

FÁY TAMÁS – YANG ZIJIAN GYŐZŐ

A gépi fordítás szerepe a fordítóképzésben

A gépi fordítás folyamatosan javuló minősége miatt a fordítóképző intézmények számára egyre sürgetőbb a kérdés, hogy hogyan viszonyuljanak a fordítómotorok oktatási célú használatához. Abban egyetértés van, hogy a témának helyet kell biztosítani a képzésen belül, viszont azzal kapcsolatban már megoszlanak a vélemények, hogy ez mikor és milyen formában történjen. Egyes vélemények szerint a gépi fordítás alkalmazását meg kell előznie a fordítási készségek kialakulásának, mert ellenkező esetben megfelelő minőségű utószerkesztést sem lehet végezni. Túl korai alkalmazásuk esetén viszont félt, hogy nem jut elegendő idő a fordítási készségek megszilárdulására, ugyanakkor – mivel szabadon hozzáférhető eszközökről van szó – szinte lehetetlen megakadályozni a képzés korai szakaszában történő használatukat. A tanulmány a gépi fordítás jellemzőinek áttekintése után ezen dilemma megoldásához kíván támpontokat nyújtani.

Kulcsszavak: gépi fordítás, fordítástechnológia, fordításoktatás, utószerkesztés, MTPE

Bevezetés

Fordításelmélet és számítógépes nyelvészet – Ha valaki napjainkban áttekintést szeretne adni a fordításelmélet interdiszciplináris kapcsolatairól, az egyik fejezetet már bizonyára ennek a témának kellene szentelnie. A gépi fordításon kívül talán nincs is más olyan területe a fordításnak, ahol ennyire nyilvánvaló lenne a két diszciplína kapcsolata. De vajon miért csak napjainkban kap kiemelt figyelmet ez a téma, amikor a gépi fordítás több mint hetvenéves múltra tekint vissza, és az 1960-as években már Magyarországon is folytak komoly kutatások a gépi fordítással kapcsolatban? (vö. Hell, 1962) A válasz nyilvánvalóan az, hogy a gépi fordítás napjainkban – köszönhetően a neurális hálózat-alapú gépi tanulás térhódításának – már olyan minőségű outputot képes létrehozni, amivel a humán fordítás méltó vetélytársává lépett elő (vö. Laki & Yang, 2022, p. 358). Ezzel párhuzamosan természetesen a fordítást oktató intézményeknek is szembe kell nézniük a kérdéssel, hogy hogyan viszonyuljanak a gépi fordításhoz. Ma már nem kérdés, hogy a gépi fordításnak helyet kell biztosítani a fordítóképzésben, a kérdés inkább az, hogy milyen formában és mélységben érdemes vele foglalkozni, és egyáltalán mit is oktassunk „gépi fordítás” címszó alatt. Jelen tanulmány célja, hogy a fordításelmélet és a számítógépes nyelvészet kapcsolatához illeszkedve támpontokat nyújtson ezen kérdések megválaszolásához.

A gépi fordítás terminológiai és fordításelméleti vonatkozásai

A gépi fordítás fogalma látszólag viszonylag könnyen definiálható (Kenny, 2020, p. 305): „[...] the automatic conversion of text from one natural language to another”¹,

¹ [...] egy szöveg automatikus átültetése egy természetes nyelvből egy másikba. [saját fordítás]

tehát a fordítás „automatikus” vagy „gépi” úton, azaz emberi beavatkozás nélkül jön létre, bár abban egyetértés van, hogy az emberi beavatkozást nem lehet teljesen kizárni a folyamatból, vö. Hutchins és Somers (1992, p. 3): „[...] computerised systems responsible for the production of translations from one natural language into another, with or without human assistance.”² Ezzel kapcsolatban lehet arra hivatkozni, hogy a gépek és a fordítást előállító algoritmusok is emberi tevékenység eredményeként jönnek létre, ugyanakkor emberi beavatkozáson többnyire az elő- vagy utószerkesztést értik (pl. Bowker & Fisher, 2010, p. 60). A szöveg létrejötte szempontjából viszont fontosabb, hogy még a neurális hálózat alapú gépi fordításnak is nagy mennyiségű olyan (egynyelvű vagy kétnyelvű) korpuszra van szüksége (vö. Laki, 2018, p. 172), amelyeket egyelőre még emberek hoznak létre (bár a mesterséges intelligencia fejlődése következtében ez hamarosan megváltozhat). A gép tehát olyan eredményt fog produkálni, mint amelyet kívülről beletápláltak. Függetlenül azonban attól, hogy mekkora mértékű és milyen jellegű az emberi beavatkozás, megállapítható, hogy a gépi fordítás és a humán fordítás között nem lehet éles határvonalat húzni. Különösen igaz ez, ha nem a célnyelvi szöveg keletkezése felől vizsgáljuk a dolgot, hanem a kész produktum felől, hiszen a gépi fordítás fogalma implicit módon a szöveg keletkezésének körülményeit sugallja, a felhasználó viszont legtöbbször már a kész szöveggel találkozik, amely nem egyszer a gépi fordítás, a fordítómemóriából és a terminológiai adatbázisból érkező találatok és a humán fordítás keveredésének visszafejthetetlen eredményeként jön létre. Nem véletlen, hogy Pym a forrásnyelvi szöveg kifejezés helyett már 2013-ban is a „start text” elnevezést tartotta időszerűnek (Pym, 2013, p. 492), amely jól kifejezi, hogy a fordítás alapjául szolgáló input már nem csak a forrásnyelvi szöveget foglalja magában.

A gépi és humán fordítás közötti éles határvonal problematikáját elkerülendő Hutchins és Somers (1992, p. 148) egy spektrum két végére helyezi a kizárólag géppel végzett fordítást és a hagyományos humán fordítást, és a kettő között pozicionálják a számítógéppel segített humán fordítást, illetve a humán segítséggel végzett gépi fordítást. Ez a megközelítés kétségtelenül közelebb jár az igazsághoz, azonban nem ad megnyugtató választ arra a kérdésre (bár nem is feladata választ adni rá), hogy a fordításelmélet és a fordításoktatás hogyan viszonyuljon vizsgálatának egyik tárgyához, a célnyelvi szöveghez. Ha még érvényesnek tekintjük Holmes eredetileg 1972-ben megjelent felosztását, aki a leíró fordítástudományon belül megkülönböztetett eredmény-, funkció- és folyamatorientált ágat (Holmes, 2000, p. 176–177), akkor az eltelt több mint 50 év, illetve a humán és gépi fordítás összefonódása fényében feltehetjük a kérdést, hogy a fordításelmélet vajon tehet-e még érvényes kijelentéseket a fordítás eredményére, folyamatára és funkciójára vonatkozóan, ha nem ismertek a szöveg létrejöttének pontos körülményei, illetve, hogy a szövegben megfigyelhető nyelvi törvényszerűségekből a fordításra mint alkotó emberi tevékenységre vagy inkább a gépek működési mechanizmusaira következtethetünk? Elképzelhető, hogy a gépi fordítóprogramok széleskörű fordítói felhasználása arra fogja ösztönözni a kutatókat, hogy a figyelmüket még inkább a műfordításra összpontosítsák, amely makacsabban

² [...] olyan számítógépes rendszerek, amelyek emberi segítséggel vagy anélkül fordítanak le egy szöveget egy természetes nyelvről egy másikra. [saját fordítás]

ellenáll a technológiai fejlődésnek, és így – ha a fordítást emberi tevékenységként kívánjuk vizsgálni – pontosabb képet adhat a fordítás folyamatáról, eredményéről.

A gépi fordítás jellemzői

Az alábbiakban elsősorban a gépi fordítás azon jellemzőit szeretnénk bemutatni, amelyek kiemelt jelentőséggel bírnak a fordítóképzés szempontjából. Az elmúlt évtizedekben a gépi fordítás módszere jelentős fejlődésen ment keresztül. A természetes nyelvek közötti fordítás megoldására többféle modell is született, a szabályokkal történő fordításoktól a statisztikai algoritmusokon keresztül a neurális módszerekig. Ezzel a fejlődéssel párhuzamosan a gépi fordítás jellemzői is folyamatosan változtak, azonban megfogalmazható néhány általános jellemző. A fordítómotorok rohamos fejlődése miatt azonban igyekeztünk minél kevesebb példával szemléltetni a leírtakat, hiszen az algoritmusok változása miatt előfordul, hogy hetek, sőt napok elteltével is más célnyelvi változatot kapunk.

A fordítás sebessége

A szakfordítók által egy nap alatt lefordított szavak átlagos száma nyilván sok tényezőtől függ, például a tapasztalatuktól, a gépelési sebességüktől, a fordítómemóriából feljövő találatok minőségétől, de az általános iparági irányadó értéket jellemzően napi kb. 2000 szóban szokták meghatározni, a lektorált szavak száma pedig ennek mintegy a felére tehető. Gépi fordítás alkalmazásával ez az érték akár megháromszorozható³. Ez a szempont ugyan a fordítóképzés során kevésbé érvényesül, hiszen egy-egy fordítás elkészítésére, kiértékelésére jóval több idő jut, de a határidők és egyéb teendők által okozott nyomás miatt a gépi fordítás előbb-utóbb a hallgatók számára is vonzó alternatívát jelent – mindenféle (indokolt vagy indokolatlan) tiltás (ezzel kapcsolatban lásd az 5. fejezetet is) és észérvekre való hivatkozás ellenére.

A fordítás minősége

A gépi fordítás megítélését illetően még mindig a minőségével kapcsolatban oszlanak meg leginkább a vélemények. A fordítóprogramok alkalmasságát és alkalmazhatóságát gyakran a szövegtípus felől szokták megközelíteni, Heltai (2014, p. 17) például a gépet nem tartja alkalmasnak versek, drámák, viccek, szójátékok, humoros írások, veretes publicisztikai szövegek, hétköznapi társalgások, tehát olyan szövegek fordítására, amelyekre jellemzően az „expresszív” vagy „operatív” címkét szoktuk ráaggatni. Lakatos-Báldy (2022, p. 66) alapvetően egyetért Heltai megállapításával, de a konklúziója, miszerint „a gépi fordítás jelenléte ma már a filmfordításban is vitathatatlanul jelen van és [sic!] helyet követel magának” egyértelműen azt mutatja, hogy a gépi fordítás használói nem feltétlenül a szövegtípus alapján mérlegelik annak alkalmazhatóságát, és még a nyilvánvaló hátrányok ellenére is használnak ilyen eszközöket. A kérdés inkább abba az irányba mozdul el, hogy milyen célra szeretnénk használni a gépi fordítást. A neurális módszerek, főleg a transzformer alapú modellek használata

³ Lásd például <https://www.oneword.de/en/whats-the-difference-between-human-translation-and-mtpe/> (megtekintve 2023. február)

óta a gép egyre finomabb összefüggéseket is képes megtanulni az elemzett szövegekben, és a doménspecifikus adatokra való hangolás után komplexebb fordításokra is képes egy adott témában. Véleményünk szerint tehát a szövegtípus felőli megközelítés egyre kevésbé fog érvényesülni a jövőben, és a gépi fordítás hasznosságát szövegtípustól függetlenül már most is inkább az alábbi képlet alapján lehet megítélni: A gépi fordítás használata célravezető, ha $MT + PE < HT$, azaz, ha a gépi fordítás (MT) az utószerkesztéssel együtt (PE) rövidebb ideig tart, mint a humán fordítás (HT), és nem célravezető a használata, ha $MT + PE \geq HT$. Nyilván a gyakorlatban kivitelezhetetlen annak mérése, hogy ugyanazt a szöveget ugyanaz a személy mennyi idő alatt fordította volna le gépi segítség nélkül és gépi segítséggel, de a gépi fordítás fejlődése egyértelműen az első képlet irányába hat, így a szövegtípus alapján történő mérlegelés valószínűleg egyre kevésbé lesz releváns. Azt is figyelembe kell venni, hogy „a fordítóiparban ma már nem csak a kifogástalan minőségnek van piaca” (Mohácsi-Gorove, 2017, p. 178), tehát adódnak olyan helyzetek, amikor egy géppel előfordított szöveg tökéletesen megfelel az adott célnak, például amikor egy szöveg főbb tartalmi elemeit szeretnénk megérteni.

Adatvédelmi szempontok

Már a képzés során fel kell hívni a figyelmet arra, hogy a gépi fordítás használatakor harmadik fél rendelkezésére bocsátjuk a szöveget, aki azt adott esetben meghatározott ideig fogja kezelni a rendszer, illetve az algoritmusok tanítása és fejlesztése céljából. A gépi segítséget nemcsak közvetlenül, az egyes szolgáltatók platformján keresztül, hanem a számítógépes fordítástámogató eszközök plugin modulján keresztül is igénybe lehet venni előfordítás formájában vagy fordítás közben, de az adatvédelmi aggályokat ezzel a megoldással sem lehet kiküszöbölni. Bruckner (2017, p. 13) a fő veszélyt abban látja, hogy a „big data” elemzési módszerek segítségével az adatgyűjtő szolgáltatók nemcsak a felhasználók keresési viselkedésének mintázatát tudják feltérképezni, hanem a feltöltött szegmensdarabokat a metaadatok alapján összeillesztve elméletileg akár egy vállalati profilt is el tudnak készíteni. Használat előtt ezért mindenképpen érdemes alaposan áttanulmányozni az adott szolgáltató használati feltételeit. Az API-kulcs használatával az adatok (szövegek) ugyan titkosított formában továbbítódnak a szolgáltató szerverére, és azokat a szolgáltató nem tárolja (huzamosabb ideig) és nem használja fel a szolgáltatás fejlesztése céljából⁴, de tapasztalataink szerint az API segítségével történő gépi fordítás nincs (sincs) még egyértelműen szabályozva adatvédelmi szempontból.

Nyelvi jellemzők

A gépi fordítás nyelvi jellemzőiről és korlátairól átfogó áttekintést ad például Veresné Valentinyi (2022, p. 77), így ezzel itt nem foglalkoznánk részletesen. A listát annyiban egészítenénk ki, hogy a gépi fordítás a kontextus hiánya, illetve a szegmensközpon-tú megközelítés következtében sok esetben a rövid (egy-, kétszavas) mondatokkal is

⁴ Vö. <https://www.microsoft.com/en-us/translator/blog/2018/03/09/api-customer-text-translations-no-longer-logged-by-microsoft-translator/> (megtekintés: 2023. február)

nehezen tud megbirkózni, illetve hosszú mondatok esetében az írásjelek, különösen a pontok mondaton belüli halmozása (pl. törvények esetében) „összezavarja” a fordítómotorokat, valószínűleg azért, mert a mondathatárok felismerését még nem lehet precízen algoritmizálni (ugyanaz a probléma jelentkezik a szövegek szegmentálásakor CAT-eszközbe történő importálás során). A terminusok következtlen fordítása szintén a gépi fordítás szegmensközpontúságából következik, de ezt egy megfelelően összeállított terminológiai adatbázissal viszonylag könnyű ellensúlyozni. Német-magyar irányban gyakori hibaforrás, hogy a németben formailag egyértelműen jelölt tegezés és magázás jelenleg teljesen véletlenszerűen alakul a magyarban, vélhetően azért, mert a fordítás valószínűleg az angol közbeiktatásával történik⁵, például:

Im schlimmsten Fall könnt Ihr mit Eurer Arbeit nicht weitermachen.

A legrosszabb esetben nem tudja folytatni a munkáját. (DeepL és Google Fordító)⁶

Ha elég sokáig próbálkozunk, találunk olyan felvezető mondatot, amellyel tegezésre tudjuk kényszeríteni a korábban még magázódó gépi fordítást:

Dann vergeht Euch das Lachen und Ihr könnt im schlimmsten Fall mit Eurer Arbeit nicht weiter machen.

Akkor abba fogod hagyni a nevetést, és legrosszabb esetben nem fogod tudni folytatni a munkát. (DeepL)

Aztán abbahagyod a nevetést, és a legrosszabb esetben nem tudod folytatni a munkát. (Google Fordító)

Ha viszont az első személyes névmás nagy kezdőbetűjét átírjuk kicsire, teljesen összezavarodnak a vizsgált fordítómotorok, amelyeknek még az előzőleg nagyjából elfogadható módon lefordított részbe is beletörik a bicskájuk:

Dann vergeht euch das Lachen und Ihr könnt im schlimmsten Fall mit Eurer Arbeit nicht weiter machen.

⁵ Ami azért furcsa, mert a neurális modelleknél minőségbeli problémák miatt egyre kevesebbszer használják a közvetítőnyelves megoldást (például a Google többnyelvű modellt használ), ugyanakkor, mivel a DeepL nem publikálja a rendszer működését, pontosan nem lehet tudni, hogy hogyan működik a háttérben, így az is elképzelhető, hogy a tanítóanyag zaja okozza a problémát. A neurális hálók (főleg az új környezetfüggő modellek) egyébként is sok mintát tanulnak, ezért nagyon érzékenyek a kis és nagy kezdőbetűre, egy-egy írásjelre, így előfordul, hogy kettő, ránézésre szinte ugyanolyan szöveget a modellek teljesen eltérő szöveggé értelmeznek. Ugyanakkor más példák egyértelműbben utalnak az angol mint közvetítőnyelv jelenlétére, például: *Die Montagefreundlichkeit der Beschlagelemente erleichtert die Arbeit --> A hardverelemek könnyű összeszerelhetősége megkönnyíti a munkát* (helyesen: *vasalatelemek*). Vö. a német mondat angol fordítását: *The ease of assembly of the hardware elements facilitates the work* (DeepL 2023. június 8.). A *hardverelemek* kifejezés csak úgy értelmezhető, ha feltételezzük, hogy a DeepL a német mondatot először angolra fordította, majd az angol fordítást átültette magyarra.

⁶ A lekérdezés dátuma minden esetben: 2023. február 11.

Akkor elveszíted a nevetésedet, és a legrosszabb esetben nem tudod folytatni a munkádat. (DeepL)

Ekkor elvész a nevetés, és a legrosszabb esetben nem tudja folytatni a munkáját. (Google Fordító)

A gépi fordításhoz és a humán fordításhoz szükséges kompetenciák összehasonlításakor egyértelmű emberi fölény rajzolódik ki (lásd Veresné Valentinyi, 2022, p. 70–71; Heltai, 2014 alapján), és a gépi fordítás jellemzőinek felsorolásakor hajlamosak vagyunk annak hiányosságaira fókuszálni, mintegy sugallva a humán fordítók tévedhetetlenségét, ugyanakkor bizonyára érdekes és meglepő tanulságokkal szolgálna a géppel fordított és utószerkesztett szövegek szisztematikus egybevetése a gépi segítség nélkül készült humán fordításokkal – az idő-minőség arányt is figyelembe véve.

A gépi fordítás használata a fordítóipari piacon

Az, hogy a fordításelméleti diskurzusban egy ideje ilyen mértékben van jelen a gépi fordítás, nyilván annak tudható be, hogy sokan használják, sőt kijelenthető, hogy a szabadúszó fordítók nagy többsége esetében a gépi fordítás használata már a mindennapi rutin része (vö. Cadwell et al., 2018 vagy „A gépi fordítás helyzete a magyar szakfordítók körében”⁷), még ha a témáról nehéz is reprezentatív felmérést készíteni – részben az alábbiakban ismertetett okok miatt. Kétségtelenül van ugyanis egy olyan látens felhasználói csoport, amelynek a tagjai előszeretettel használják ugyan, de az adatvédelmi aggályok és a szabályozatlanság miatt tiltott gyümölcsnek tekintik a gépi fordítást, amelybe jólesik beleharapni, de ezt nem szívesen vallják be a megrendelők által – a szöveg bizalmassági szintjétől függetlenül – elrendelt általános tilalom miatt. Vannak olyan szakfordítók, akik eleve bizalmatlanul tekintenek mindenféle technológiai fejlődésre, és még a számítógépes fordítástámogató eszközöket is elutasítják (azok pontos funkciójának ismerete nélkül), mások a munkájukat veszélyeztető ellenséget látják a gépben⁸, és bizonyára vannak olyanok, akiket a gépi fordítás kezdeti verzióival szerzett negatív tapasztalatok riasztottak el a technológia használatától, és arra az általuk helyesnek vélt következtetésre jutottak, hogy „egyszerűbb újrafordítanom a szöveget” – az előző fejezetben leírt képlet alapján azonban egyáltalán nem biztos, hogy az ilyen kijelentés minden esetben megállja a helyét. A gépi fordítórendszerek iránti (különbféle online fórumokon megnyilvánuló) ellenszenvet az információhiány és a gépi fordító rendszerek nem megfelelő használata is táplálhatja (sokan nem arra használják, amire való). Ha a szakfordítók jobban megértenék a gépi fordítómotorok működését, és figyelembe vennék azok előnyeit és hátrányait, akkor bizonyára elfogadóbb magatartást tanúsítanának, és hatékonyabban tudnák integrálni ezeket az eszközöket a munkafolyamatba.

⁷ <https://hunnect.com/hu/gepi-forditas-helyzete-magyar-szakforditok-koreben> (megtekintés: 2023. február)

⁸ A gépekkel szembeni ellenérzés nem újkeletű dolog, ennek kapcsán könnyen eszünkbe juthat a 19. század elején tomboló géppromboló mozgalom, amikor a ludditák a szövőgépek megrongálásával adtak nyomatékot abbéli félelmüknek, hogy a gépek miatt elveszítik a munkájukat.

Azt gondolhatnánk, hogy a szakfordítók paradox helyzetbe kerültek: Ha használják a gépi fordítást, maguk is hozzájárulnak az algoritmusok fejlődéséhez, aminek következtében egyre kevesebb munkához jutnak. A világpiacon a trendek viszont arról tanúskodnak, hogy a fordítóipari piacra folyamatos növekedés jellemző, amit még a koronavírus-járvány sem tudott megfékezni – sok más iparággal ellentétben. 2021-ben a fordítóipar globális forgalma két év alatt ötmilliárd dollárral mintegy 56 milliárd dollárra nőtt, 2025-re pedig 73,6 milliárd dolláros forgalmat prognosztizálnak⁹. Az persze nem derül ki az adatokból, hogy ezek a változások miként érintik a szabadúszó fordítók anyagi helyzetét vagy számát.

Vállalatok és fordítóirodák esetében még fontosabb a 3. fejezetben tárgyalt képlet, hiszen az emberi fordítás nagyságrendekkel drágább. A gépi fordítómotorok használatával időt és ezzel pénzt lehet megspórolni, a minőségi fejlődésük pedig indokoltá tette a fordítási munkafolyamatba való integrálásukat. A jelentőségüket jól mutatja, hogy külön kutatási irányok és munkahelyek jöttek létre az utószerkesztéssel, az utószerkesztés mérésével, a gépi fordítás kiértékelésével és ezek automatizálásával kapcsolatban.

A gépi fordítás a fordítás oktatásában

Gyakran felmerül a kérdés, hogy a fordítóképzés hogyan viszonyuljon a gépi fordításhoz: Tiltsa, tűrje vagy támogassa? A három lehetséges megközelítés közül véleményünk szerint az első értelmezhetetlen ebben a kontextusban, hiszen szabadon hozzáférhető (legális) eszközökről van szó. A tiltás (megszegése) ráadásul mindig együtt jár a bizonyítás kényszerével, a gépi fordítás használatát viszont – már csak a folyamatosan változó algoritmusok miatt is – nehéz bizonyítani, és a folyamatos gyanúsítgatás bizonyára az oktatás bizalomra épülő légkörére sem lenne jó hatással. A tűrés azt sugallja, hogy ne vegyünk tudomást a gépi fordításról, ami az előző fejezetek tükrében minden racionális érvek ellentmond. A különféle szakmai fórumokon megosztott tapasztalatok alapján úgy tűnik, a képzőintézmények többsége a támogatás mellett tette le a voksát. Ezzel kapcsolatban az alábbi három kérdést érdemes tisztázni: A képzés melyik szakaszában (már az elején, vagy csak a fordítástechnikai készségek begyakorlása után), milyen formában (külön kurzusként vagy egy fordítási/lektorálási kurzus részeként, ezzel kapcsolatban lásd Mohácsi-Gorove, 2017, p. 186 is) és milyen mélységben (a gépi fordítás története, fajtái, működése stb.) érdemes foglalkozni a gépi fordítással (vö. Kóbor 2023 és Szlávik 2023 ebben a kötetben). Mindegyik szempont ellen és mellett is fel lehet hozni érveket, tehát nem könnyű konszenzusra jutni¹⁰: A tiltás és a támogatás sajátos hibrid modelljeként jó megoldás lehet például, ha a fordítástechnikai készségek kialakulása és erősítése érdekében kezdetben több időt szánunk a fordítások tanórai elkészítésére és kiértékelésére (oktatói irányítással, gépi segítség nélkül), és ezzel párhuzamosan vagy a képzés egy későbbi szakaszában

⁹ <https://www.mylanguageconnection.com/translation-industry-trends-and-statistics/> és <https://www.truelanguage.com/translation-services-are-growing-2022-translations-predictions/> (megtekintés: 2023. február)

¹⁰ Ezzel kapcsolatban a felsőoktatási intézmények a 2022. őszi Tanár–Tanár Találkozó eredményeként elfogadott ajánlások formájában törekednek egy egyéges álláspont kialakítására.

külön kurzus keretében foglalkozunk az utószerkesztéssel. Az is célravezető lehet, ha a képzés korai, készségfejlesztő szakaszában olyan szöveget adunk ki otthonra, amelynek a fordítására a 3. fejezetben ismertetett képlet alapján a gépi fordítómotorok kevésbé alkalmasak.

A géppel fordított szövegek esetében elő- és utószerkesztést szoktak megkülönböztetni: Az előbbi a szöveg előkészítését jelenti azzal a céllal, hogy kiküszöböljük a félrefordításra lehetőséget adó szövegrészeket, az utóbbi pedig a gép által lefordított szöveg tökéletesítésére irányul (minimális vagy teljes utószerkesztés formájában, bár ez a megkülönböztetés inkább utólag, a szerkesztés elvégzése után nyer létjogosultságot). A gépi fordítás oktatása során az utószerkesztés mint kompetencia jóval nagyobb hangsúlyt kap, pedig a szöveg előszerkesztési szempontok szerint történő előkészítése az utószerkesztést és a fordítómotorok által elkövetett hibák felismeréséhez szükséges készségek kialakulását is megkönnyíthetné¹¹. De vajon beszélhetünk-e külön utószerkesztési kompetenciáról? És ha igen, az mennyiben tér el a lektorálási kompetenciától? Vagy esetleg a két fogalom között csak a szöveg elkészültének körülményeit tekintve van különbség (humán vs. gépi fordító), és a különbségtételt még a neurális fordítómotorok előtti időszak ihlette? A kérdés egy külön tanulmányt is megérdemelne, de véleményünk szerint a látszat és az eltérő megnevezés ellenére a kettő között nincs akkora különbség (és az olló két szára egyre inkább zárul), hogy az utószerkesztést ne lehessen kényelmesen elhelyezni a lektoráláson belül: A lektori kompetencia összetevőit Horváth (2011, p. 44) alapján megvizsgálva semmi olyat nem találunk, ami ne lenne érvényes az utószerkesztésre is. Ha pedig megnézzük a de Almeida & O'Brien (2010¹²) által összeállított, utószerkesztéshez szükséges általános készségek listáját, akkor megállapíthatjuk, hogy a benne foglaltak változatlan formában a lektorok munkájára is érvényes:

„1 – The ability to identify issues [...] that need to be addressed and to fix them appropriately; 2 – The ability to carry out the post-editing task with reasonable speed [...]; 3 – The ability to adhere to the guidelines, so as to minimise the number of preferential changes”¹³ (Almeida & O'Brien, 2010, p. 2)

A kérdést a humán és a gépi fordító kompetenciái közötti különbségek (lásd 3. fejezet) felől is megközelíthetjük, de ezek egy ideáltipikus állapotot írnak le, a gyakorlatban a rengeteg kompetenciabeli különbség nem mindig eredményez „rossz” fordítást. Mint Mohácsi-Gorove (2017, p. 183) megjegyzi, „a géppel fordított szövegben egészen más típusúak a hibák, mint az emberi fordításban”, valójában azonban a géppel magyarra fordított szövegek minőségét elemző tanulmányok (pl. Heltai, 2014; Veresné, 2022 stb.) eredményeit vizsgálva megállapíthatjuk, hogy nem a hibák típusa tér el, hanem –

¹¹ Fordítói körülmények között az előszerkesztést jellemzően akkor szokták elvégezni, ha a szöveget több (legalább három) célnyelvre kell lefordítani.

¹² Lásd <https://aclanthology.org/2010.eamt-1.19.pdf> (megtekintés: 2023. február)

¹³ 1. A beavatkozást igénylő problémák [...] azonosításának és megfelelő javításának képessége; 2. Az utószerkesztés észszerű gyorsasággal történő elvégzésének képessége [...]; 3. Az utószerkesztési irányelvek betartásának képessége, hogy a lehető legkevesebb preferenciális változtatást hajtsuk végre. [saját fordítás]

a humán fordító felkészültségétől függően – legfeljebb azok gyakorisága: Helyesírási hibák például a gépi és humán fordításokban is előfordulnak¹⁴, talán az utóbbiakban gyakrabban. A 3. fejezetben említett terminológiai következetlenségek és szórendi hibák szintén megfigyelhetők mindkettőben, bár a gyakoriságra vonatkozó állásfoglalás ezeknél is (naprakész) empirikus megalapozottságot igényelne. Véleményünk szerint tehát az utószerkesztési kompetencia nem az egyes hibák minőségi, hanem legfeljebb mennyiségi szempontjából tér el a lektori kompetenciától¹⁵, de a kérdés végleges megválaszolásához további kutatásokra lenne szükség, amelyek a fordításelmélet és a nyelvtechnológia együttműködésében valósulhatnak meg: Az előbbi az output felől, az utóbbi az input felől járulhatna hozzá a kérdés tisztázásához.

Összegzés

A fordítóképzés ambivalens módon viszonyul a gépi fordításhoz: Egyfelől szeretné elsajátíttatni a fordításhoz szükséges készségeket, másfelől, mivel a képzés egyik és talán legfontosabb célja a munkaerőpiaci elvárásoknak való megfelelés, kénytelen (lenne) beemelni a tantervbe a gépi fordítást, de félt, hogy ezáltal tompítja vagy kioltja a képzés fejlesztő hatását. Ugyanakkor nem biztos, hogy a kettő teljesen kizárja egymást: Szakmai fórumokon gyakran elhangzó érv, hogy megfelelő minőségű utószerkesztés (a lektoráláshoz hasonlóan) csak jól megalapozott fordítási készségek birtokában végezhető, mások véleménye szerint viszont a gépi fordítás a fordítási tudás vagy akár „a lektori kompetencia fejlesztésére” is alkalmas (Heltai, 2014, p. 23), tehát a gépi fordításra (akár önmagában) készségfejlesztő eszközként is lehet tekinteni. Az intenzív együttgondolkodás és az egységesítésre irányuló törekvések ellenére jelenleg nem látni, hogy a kérdést hogyan lehetne megnyugtató módon rendezni, ugyanakkor feltételezhető, hogy a jövőben a hangsúly egyre inkább áthelyeződik a fordítások lektorálására, utószerkesztésére. Ha elfogadjuk, hogy a képzés nem légtüres térben zajlik, akkor kijelenthetjük, hogy az intézményeknek kötelességük választ adni a munkaerőpiaci folyamatokra, és nagyobb mértékben integrálni a gépi fordítást a képzésbe, ami előbb-utóbb óhatatlanul a fordítási készségek fejlesztésének háttérbe szorulásával fog járni. Ezt a súlyponteltolódást a nyelvtechnológusok és a fordítás elméletével/gyakorlatával foglalkozó szakemberek közötti intenzívebb együttműködésnek kellene tükröznie, amely során nagyobb figyelmet kellene szentelni annak, hogy a gépi fordítást miként lehetne hatékonyabban és az adatvédelmi szempontoknak megfelelően hasznosítani. A géppel fordított szövegek utószerkesztése jelenleg önmagában is óriási feladat, tehát nem kell attól tartani, hogy a fordítóképzés kiüresedik (és hogy a nyelvi

¹⁴ Ebből a szempontból Mohácsi-Gorove megállapítása, miszerint géppel fordított szövegekben „helyesírási hiba [...] például sosem fordul elő” nem helytálló, de például a betűtívesztések valóban ritkábbak. A (nehezen algoritmizálható) egybeírás-különírás témakörével például könnyen zavarba lehet hozni a fordítómotorokat: *Social-Media-Kanal* --> *Közösségi média csatorna*, helyesen: *közösségimédia-csatorna* vagy *Spielzeugauto* / *Spielzeugmuseum* / *Spielzeugpistole* --> *Játékautó* / *Játékmúzeum* / *Játékpisztoly*, helyesen: *játék autó* / *játékmúzeum* / *játék pisztoly* stb. (DeepL, 2023. június 8.)

¹⁵ Legalábbis ha eltekintünk a fordítás sajátos esetének tekinthető műfordítástól, ahol nagyobb valószínűséggel fordulnak elő olyan szövegjellemzők (dialektusok, szociolektusok, egyéni beszéd-sajátosságok stb.), amelyeknek a felismerésére és reprodukálására a gép (egyelőre) eleve nem alkalmas.

szakemberek munka nélkül maradnak). Kétségtelen, hogy a hagyományos képzési módszerektől nehéz megválni, de ez a folyamat már a fordítástámogató eszközök elterjedésével elkezdődött, a szakfordítások címzettjei, a piaci szereplők pedig alapvetően a kész fordítás minősége és nem a felhasznált eszközök felől ítélik meg az elvégzett munkát. Sokkal nagyobb dilemmát okoz az a kérdés, amellyel a tanulmányban csak érintőlegesen foglalkoztunk, azaz, hogy a fordításelmélet számára milyen következményekkel járhat a gépi fordítás előtérbe kerülése. Szintén megválaszolásra vár az a kérdés, hogy az utószerkesztéshez valóban szükséges-e fordítási rutin, vagy – a lektorálás különleges altípusaként – attól függetlenül is végezhető.

IRODALOM

- Bowker, L. & Fisher, D. (2010). Computer-aided translation. In Gambier, Y. & van Doorslaer, L. (Eds). *Handbook of Translation Studies* (pp. 60–65). Benjamins. <https://doi.org/10.1075/hts.1.comp2>
- Bruckner, Ch. (2017). Maschinelle Übersetzungs-Plug-ins in CAT-Tools, Sicher (nicht) nutzen? *MDÜ, 2017(4)*, 10–15.
- Cadwell, P. & O'Brien, S. & Teixeira, C. S. C. (2018). Resistance and accomodation: factors for the (non-) adoption of machine translation among professional translators. *Perspectives, 26(3)*, 301–321. <https://doi.org/10.1080/0907676X.2017.1337210>
- de Almeida, G. & O'Brien, S. (2010). Analysing post-editing performance. Correlations with years of translation experience. In European Association for Machine Translation. *Proceedings of the 14th annual conference of the European Association for Machine Translation*.
- Hell, Gy. (1962). *A gépi fordítás néhány kérdése*. Műszaki Könyvtárosok Tájékoztatója, 9(5), 1–14.
- Heltai, P. (2014). Mi az, amit a gép nem tud fordítani? In Fata, I. & Veresné Valentinyi, K. (Szerk.). *Szaknyelv és Szakfordítás. Tanulmányok a szakfordítás és a fordítóképzés aktuális témáiról* (pp. 7–29). Szent István Egyetemi Nyomda.
- Holmes, J. S. (2000) [1972]. The Name and Nature of Translation Studies. In Venuti, L. (Ed.) 2000. *The Translation Studies Reader* (pp. 172–185). Routledge.
- Horváth, P. I. (2011). *A szakfordítások lektorálása. Elmélet és gyakorlat*. Tinta Könyvkiadó.
- Hutchins, W. J. & Somers, H. L. (1992). *An Introduction to Machine Translation*. Academic Press.
- Kenny, D. (2020). Machine translation. In Baker, M. (Ed.). *Routledge Encyclopedia of Translation Studies* (pp. 305–310). Routledge. <https://doi.org/10.4324/9781315678627-65>
- Kóbor, M. (2023). Gépi fordítás (nem csak) a fordítóképzésben – Módszertani adalék a tudatos MT-használat oktatásához. *Modern Nyelvoktatás, 29(1–2)*, 7–23. <https://doi.org/10.51139/monye.2023.1-2.7.23>
- Lakatos-Báldy, Zs. (2022). Tud-e a gép filmet fordítani? *Porta Lingua, 2022(1)*, 59–67. <http://doi.org/10.48040/PL.2022.1.6>
- Laki, L. J. (2018). Mesterséges intelligencia a gépi fordításban. In Tolcsvai Nagy, G. (Szerk.). *A humán tudományok és a gépi intelligencia* (pp. 156–183). Gondolat.
- Laki, L. J. & Yang, Z. Gy. (2022). Jobban fordítunk magyarra, mint a Google! In Berend, G. & Gosztolya, G. & Vincze, V. (Szerk.) *XVIII. Magyar Számítógépes Nyelvészeti Konferencia* 18. köt. (pp. 357–372). Szegedi Tudományegyetem.
- Mohácsi-Gorove, A. (2017). Kifogástalan vagy csak elég jó? A gépi fordítás oktatása a minőség tükrében. In Kóbor, M. & Csikai, Zs. (Szerk.): *Iránytű az egyetemi fordítóképzéshez. A kompetenciafejlesztés új fókuszai* (pp. 173–196). Kontraszt Kiadó.
- Pym, A. (2013). Translation Skill-Sets in a Machine-Translation Age. *Meta: Translators' Journal, 58*, 487–503. <https://doi.org/10.7202/1025047ar>

-
- 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>
- 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>
-

The role of machine translation in translation training

Due to the ever-improving quality of machine translation, the question of how higher education institutions offering translation training should deal with the use of translation engines in training is becoming increasingly urgent. There is almost complete agreement that this topic should be made an integral part of training, but opinions differ as to when and in what form it should be implemented. One argument frequently voiced at professional forums is that the use of machine translation should precede the development of translation skills because this is considered a prerequisite for high-quality post-editing. In this context, however, there is concern that using machine translation engines at too early a stage could result in the inadequate development of translation skills. At the same time, it is unrealistic to prevent their use in the early stages of training, as these tools are open access. After an overview of the features of machine translation, this paper aims to offer some hints on solving this dilemma.

Keywords: machine translation, translation technology, translation training, postediting, MTPE