardening.</p> <p>Well into the early hours of Friday morning, EU leaders made proposal after proposal, all seeking to stem the rise in people arriving to the Continent outside legal channels. Some wanted Brussels to help pay for border fences. Others stressed a focus on returning rejected asylum seekers.</p> <p>German Chancellor Olaf Scholz at one point tried to calm passions, according to one official familiar with the discussion. Did the EU want to turn itself into a fortress? Walls, simply put, do not work, he said, pointing to the U.S.-Mexico border, where a fractious debate over a wall hasn’t abated crossings.</p>	<p>Az EU több pénzt ígér a határrendészet-re, miközben újraéled a határkerítésről szóló vita</p> <p>Az EU, amely egykor nagyrészt elutasította a határfalakat, mint durva trumpi megoldást, ígéretet tett arra, hogy “jelentős” pénzeszközöket irányít a határőrök és a határőrizeti berendezések finanszírozására, miközben az országok egyre gyakrabban kérnek segítséget a határkerítések kifizetéséhez.</p> <p>A csütörtöki uniós tanácsi csúcstalálkozón több órán át csak Volodimir Zelenskij volt a főszereplő. Aztán az ukrán vezető távozott, és a migrációval kapcsolatos jól ismert összecsapás azonnal visszatért. Éles jelei voltak annak, hogy a blokk álláspontja keményedik.</p> <p>Péntek kora reggelig az uniós vezetők javaslatról javaslatra tettek javaslatokat, amelyek mindegyike a legális csatornán kívül a kontinensre érkező emberek számának növekedését igyekezett megfékezni. Egyesek azt akarták, hogy Brüsszel segítsen a határkerítések finanszírozásában. Mások az elutasított menedékkérők visszaküldésére helyezték a hangsúlyt.</p> <p>A megbeszélést ismerő egyik tisztviselő szerint Olaf Scholz német kancellár egy ponton megpróbálta csillapítani a kedélyeket. Az EU erődde akarta változtatni magát? A falak, egyszerűen szólva, nem működnek - mondta, és az amerikai-mexikói határa mutatót, ahol a falról folytatott heves vita nem csökkentette az átkeléseket.</p>	<p>Az EU több készpénzt ígér a határrendészet számára, ahogy a határkerítésről szóló vita élénkül</p> <p>Az EU, amely egykor nagyrészt elutasította a határfalakat, mint durva Trump-féle megoldást, azt ígéri, hogy „jelentős” pénzeszközöket fordít a határőrökre és a határőrizeti berendezésekre, miközben az országok egyre több segélykérést kérnek határkerítéseikért.</p> <p>Az EU Tanácsának csütörtöki csúcstalálkozóján több órán át az egyetlen előadás a városban Volodimir Zelenskij volt. Aztán az ukrán vezető távozott, és azonnal visszatért a jól megkopott összecsapás a migráció miatt. Éles jelek mutatkoztak annak, hogy a blokk álláspontja megkeményedik.</p> <p>Péntek hajnalban az EU vezetői javaslatról javaslatra tettek javaslatot, és mindegyik arra törekedett, hogy megállítsák a legális csatornán kívül a kontinensre érkezők számának növekedését. Néhányan azt akarták, hogy Brüsszel segítsen kifizetni a határkerítéseket. Mások hangsúlyozták, hogy az elutasított menedékkérők visszatérésére összpontosítsanak.</p> <p>Olaf Scholz német kancellár egy ponton megpróbálta csillapítani a szenvedélyeket – állítja egy tisztviselő, aki jól ismeri a vitát. Az EU erődde akarta változtatni magát? Egyszerűen fogalmazva, a falak nem működnek – mondta, és az Egyesült Államok-Mexikó határa mutatót, ahol a fal körül heves vita nem csökkentette az átkelők számát.</p>

Source: <https://www.politico.eu/article/euco-eu-crosses-into-the-border-fence-game-migration/>

Az egyik csoport előbb a DL kimenetében dolgozott, majd 10 perccel később megkapta mellé a GF gépi kimenetét is, a kontrollcsoportnak viszont csak a GF-kimenet állt rendelkezésére. Hipotézisünk az volt, hogy több gépi fordítás egymás mellé állításakor a motorok eltérő megoldásai mintegy ráirányítják a hallgatók figyelmét a problémás szövegrészekre, ami segít nekik a hibák észrevételében. A felmérés eredményei igazolták a várakozásokat: a kétféle kimenettel dolgozó csoport tagjai a GF-kimenetben általunk előzetesen azonosított 37 (megfeleltetési, terminológiai, nyelvi és stilisztikai) hiba 42%-át megtalálták (átlagosan 15,41 db-ot), míg a kontrollcsoport tagjai csupán a hibák 33%-át azonosították (12,12 db-ot). A két gépi fordítás párhuzamos jelenléte a DL megoldásaival szemben is kritikusabbá tette a hallgatókat: a GF-kimenet elemzése közben átlagosan további 3,24 db hibát fedeztek fel a DL fordításában, tehát a dupla kimenet mindkét gépi fordítás iránti hibaérzékenységet növelte, gyakorlatilag minden hibakategóriában (de leginkább a nyelvi és tipográfiai hibák terén).

Egy ilyen, pl. óra elején, bemelegítő feladatként végzett „hibavadászat” nyelvórán is használható a kritikus szemlélet kialakítására. A fordítóképzésekben ezek a feladatok az utószerkesztés előkészítésére szolgálhatnak: a nyers MT-kimenet hibáinak észlelése ugyanis alapvető feltétele a gépi fordítás sikeres utószerkesztésének.

Több motor fordításának párhuzamba állítása azonban nem csak a hibák és korlátok, hanem a gépi fordításban rejlő lehetőségek tudatosításában is segíthet: hibák helyett a pozitívumokra, a jónak/jobbnak tűnő megoldásokra is „vadászhatunk” a diákokkal. Egy másik kismintás felmérésünk szerint (Kóbor, 2023) L2–L1 nyelvirányban már az első éves fordítóhallgatók nagy biztonsággal ki tudják választani a „legjobb” megoldást három gépi kimenet közül. Annak tudatosítása, hogy a megoldások között vannak minőségi különbségek, kevésbé jó és jobb megoldások, aktivitásra ösztönzi az egyetlen gépi kimenetet sokszor (mint a számológép-eredményt) passzívan elfogadó diákokat. A fordítóképzésben a kimenetek közötti választás fejleszti a döntéshozatali képességet, amire napjaink memóriatalálatokban és gépi javaslatokban bővelkedő fordítói gyakorlatában különösen nagy szükség van. A hatékony utószerkesztésnek is előfeltétele, hogy a fordító minél gyorsabban ki tudja választani a lehető leggyorsabban javítható találatot. Ennek gyakorlásához olyan alkalmazások (akár ingyenes próbaverzióinak) használatát javasoljuk, mint a GT4T, amely számos motor (és már a GPT-4) kimenetét is képes egyetlen gombnyomásra, egyetlen képernyőn megjeleníteni (lásd a 4. ábrát).

4. ábra

Példa a GT4T által kínált – tetszőleges számú és eredetű – MT-alternatívákra

Az ISO-18578:2017-szabvány-által-szabályozott, ám-a-becslések-nél-valójában-kiseb-
 arányban-jelen-lévő-„tisztán”-MTPE-munkafolyamatban-már-nem-is-jelennek-meg-a-
 hagyományos-értelemben-vett-fordítási-folyamatnak-sem-a-szereplői-(fordító,-lektor),-sem-pedig-
 olyan-meghatározó-dokumentumai,-mint-a-(lektorálatlan,-illetve-lektorált)-fordítás.-A-
 forrásnyelvi-szöveg-mellett-itt-egy-nyers-gépileg-fordított-fájl-kap-központi-szerepet,-melynek-
 részleges,-illetve-teljes-utószerkesztése-eredményeképpen-születik-meg-a-célnyelvi-szöveg.-
 Ennek-minőségellenőrzéséről-a-jelenleg-felülvizsgálat-alatt-álló-szabvány-eredeti-kiadása-
 nem-rendelkezik.

2 → ábra

Az ISO 18578:

A nyelvi-közvetít-
 gyakorlatban-a-gépi-
 által-meghatározó-
 segéd-eszközök-eg-

1 In the 'pure' MTPE workflow, which is regulated by the ISO 18578:2017 standard, but is actually present to a lesser extent than the estimates, neither the actors (translator, proofreader) nor the decisive documents of the traditional translation process, such as (unrefined or proofread) translation, are shown. In addition to the source language text, a raw machine translated file plays a central role here, and as a result of partial or complete post-editing, the target language text is created. This quality control is not provided for in the original edition of the standard, which is currently under review.

2 In the 'pure' MTPE workflow, which is governed by ISO 18578:2017, but which is actually less present in the estimates, neither the actors (translator, proofreader) nor the key documents such as the (unproofread or proofread) translation are present in the traditional translation process. In addition to the source language text, the focus is on a raw machine-translated file, which is partially or fully post-edited to produce the target language text. The original version of the standard, which is currently under revision, does not provide for quality control.

3 In the 'pure' MTPE workflow, regulated by the ISO 18578:2017 standard, but actually present in a smaller proportion than the estimates, neither the actors of the translation process in the traditional sense (translator, proofreader) nor such defining documents as the (unreviewed or reviewed) translation. In addition to the source language text, a raw machine-translated file plays a central role here, and the target language text is created as a result of its partial or complete post editing. The original edition of the standard, which is currently under review, does not provide for quality control.

4 The 'pure' MTPE workflow, regulated by ISO 18578:2017, but present in a smaller proportion than the estimates, no longer contains either the actors (translators, revisers) or the key documents of the traditional translation process, such as (unrevised or proofread) translation. In addition to the source language text, a raw machine-translated file plays a central role, resulting in partial or complete post-editing of the target language text. This quality control is not foreseen in the original edition of the standard, which is currently under revision.

A negatívumok és pozitívumok tudatosítására szolgáló gyakorlatokkal nem csak egy reális MT-kép kialakításához, hanem az *MT literacy* harmadik kérdéséhez is eljuthatunk:

Mikor és hogyan lehet, illetve kell módosítani a gépi fordítás kimenetét?

Ez a kérdés már túlmutat az MT-használat keretein és az utószerkesztés (PE) területére vezet minket (erről lásd Szlávik, 2023). A hibakeresési gyakorlatok minden szinten tudatosítják a diákokban a gépi fordítás „tökéletlenségét”, melynek kapcsán felhívhatjuk figyelmüket arra, hogy a kimenetek bizonyos helyzetekben módosításra szorulhatnak. Már említettük, hogy a nyelvtanulóktól nem várható el az L2-kimenet érdemi módosítása, L1 irányban pedig általában szövegértési célból alkalmaznak gépi fordítást, amelyet nem szükséges javítaniuk. A gépi fordítást jellemzően L2–L1 irányban használó fordítóhallgatókkal is csak az alapvető fordítástechnikai és célnyelvi ismeretek elsajátítása után érdemes konkrét utószerkesztési műveleteket végeztetni, de a gépi fordítást addig sem célszerű tabuként kezelni: a képzés első (fél)éve remek alkalmat teremthet az MT-eszközök használatával kapcsolatos instrumentális készségek fejlesztésére és egy egészségesen kritikus, reális MT-kép kialakítására.

Konklúzió

Az MT-használat oktatásba integrálásának kérdéseit a fordítóképzések, az idegennyelv-
 oktatás és a tanárképzés szintjén is alaposan át kell gondolni. Szem előtt kell tartani azt a

tényt, hogy a diákok a tiltások ellenére is használják a számos tekintetben kockázatokat rejtő gépi fordítóeszközöket, ugyanakkor azt is, hogy az MT-kimenetek mai színvonala akár megfelelő inputként is szolgálhat nyelvtanulóknál pl. a szókincs, mondatalkotási stb. készségek fejlődéséhez, a fordítóképzésben pedig az alternatív fordítási lehetőségek mérlegeléséhez, s ezáltal akár a fordítási kompetencia fejlesztéséhez is. Ilyen körülmények között tiltás (vagy tabusítás) és túrés (vagy tétlenkedés) helyett a nyelvtanításban és a fordítóképzések első szakaszában is érdemesnek tűnik „tudatosan tudatosítani” diákjainkban a gépi fordítás korlátait és lehetőségeit, egyetlen eszköz helytelen módon (pl. szótárként) történő, ellenőrizetlen használata helyett inkább többféle eszköz párhuzamos, kontrollált használatának szorgalmazásával. A fordítási munkafolyamatok bemutatásakor láthatuk, hogy a többféle forrásból (TM-ek, MT stb.) származó fordítási lehetőségek közötti választás fontos része a mai fordítók munkájának, ezért fordítóképzéseinkben a tudatos MT-használatot elősegítő gyakorlatok a reális MT-kép kialakítása mellett a fordítástól ma már elválaszthatatlan utószerkesztési kompetencia fejlesztéséhez is hozzá tudnak járulni.

IRODALOM

- Bowker, L. & Buitrago-Ciro, J. (2019). *Machine Translation and global research: Towards improved machine translation literacy in the scholarly community*. Emerald Publishing. <https://doi.org/10.1108/9781787567214>
- Bowker, L. (2021). Promoting Linguistic Diversity and Inclusion: Incorporating Machine Translation Literacy into Information Literacy Instruction for Undergraduate Students. *The International Journal of Information, Diversity, & Inclusion*, 5(3), 127–151. <https://doi.org/10.33137/ijidi.v5i3.36159>
- Bin Dahmash, N. (2020). 'I Can't Live Without Google Translate': A Close Look at the Use of Google Translate App by Second Language Learners in Saudi Arabia. *Arab World English Journal (AWEJ)*, 11(3). <https://doi.org/10.24093/awej/vol11no3.14>
- Dorst, A. G., Valdez, S. & Bouman, H. (2022). Machine translation in the multilingual classroom: How, when and why do humanities students at a Dutch university use machine translation? *Translation and Translanguaging in Multilingual Contexts*, 8(1), 49–66. <https://doi.org/10.1075/ttmc.00080.dor>
- Ducar, C. & Houk Shocket, D. (2018). Machine translation and the L2 classroom: Pedagogical solutions for making peace with Google Translate. *Foreign Language Annals*, 51, 779–795. <https://doi.org/10.1111/flan.12366>
- ELIS (2021, 2022, 2023). <https://elis-survey.org/>
- EMT (2022). https://commission.europa.eu/system/files/2022-11/emt_competence_fw_k_2022_en.pdf
- Fáy, T. & Yang, Z. Gy. (2023). A gépi fordítás szerepe a fordítóképzésben. *Modern Nyelvoktatás*, 29(1–2), 24–34. <https://doi.org/10.51139/monye.2023.1-2.24.34>
- Gene, V. & Guerrero, L. (2023). *A Common Machine Translation Post-Editing Protocol for Academia, Clients, LSPs and Post-Editors - Findings, Resources, and Conclusions from GALA MTPE Training SIG*, <https://gala-global.notion.site/gala-global/A-Common-Machine-Translation-Post-Editing-Protocol-for-Academia-Clients-LSPs-and-Post-Editors-2a137e7abdc6455b8dcc79ef608f1b65>
- Hunnect Kft. (2021). *A gépi fordítás helyzete a magyar szakfordítók körében*. <https://hunnect.com/hu/gepi-forditas-helyzete-magyar-szakforditok-koreben/>
- Kasperé, R., Horbačauskienė, J., Motiejūnienė, J., Liubinienė, V., Patašienė, I. & Patašius M. (2021). Towards sustainable use of machine translation: Usability and perceived quality from the end-user perspective. *Sustainability*, 13(23), 13–30. <https://doi.org/10.3390/su132313430>

- Kóbor, M. (2023). La traduction automatique dans nos formations : sensibilisation aux limites. In Németh, L., Nyúl, E. A. & Salamon, E. (szerk.), *Nemzetközi újlatin tudományos konferencia 2023. Libretto abstract / Livre des résumés / Libro de resúmenes*, Pécsi Tudományegyetem BTK.
- Lee, S-M. (2020). The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning*, 33(3), 157–175. <https://doi.org/10.1080/09588221.2018.1553186>
- Lesznyák, Á. (2019). Hungarian translators' perceptions of Neural Machine Translation in the European Commission. In Forcada, M., Way, A., Tinsley, J., Shterionov, D., Rico, C. & Gaspar, F. (eds.), *Proceedings of Machine Translation Summit XVII: Translator, Project and User Tracks* (pp. 16–22), European Association for Machine Translation.
- Lesznyák, M., Bakti, M., Sermann, E. (2022). Humán fordítás és utószerkesztés hallgatói szemmel. In Navracsics, J. & Bányi, Sz. (eds.), *Nyelvek, nyelvvaltozatok, következmények II.: Fordítástudomány, terminológia, retorika, kognitív nyelvészet, kontrasztív nyelvészet, interkulturális kommunikáció, névtan* (pp. 86–102). Akadémiai Kiadó.
- Liubinienė, V., Lisaitė, D., Motiejūnienė, J. (2022). A Snapshot of Children's Attitudes toward Machine Translation. *MDPI*. 13(7), 1–14. <https://doi.org/10.3390/info13070317>
- Liu, K., Kwok, H. L., Liu, J. & Cheung, A.K.F. (2022). Sustainability and Influence of Machine Translation: Perceptions and Attitudes of Translation Instructors and Learners in Hong Kong. *Sustainability*, 14, 63–99. <https://doi.org/10.3390/su14116399>
- Loock, R. & Léchauguette, S. (2021). Machine translation literacy and undergraduate students in applied languages: report on an exploratory study. *Revista Tradumàtica. Technologies de la Traducció*, 19, 204–225. <https://doi.org/10.5565/rev/tradumatica.281>
- Loock, R. (2020). No more rage against the machine: How the corpus-based identification of machine-translationese can lead to student empowerment. *The Journal of Specialised Translation*, 34, 150–170.
- Niño, A. (2020). Exploring the use of online machine translation for independent language learning. *Research in Learning Technology*, 28. <https://doi.org/10.25304/rlt.v28.2402>
- Nitzke, J., Hansen-Schirra, S., Canfora, C. (2019). Risk management and post-editing competence. *The Journal of Specialised Translation*, 31.
- O'Brien, S., Ehrensberger-Dow, M. (2020). MT Literacy: A cognitive view. *Translation, Cognition & Behavior*, 3(2), 145–164. <https://doi.org/10.1075/tcb.00038.obr>
- Prószék, G. (2021). A gépi fordítás hetvenéves története. In *A korpusznyelvészettől a neurális hálókig: Köszöntő kötet Váradi Tamás 70. születésnapjára*, pp. 141–156. MTA Nyelvtudományi Kutatóközpont. <https://doi.org/10.18135/VT70.17>
- Resende, N., Way, A. (2021). Can Google Translate rewire your L2 English processing? *Digital*, 1(1), 66–85. <https://doi.org/10.3390/digital1010006>
- Robin, E. & Eszenyi, R. (2022). BME TFK Őszi Konferencia Translating Europa Workshop. *Fordítástudomány*, 24(2), 184–191. <https://doi.org/10.35924/fordtud.24.2.10>
- Robin, E., Eszenyi, R., Kóbor, M. & Seidl-Péché, O. (2023). *Human in the Translation Loop: az ELTE FTT és a DGT kutatási projektje*. Elhangzott: TransELTE 2023.
- Rossetti, A., O'Brien, S. & Cadwell, P. (2020). Comprehension and trust in crises: Investigating the impact of machine translation and post-editing. In *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation* (pp. 9–18). European Association for Machine Translation.
- Seidl-Péché, O. (2018). Milyen alapelvek mentén oktassuk a terminológiai ismereteket a szakfordítóképzésben? *PORTA LINGUA*, 2018, 153–163.
- Seresi, M., Koponen, M., Pareja, A., Bowker, L., Delorme Benites, A., Şahin, M. et al. (2022). *LITHME WG7's Survey on Machine Translation Use Habits and Practices*. Figshare. Dataset. <https://doi.org/10.6084/m9.figshare.21165775.v1>

- Szlávik, Sz. (2022a). A gépi fordításhoz kötődő alapvető terminusok, definíciók és a közöttük lévő ellentmondások. *Fordítástudomány*, 24(1), 87–103. <https://doi.org/10.35924/fordtud.24.1.6>
- Szlávik, Sz. (2022b). A neurális gépi fordítómotorok típusának és fejlődésének jelentősége az utószerkesztés kutatásában. Elhangzott: *TransELTE 2022*.
- Szlávik, Sz. (2023). Az utószerkesztési kompetencia (PE) fejlesztésének lehetőségei. *Modern Nyelvoktatás*, 29(1–2), 35–48. <https://doi.org/10.51139/monye.2023.1-2.35.48>
- Teixeira, C. (2020). Revising computer-mediated translations. In Mossop, B., Hing, J. & Teixeira, C. (Eds.), *Revising and Editing for Translators* (pp. 207–224). Routledge. <https://doi.org/10.4324/9781315158990-16>
- Veresné Valentinyi, K. (2022). *A gépi fordítás kihívásai a szakfordítóképzésben*. *PORTA LINGUA*, 2022(1), 69–80. <https://doi.org/10.48040/PL.2022.1.7>
- Yamada, M. (2019). The impact of Google Neural Machine Translation on post-editing by student translators. *The Journal of Specialised Translation*, 31. 87–106. https://jostrans.org/issue31/art_yamada.php
- Yamada, M. (2020). Language learners and non-professional translators as users. In O’Hagan, M. (Ed.), *The Routledge Handbook of Translation Technology* (pp. 183–199). Routledge. <https://doi.org/10.4324/9781315311258>
- Yang, Z. Gy. (2018). A gépi fordítás és a neurális gépi fordítás. *Modern Nyelvoktatás*. 24(2–3), 129–139. <https://ojs.elte.hu/modernnyelvok/article/view/1625/1380>

Machine Translation (not only) in Translation Training. Methodological contribution to the development of MT education

State-of-the-art neural machine translation (NMT) tools can now produce more and more convincing outputs even for "low resource" languages like Hungarian but machine translated texts still require post-editing (PE). As MT/PE is integrated into professional workflows, translator training programs must adapt to the constantly evolving new reality. Despite fears that premature use of MT impedes the development of translation competence, training programs should not ignore the fact that students now use machine translation tools at all levels of language learning, often well before translation training. Therefore, instead of banning MT-use, trainers should strive to raise awareness of the limitations of machine translation and to develop students' MT literacy – not only in translation training but also in L2 classes. Our data suggests that supervised parallel use of two or more engines can help to develop the critical attitude needed to acquire the appropriate competences.

Keywords: machine translation, NMT, MT literacy, awareness raising, translator training