

Eötvös Lóránd Tudományegyetem
Bölcsészettudományi Kar
Nyelvtudományi Doktori Iskola
Fordítástudományi Doktori Program

A DOKTORI ÉRTEKEZÉS TÉZISEI

LENGYEL ISTVÁN

A fordítási hiba fogalma funkcionális
megközelítésben

Témavezető: Dr. Prószéky Gábor, DSc, egyetemi tanár

Budapest, 2013

1. A kutatás tárgya, az értekezés célja

Mi a fordítási hiba? Milyen jellegű fordítási hibák vannak? Mi határozza meg, hogy mikor tekintünk hibának egy jelenséget a fordításban? A fordítási hiba kategóriája önmagában nem értelmezhető: „A vélekedés arról, hogy mi számít fordítási hibának, a fordítási elméletek és normák szerint más és más” (Hansen 2010). Hiba az, ami az elvárásoktól eltér, tehát először az elvárásokat kell meghatározni. Az elvárás lehet elméleti jellegű, mint például az ekvivalencia valamely fajtájának megvalósítása, gyakorlatibb jellegű, például megfelelés a fordítói útmutatónak, adott szövegtípus követelményeinek, vagy társadalmi jellegű, például megfelelés a hivatásos fordítók közössége által meghatározott viselkedési normáknak (Chesterman 1993).

Vannak azonban egyértelmű és nem egyértelmű hibák – ezeket Pym (1992) bináris és nem bináris hibáknak nevezi. Nem minden hiba egyformán fontos viszont: míg egy helyesírási hiba sok nyelven bináris jellegű, aligha csorbítja annyira a fordítás értékét, mint egy félrefordítás – hacsak a másik szó nem értelmes és félrevezető.

Kutatásom végső célja egy automatikus, nyelvfüggetlen statisztikai módszer megalkotása volt valamilyen hibatípus észrevételére. A hibatipológiák és a vállalati hibakeresési gyakorlatok felmérése, továbbá az automatikus hibakeresési szoftverek áttekintése után vizsgált hibatípusként kiválasztottam a kihagyásokat és beszúrásokat, azaz az információs aszimmetria észrevételét. Kérdésként felmerült azonban, mennyire lehet ezeket binárisnak nevezni: erre mutat rá a négy korpuszon, két nyelvpárban, három-három lektorral végzett kísérletem.

2. Az értekezés felépítése

Az értekezés szerkezetében először feltérképezi a lehetséges hibatípusokat, ezután a gépi és emberi hibakeresés lehetőségeit, végül pedig kísérletet tesz a beszúrások és kihagyások automatikus felderítésére. A lehetséges hibatípusok feltérképezését a második fejezetben a fordítástudományi irodalom feldolgozásával kezdem, majd a 3. fejezetben a nyilvánosan elérhető hibatipológiákat veszem górcső alá. A 4. fejezetben megvizsgálom hét cég és egy non-profit szervezet minőségértékelési módszereit. A gépi és emberi hibakeresés lehetőségeinek

feltérképezése a gépi minőségellenőrző szoftverek funkcióinak vizsgálatával kezdődik az 5. fejezetben, amelyet a 6. fejezetben egy kísérlet követ, a kihagyások és beszúrások lektori vizsgálatának pontosságára vonatkozóan. A 7. fejezetben tesztek kísérletet a beszúrások és kihagyások automatikus felderítésére egy módszer, az ún. múzsa segítségével. Ezt az eredmények összegzése, majd az irodalomjegyzék követi.

3. A kutatási módszerek

A kutatás során többféle, egymást kiegészítő módszert alkalmaztam a különböző kutatási kérdések megválaszolására. Hagyományos, könyvtári és internetes kutatómunkát végeztem a szakirodalmi előzmények összegyűjtéséhez. A szakirodalmi kutatás során az értekezés szűkebb területére, a vállalati „emberi” (azaz nem géppel végzett) fordítások értékelésére koncentráltam, és nem vizsgáltam részletesen sem a pedagógiai célú értékelések, sem a gépi fordítás szakirodalmát.

A nyilvánosan elérhető hibatipológiákat először össze kellett gyűjtenem, amelyben segítségemre volt az internet és több fordítóiroda képviselője is. Így jutottam hozzá például a LISA QA és a SAE J2450 leírásához, amelyek nem nyilvánosak. A hibatipológiák részletesebb tanulmányozásához főleg internetes hivatkozásokat használtam fel, de sokat segített a John Benjamins kiadó ATA Scholarly Monograph Series kiadványsorozata is, amely sokszor gyakorlatias témákkal foglalkozik. A TAUS DQF modell csak a TAUS tagjai számára nyilvános, ennek zárt részéhez hozzáférést kaptam, és beszélgetést is folytattam Rahzeb Choudhury-val, aki a DQF modell egyik szerzője.

A vállalati hibatipológiák megismeréséhez, feldolgozásához először fordítóirodáktól kértem javaslatokat olyan vállalatokra, amelyek a fordításaikat rendszeresen értékeltek. Ezután kapcsolatba próbáltam lépni 25 vállalattal, amelyből 8 mutatkozott nyitottnak arra, hogy értékeljem módszerüket a dolgozatban. Az ő modelljeiket telefonos interjú útján ismertem meg, de sokan közülük az értékelésre használt Excel-táblázatot is átküldték.

Az automatikus fordításellenőrző szoftverek funkcióinak felderítéséhez egyrészt szakirodalmi alapokra helyezkedtem – a legjobb összefoglalást erre Makoushina (2007) adja –, másrészt pedig letöltöttem a programok próbaverzióit, illetve a programok online súgójából tájékozódtem.

Az értekezésben két kísérletet végeztem el. Mindkét kísérletben ugyanazt a korpuszt használtam, de míg az egyik kísérletben az emberi lektori döntések számszerűsítése került a középpontba, a másik kísérletben kétféle módszert hoztam létre egy hibafajta automatikus megállapítására, és ezt vettem össze az elrontott korpussszal. Az első kísérlet során valós lektorálási feladatot szimuláltam korpuszonként 3-3 lektor (összesen 12 lektor) bevonásával, feltételeztem jó fordítások egy részének elrontásával (részletesebben lásd alább). Azért három főre esett a választásom, mert a gépi fordítások emberi fordítással történő összehasonlítását szolgáló metrika, a BLEU is három referenciafordításból indul ki (Papineni et al 2002). A lektorok feladatukat tényleges, órabérben fizetett megbízásként, fordítóirodától kapták, nem történt említés sem arról, hogy ez egy tudományos kutatáshoz szükséges. A kísérlet eredményeit egy ugyanolyan kétnyelvű XLIFF-fájlban (Krenz – Ramlow 2008) kaptam meg, mint amelyet lektorálás után a lektorok egyébként is leadnak. Ezzel próbáltam elkerülni, hogy a lektorok gyanakodjanak arra, hogy ez egy kísérlet, és megpróbálják megérteni, hogyan készítettem elő a dokumentumokat, hiszen ez torzíthatta volna az eredményt – a mintavételezéses lektorálás nem ismeretlen gyakorlat.

A másik kísérlet, az automatikus hibakeresés megvalósítása két fő lépésből állt: a fejlesztésből és az eredmények kiértékeléséből. A fejlesztés során kialakítottam az algoritmust egy meglévő eszköz, az ún. múzsa alapján. A múzsa kiválasztása az eszköz tulajdonságai miatt logikus lépés volt. A múzsa ugyanis forrás- és célnyelvi kifejezéspárokat rendel össze, azaz nyelvfüggetlenül kivonatol kifejezéspárokat egy-egy korpuszból. A fejlesztés tényleges része a kihagyás/beszúrás múzsaalapú kiértékelése volt, melynek során a szegmenst a múzsa találatai alapján próbáljuk lefedni. Az értékelés automatizált mérésekkel történt, az ellenőrzés által megjelölt szegmenseket összehasonlítottam a korpuszba szándékosan bevezetett hibákat tartalmazó szegmensek listájával. Mivel az empirikus eloszláson alapuló korreláció nem volt annyira erős, hogy jóslóerővel bírjon, a kísérlet eredményeit nem volt értelme a lektori eredményekkel összehasonlítani, és az sem jelentett jelentős torzítást, hogy a tanítókorpusz tartalmazta az elrontatlan szegmenseket is (ez az eredményt csak tovább ronthatna volna).

A következőkben bemutatom a kísérletekhez használt korpuszt.

4. A korpusz összetétele, mérete és előkészítése

A két kísérlethez négy párhuzamos korpuszt használtam, ebből kettőt angol-német, kettőt angol-magyar nyelvpárban. A kísérletben céлом volt az angol-magyar és angol-német értékelések összehasonlítása, ezért két-két hasonló témájú és bonyolultságú korpuszt kerestem. Így esett végül a választás a következő korpuszokra:

1. A DGT angol-magyar fordítómémória 2012-es kiadása (Steinberger et al 2012). A DGT az Európai Bizottság Fordítási Főigazgatósága, ahol 2007 óta közzéteszik a kézzel készített fordításokból készült fordítómemóriát 23 nyelven. A fordítást professzionális fordítók meghatározott szabályok és meghatározott terminológia alapján készítették, nem párhuzamosítással keletkezett a memória. Erre EU EN-HU-ként is hivatkozom.

2. A DGT angol-német fordítómémória 2012-es kiadása (EU EN-DE). Ez szinte megegyezik az angol-magyar fordítómemóriával az angol nyelvű szegmensek tekintetében.

3. A Hunglish angol-magyar korpusz informatikai részhalmazának egy adott része (Varga et al 2005). Ez a korpusz szövegpárhuzamosítással keletkezett, nem fordítástámogató szoftverrel történő fordítással.

4. A memoQ fordítástámogató szoftver német sűgójának fordítása angolból. Ez a szöveg szintén informatikai jellegű, de fordítástámogató szoftverrel készült.

A korpuszok kiválasztása után a korpuszokat „egyeltem” (eltávolítottam az alternatív fordításokat), és eltávolítottam a gyorsan észrevehető rossz fordításokat is. Így kaptam meg a kísérletben felhasznált korpuszokat, amelyeknek a kísérlet szempontjából legfőbb adatait a következő táblázat mutatja be:

Korpusz	Szegmen- sek száma	Forrás- szegmens átlagos hossza (szó)	Forrásszeg- mens átlagos hossza (leütés)	Célszegmens átlagos hossza (szó)	Célszegmens átlagos hossza (leütés)
Hunglish EN-HU	14 518	15,60	92,05	13,04	97,15
memoQ EN-DE	20 195	10,25	62,72	9,58	75,89
EU EN-DE	203 313	19,59	128,51	17,40	136,05
EU EN-HU	203 617	19,57	128,74	15,78	132,20

Feltételezvé, hogy a korpuszokban a szegmensek fordításai helyesek, a korpuszokból nagyjából 100-100 szegmenst véletlenszerűen elrontottam, úgy, hogy szándékosan kitöröltem a fordításból néhány szót, vagy beszúrtam pár szót egy vagy több helyre. A beszúrásnál és a kihagyásnál próbáltam a hibákat úgy beszúrni, hogy a mondat grammatikus és értelmes maradjon.

A fenti korpuszokat teljes méretben csak a gépi módszer előkészítéséhez használtam fel. A lektorok döntéseinek összehasonlításához és a gépi módszer ellenőrzéséhez a korpuszokból kb. 300 szegmens részhalmazokat képeztem, amelyekben szerepelt az összes elrontott szegmens és még kétszer annyi érintetlen szegmens. Az elrontott szegmensekből eltávolítottam a hibafajtról szóló információt, így nem lehetett megkülönböztetni formailag a hibás és az érintetlen szegmenseket.

5. A kutatás újszerűsége

A fordításértékelési szabványok 2012-2013-ban ismét aktuálissá váltak a gépi fordítás és az utószerkesztés előretörése miatt, mivel ezeknek az értékelése, a hibáinak a kiszűrése némiképp eltér az emberi fordításétól. Ez a kutatás adja ezen kategóriáknak a legrészletesebb bemutatását és értékelését, továbbá szempontokat ad a hibakategóriák átgondolására, értékelésére is. Az értékelt szabványok között több új, 2012-2013-as vagy még tervezés alatt álló szabvány (ISO WD 14080, TAUS DQF, MQM) szerepel, amelynek értékelését nemzetközi szakirodalomban sem találtam.

Korábban a vállalati minőségértékelési gyakorlatot tudományos szempontból – tudomásom szerint – nem vizsgálták, ezek az interjúk teljesen új ismeretanyagot hoztak létre.

Az emberi lektorok összhangjának mérésére kevés kutatást találtam, hasonló jellegű kutatást csak a GREVIS projekt (Brunette et al 2005) és Horváth Péter Iván (2011) végzett. Ez a kutatás a kétnyelvű lektorálást mérte, annak egyezését több lektor esetében, és ezt vetette össze egy „aranystandarddal”, amely a szándékosan bevezetett hibák megtalálása volt.

A tudomány számára új, de az iparágban már van rá példa, hogyan lehet a tényleges nyelvi fordítási hibákat nyelvtechnológia segítségével megítélni. Ennek módszerei, elvei számomra ismeretlenek voltak a kutatás során és jelenleg is azok, de tudok róla, hogy a Lionbridge és a Welocalize, két igen nagy fordítóiroda, hasonló eszközök kifejlesztésén fáradozik.

6. Az értekezés által kijelölt új kutatási irányok

A dolgozat kijelöl néhány kutatási irányt: érdekes lenne megtudni, hogy egy „optimális” szöveghez hány lektor kell. Hagyományos fordítóirodai munkafolyamat az ún. TEP, *translation-editing-proofreading*, amelynek hasznosságát sok fordítási szakember megkérdőjelezte már, de kutatás még sosem vizsgálta. Vajon két lektor megtalál minden jelentős hibát? Illetve milyen hibafajtákat talál meg jó eséllyel egy lektor, két lektor, három lektor stb., technológiahasználattal vagy anélkül? Az értekezés megadja a módszertani keretet ennek méréséhez.

A dolgozatomban eredetileg még egy kérdést kívántam vizsgálni, ez pedig a fordítási hibák megítélése a hibatipológiák alapján. Ugyanazt a hibát mennyien azonosítják egyformán, egy hiba hány hibatípusba tartozhat bele lektoroktól és hibatipológiáktól függően? A CD-ROM-on található adatok között már előremutató kísérleti adatot is találhatunk, mivel a négy korpuszból három esetében a fordítók nem csupán a kihagyást/beszúrást, hanem a félrefordítást is megjelölhették hibaként (noha az előkészítés során másféle jelentésmódosítást nem vezettünk be). Ebben az esetben 2-es számérték szerepel az Excel-fájlokban az 1-es helyett. Mivel szándékosan csak kihagyás/beszúrás szerepelt a dokumentumokban, ezért a vizsgálatnál erre nem voltam figyelemmel, akkor is kihagyást/beszúrást számoltam, amikor a lektor félrefordítást jelölt. Ezt a kísérleti adathalmazt szívesen felajánlom további kutatásra.

Érdeemes lenne tovább vizsgálni a beszúrási/kihagyási hibafajtákat is elemző módszerrel: nem törekedtem arra, hogy a különböző beszúráásokat és kihagyásokat csoportba rendeljem valamilyen szempont alapján, de elképzelhető, hogy például a zárójelbe tett kifejezések kihagyását másképpen ítélik meg a lektorok, mint egy kihagyást egy többszavas tulajdonnévből. Ezt a meglévő szövegek elemzésével már lehet vizsgálni.

A kísérletben érdekes eredmény, hogy kialakult egy angol-magyar és egy angol-német lektori „profil”: jelentős eltéréseket találtam a megjelölt szegmensek száma és – ezután már kevésbé meglepően – a szegmensek pontossága között. Érdeemes lenne ezt a vizsgálatot továbbvinni, és elegendő adattal alátámasztani (vagy épp megcáfolni).

Az automatikus módszer nem lett igazán sikeres, bár nagyobb eséllyel megjelöli az ilyen jellegű hibákat, mint a véletlen mintavétel, és így időt takaríthat meg, ez nem

elegendően megbízható arra, hogy a fordítástámogató szoftverekbe ezt a módszert bevezessük. Úgy gondolom azonban, hogy ha a múzsát kicserélnénk egy másik indexelési módszerre, akár csak a feltételes valószínűség módszerére, más eredményeket kapnánk. Elképzelhető, hogy ennek a módszernek a finomhangolásával jelentős javulást érhetnénk el, de mivel a múzsából kiinduló értékelés is igen logikusnak tűnt számomra, és ennél jobb eredményekre számítottam, ebben egyáltalán nem vagyok biztos.

7. A kutatás eredményei

1. tézis: Dolgozatomban magyar nyelven először mutatom be a minőségbiztosítással kapcsolatos szabványokat, az általuk alkalmazott hibatipológiákat, továbbá rámutatok, hogy a minőségmérésben alkalmazandó hibatipológiát a szöveg sajátosságai alapján érdemes kiválasztani. Azt állítom, hogy a hibatipológia akkor megfelelő, ha a munkafolyamatbeli szerepeket és az előállítási technológiát nem keveri össze a fordítási hibák észrevételével, azaz vagy okokat, vagy hibákat keres.

Részletesen megvizsgáltam a SAE J2450, a LISA QA modellt, az MQM és a DQF, továbbá az ISO 14080 WD hibakategóriáit, értékelve az egyes hibákat. Bemutattam az ATA és az ELTE vizsgafordítás-értékelési hibakategóriáit, az ITS Tag Set 2.0 hibakategóriáit, továbbá az SDL TMS és a memoQ beépített hibakategóriáit. Rámutattam, hogy egyes hibakategóriák, például a nyelvtani hiba, a terminológiai megfelelési hiba stb. jelenségeket mutatnak, míg mások, például a félrefordítás a forrásszöveg hibája miatt, vagy a nem megfelelő 100%-os találat, okokat keresnek. Bemutattam, miért okoznak ezek többértelműséget a modellben. A tézis állításának bizonyításán túl javaslatokat is tettem gyakorlatias, egyértelmű hibatipológiák összeállításához.

2. tézis: Megvizsgálom, hogy az egyes nagyvállalatok, főleg az informatikai nagyvállalatok, hogyan értékelik a fordításokat hibatipológiákkal, milyen hibákat különböztetnek meg és milyen információkat gyűjtenek be a lektoroktól. Rámutatok, hogy a vállalati hibatipológiák ihletet merítenek a szabványokból, de ritkán alkalmazzák tisztán a szabványokat.

Az értekezésben először mutattam be akár a magyar, akár a nemzetközi szakirodalomban ennyi vállalat hibatipológiáját. Ezeket interjúk és megkeresések

alapján tudtam leírni. Láthattuk, hogy csak a HP és a Google mondja magáról, hogy a LISA QA modellt követi, de ezeknél is vannak eltérések, például a Google a forrásszöveg félreértését is bevezeti a modellbe. A többinél – például VeriSign, Microsoft – jelentős hasonlóság van a LISA QA-val, a két igazán eltérő a McAfee és a Symantec modellje.

3. tézis: Bemutatom az automatikus hibafelismerés formáit, használati körét, és rámutatok, hogy csakis egyszerű, szabályalapú hibakeresést valósítottak meg eddig ezen a területen. Megvizsgálom és összegzem, hogy milyen hibatípusokat lehet és milyen hibatípusokat nem lehet jelenleg számítógépes eszközökkel vizsgálni.

A dolgozatban bemutattam az automatikus hibafelismerés irodalmát, amely igen szűk, továbbá adtam egy összegzést a számítógépes eszközökkel vizsgálható és nem vizsgálható fordítási hibákról.

4. tézis: Kísérlettel megvizsgálom, hogy tapasztalt lektorok milyen pontossággal állapítják meg a beszúrás/kihagyás tényét angol-német és angol-magyar fordításokban, azaz milyen pontosság várható el a hibák észrevételénél számítógépes eszközökkel. Bizonyítom, hogy attól függően, hogy mit tekintünk fordítási hibának, jelentős különbségek lehetnek az eredményekben.

A disszertációban bemutatott kísérlet alapján attól függően, hogy mit tekintettünk hibának (hiba-e az, amit mi szándékosan elrontottunk, illetve hány lektornak kell egy szegmenst hibásnak minősítenie ahhoz, hogy azt hibásnak fogadjuk el), a helyesen észrevett hibák száma között legalább 2,34-szeres, legfeljebb 5,1-szeres, az észre nem vett hibák között legalább 4,18-szoros, legfeljebb 8,46-szoros, a helyesen nem jelölt hibák között (a legszámosabb kategóriában) legalább 1,07-szeres, legfeljebb 1,68-szoros, míg a tévesen bejelölt szegmensek között legalább 3,8-szoros, legfeljebb 16,6-szoros különbség volt a négy korpusz között a legjobb és a legrosszabb esetben. Észrevettük, hogy a lektorok a kihagyási és beszúrási hibák jelentős részét, legalább 35-45%-át nem találják meg közepesen engedékeny hibadefiníció mellett. A legalaposabb lektor a hibák 26%-át nem találta meg, míg a legkevésbé alapos lektor a hibák 57%-át nem találta meg. Bizonyítottuk, hogy a szegmensek közül nagy valószínűséggel azokat jelöli meg egy lektor, amelyek hibásak, viszont olyanokat is megjelöl, amelyek nem hibásak, és a hibázási arány növekvő szegmensmegjelölés mellett növekszik: tehát először a biztos hibákat jelölik meg. Ez sugallhatja a hibadefiníció konszenzusjellegét is, és rávilágít arra, hogy Pym (1992) két elméleti hibakategóriája tényleg létezik, azonban ezek nem feltétlenül

kapcsolódnak a hiba típusához: vannak olyan karakterisztikák, amelyek mentén egyetértés van, és vannak olyanok, amelyek mentén nincsen.

5. tézis: Bemutatok egy próbálkozást a beszúrási/kihagyási számítógéppel történő nyelvfüggetlen kiértékelésére, értékelem ennek megbízhatóságát az emberi lektoráláshoz képest.

A dolgozatban bemutattam két módszert, a múzsaalapú módszert és a szegmenshosszarányon alapuló módszert, amely nem közelítette meg az emberi elemzés hasznosságát. A szegmenshosszszelvény jobban tudta jósolni a hibák előfordulását, és jobban teljesített beszúrási, mint kihagyási, míg a múzsaalapú módszer a kihagyást találta meg jobban. Némi eredményként szolgál, hogy ha ezekkel a módszerekkel választunk ki szegmenseket, azok 7%-128%-kal nagyobb valószínűséggel tartalmazzák a hibás szegmenseket, mint a véletlenszerű kiválasztás. Ez mintavételezési algoritmusnak hasznos lehet, ha a mintavételezésnél a cél a hibák kiszűrése minél kisebb idő- és pénzbefektetéssel.

8. A tézisekhez felhasznált irodalom

- Chesterman, A. 1993. From 'is' to 'ought': Laws, norms and strategies in translation studies. *Target*. Vol 5. No. 1. 1–20.
- Hansen, G. 2010. Translation 'errors'. In: Gambier, Y., van Doorslaer, L. (eds). *Handbook of Translation Studies*. Amsterdam: John Benjamins. 385–388.
- Horváth Péter Iván 2011. *A szakfordítások lektorálása. Elmélet és gyakorlat. Segédkönyvek a nyelvészet tanulmányozásához 117*. Budapest: Tinta Könyvkiadó.
- Krenz, M., Ramlow, M. 2008. *Maschinelle Übersetzung und XML im Übersetzungsprozess*. Berlin: Frank & Timme Verlag.
- Makoushina, J. 2007. Translation Quality Assurance Tools: Current State and Future Approaches. In: *Proceedings of the 29th International Conference on Translating and the Computer*. London: ASLIB.
- Papineni, K., Roukos, S., Ward, T., Zhu, W.-J. 2002. BLEU: A Method for Automatic Evaluation of Machine Translation. In: *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*. Philadelphia: ACL. 311–318.

- Pym, A. 1992. Translation Error Analysis and the Interface with Language Teaching. In: Dollerup, C., Loddegaard, A. (eds.). *Teaching Translation and Interpreting*. Amsterdam: John Benjamins. 279–290.
- Steinberger, J., Eisele, A., Kloczek, S., Pilos, S., Schlüter, P. 2012. DGT-TM: A freely Available Translation Memory in 22 Languages. In: *Proceedings of the 8th international conference on Language Resources and Evaluation (LREC'2012)*. Istanbul: LREC.
- Varga, D., Németh, L., Halácsy, P., Kornai, A., Trón, V., Nagy, V. 2005. Parallel corpora for medium density languages. In: *Proceedings of the RANLP 2005*. Borovets: RANLP. 590–596.

9. Publikációs jegyzék

Az értekezés témájához kapcsolódó publikációk és tudományos előadások

- Lengyel I., Kis B. 2003. Új módszerek az emberi fordítás gépi támogatásában. In: Gyimóthy Tibor (ed.) *Az I. Magyar Számítógépes Nyelvészeti Konferencia gyűjteményes kötete*. Szeged: Szegedi Tudományegyetem. 268–274.
- Lengyel I. 2003. Dynamic dictionaries: similarity vs. equivalence. In: Ballard, Michel (ed.) *Qu'est-ce que la traductologie?* Arras: Université d'Artois.
- Lengyel I., Kis B., Ugray G. 2004. MemoQ – új megközelítés a fordítástámogatásban. In: Gyimóthy Tibor (ed.) *A II. Magyar Számítógépes Nyelvészeti Konferencia gyűjteményes kötete*. Szeged: Szegedi Tudományegyetem. 268–274.
- Lengyel I. 2004. Group translation – exploiting synergy. In: Birgit Strotmann (ed.) *Proceedings of IV Jornadas sobre la Formación y Profesión del Traductor e Intérprete*. Madrid: Universidad Europea de Madrid. Available online: <http://www.uem.es/web/fil/invest/publicaciones/web/EN/RED.HTM>
- Lengyel I., Kis B. 2005. A fordítás számítógépes segédeszközéről. In: Dobos Csilla et al. (ed.) *Mindent fordítunk, és mindenki fordít. Tanulmánykötet Klaudy Kinga tiszteletére*. Bicske: SZAK Kiadó. 53–60.
- Lengyel I. 2004. Az Európai Unió és a túlterminologizálás. In: Cs. Jónás Erzsébet és Székely Gábor (ed.) *Nyelvek és nyelvoktatás Európa és a Kárpát-medence régióiban. 1/2. kötet. A XIV. Magyar Alkalmazott Nyelvészeti Kongresszus Előadásai*. Pécs–Nyíregyháza: MANYE – Bessenyei Kiadó. 213–218

Recenzió:

Lengyel I. 2005. Ljuba Tarvi: Comparative Translation Assessment: Quantifying Quality. *Fordítástudomány* VII. évf. 1. szám. 112–115.

Előadás:

Lengyel I. 2006. (előadás) A fordító, a fordítómemória meg a konzisztencia. Elhangzott: XVI. MANYE Kongresszus. Gödöllő: SZIE, 2006. április 10–12.

Lengyel I. 2009. (előadás) The Future of TM Management. Elhangzott: tekomp Jahrestagung. Wiesbaden: Rhein-Main Hallen, 2009. november 5.

Lengyel I. 2011. (előadás) Catalyzing Business Model Innovation. Elhangzott: TAUS. Enabling better translation. User conference. Santa Clara: Hyatt Regency, 2011. október 6.

Lengyel I. 2013. (előadás) Translation Tools Processing Each Other's Formats – Dream or Not? Elhangzott: AFIT Romanian Translation Forum. Bukarest: Capital Plaza, 2013. október 25.

Lengyel I. 2013. (előadás) Terminology in the Cloud with memoQ and TAAS. Elhangzott: Creation, Harmonization and Application of Terminology resources. Wiesbaden: Rhein-Main Hallen, 2013. november 7.