



SEIDL-PÉCH OLÍVIA. A FORDÍTÁSOK LEXIKAI KOHÉZIÓS MINTÁZATA. FORDÍTOTT ÉS AUTENTIKUS MAGYAR SZÖVEGEK KORPUSZNYELVÉSZETI VIZSGÁLATA

Budapest, ELTE BTK Fordító- és Tolmácsképző Tanszék. 2024.

ISBN 978-963-489-735-4 (online). ISSN 3057-8043 (online). <https://doi.org/10.21862/fordtudert4>

Csatár Péter

csatarpeter@unideb.hu

Debreceni Egyetem, Germanisztikai Intézet

Német Nyelvészeti Tanszék

<https://orcid.org/0000-0001-5157-2840>

Köztudott, hogy a magyar a legtöbbet fordított és a legtöbbet fordító nyelvek közé tartozik az UNESCO gondozta Index Translationum adatbázisának az adatai alapján (<https://www.unesco.org/xtrans/bsform.aspx>). A top 50 célnyelv közül a magyar a 15. helyet foglalja el, ami azt jelenti, hogy a fordításkultúránk (elsősorban a tudományos szakkönyvek, és a szépirodalmi alkotások terén) mind tartalmi, mind nyelvi tekintetben nagyon aktív alakítója a magyar kultúrának és a tudományos közéletnek is. Fontos tény az is, hogy a magyar mint forrásnyelv is az előkelő, 19. helyen áll a legtöbbet fordított nyelvek rangsorában, amiből nyugodtan következtethetünk arra, hogy a magyar szellemi élet eredményeit más nyelvek és kultúrák is szívesen fogadják be.

Talán meglepő, de ezen adatok tükrében is fontos és tanulságos Seidl-Péch Olívia monográfiája, amely azt vizsgálja, hogy a magyarra fordítás során a fordított szöveg lexikális kohéziója mennyiben viseli nyomán a forrásnyelvi szövegekre jellemző lexikális-kohéziós mintázatokat, vagyis azt, hogy a lexikális kohéziós mintázataik tekintetében a magyarra fordított szövegek vajon eltérnek-e, és ha igen mennyiben és milyen értelemben az autentikus magyar szövegek lexikai mintázataitól? A Seidl-Péch által elvégzett nagymintás korpusznyelvészeti vizsgáldások egyik messzemenő tanulsága, hogy az eltérések arra (is) rávilágítanak, hogy az idegen nyelvekről magyarra fordított szövegek nem csak a tartalmuk, de a nyelvi megformáltságuk

mintázatai révén is hatással vannak (vagy lehetnek) a magyar nyelvű szövegalkotás normáira, stílusára és ezeken keresztül a magyar nyelvi szövegkultúrára általában is (Seidl-Péché 2024: 13).

Seidl-Péché lexikaikohéziós vizsgálatai egy *nyolc részből álló* kötetben kapnak helyet. A kötet első, *bevezető része* egyfelől explikálja a szerző által végzett korpusznyelvészeti kutatás célkitűzéseit, másfelől pedig kitér azokra a korábbi vizsgálódásokra, amelyek a kötet fókuszában álló ún. *fordított szövegek* sajátosságainak, és ezekből kiindulva a fordított szövegek és az autentikus (azaz nem fordított) szövegek eltérő nyelvi mintázatainak feltárására, továbbá az eltérésekért felelős nyelvi és nem nyelvi tényezők azonosítására irányultak. Szakirodalmi eredményekre támaszkodva a szerző abból indul ki, hogy a fordított és nem fordított szövegek közötti eltérések „nem csupán a nyelvi struktúrák szintjén jelentkeznek, hanem mélyebb, szemiotikai és pragmatikai szinteken is” (Seidl-Péché 2024: 7) jelen vannak. Ezen szintek közül a szerző érdeklődésének homlokterében a nyelvi szint, azon belül is a lexikai kohéziós mintázatok eltérése áll, mivel célja „egy nyelvtechnológiai alkalmazásokkal támogatott elemzési módszer” (uo.) kifejlesztése és alkalmazhatóságának tesztelése „nagy mintájú szövegvizsgálathoz” (Seidl-Péché 2024: 8).

A kötet *második részének* fejezetei – összesen hat fejezet – olvashatók tudománytörténeti eszmefuttatásokként is, amelyek kitérnek a fordítástudomány paradigmatis elméleti és módszertani irányzataira. Seidl-Péché kiemeli a fordítástudomány leíró jellegét és tárgyának komplexitását, a nyelvi és kulturális rendszerek különbségeit, rámutatva arra, hogy a fordítást befolyásoló forrás- és célnyelvi tényezők normalizáló hatása mellett fordítási univerzálék (nyelvpartól független fordítási mintázatok) is megjelennek a fordítás produktumában, mint amilyen a forrásnyelvi szöveg lexikai ismétléseinek kerülése a célnyelven, vagy éppen az explicitáció (Seidl-Péché, 2024, p. 33).

Azonban ezek a történeti szálak a *második részben* mind egy irányba, pontosabban egyazon kutatási tárgyra, és pedig a *fordított szövegre* mutatnak. A *fordított szöveget* Klaudy (2007) tanulmányára és saját megelőző kutatásaira (Péché 2010) támaszkodva a „nyelvek fordítási viselkedéseként” határozza meg: „A ’nyelvek fordítási viselkedése’ ebben az esetben azt a jelenséget jelöli, amikor az adott forrásnyelvű szöveg másképpen viselkedik, ha különböző célnyelvekre kell lefordítani” (Seidl-Péché, 2024, p. 20). Az eredmény pedig a fordított szöveg *kvázi helyessége* (Klaudy, 2007, p. 139) lesz, melyet Heltai Pál így hoz fogalmi formára: „A fordított szövegek eltérései a nem-fordított szövegektől disztribúciós jellegűek: a fordított szövegek nyelvtanilag helyesek, de bizonyos szerkezetek vagy szavak nagyobb vagy kisebb gyakorisággal fordulnak elő, mint a hasonló típusú nem fordított szövegekben” (Heltai, 2009, p. 23, id. Seidl-Péché, 2024, p. 23). A kvázi helyesség és a fordításnyelv (vagy ahogyan Frawley (1984) nevezi, a *harmadik kód*) terminusai tehát tulajdonképpen ugyanannak a jelenségnek több aspektusból való elemzésére utalnak: míg a kvázi helyesség a célnyelvi szöveg strukturális értelemben vett (kvázi) normasértő voltára utal, addig a *fordításnyelv* (*translationese*) a jelenség hatását domborítja ki a célnyelvi kultúrában, egyúttal arra is utalva, hogy ez a hatás megmaradhat és gyökeret verhet a célnyelvi szövegalkotási folyamatokban (lásd pld. a magyar jogi szaknyelv történetének egyes jelenségeit, Szép 2009).

A kötet tudománytörténeti eszmefuttatásaiból az világlik ki, hogy a kutatók a *fordított szöveget* sokféleképpen értelmezik, pl. mint az interkulturális kommunikáció hatását a célnyelvre, amely célnyelvi fordításokban (és kulturális hatásokban) ölt testet, vagy mint „rontott” célnyelvi szöveget, amely kényszerűen magán viseli a forrásnyelvi megformálás sajátosságait (idegenszerűségét), és ami ezáltal nem is tagozódik be megfelelően a célnyelvbe, és negatív konnotációt kap a legtöbb kutatónál. A kötet *második része* alaposan körüljárja a fordított szöveg tematikáját. Külön alfejezet foglalkozik a fordításoknak az anyanyelv használatára, azon belül is a célnyelvi normára gyakorolt hatásával, valamint a fordítói kompetenciával és az ekvivalencia fogalmával is. Seidl-Péché a vizsgálódásai során a bőséges idegennyelvű szakirodalom mellett elsősorban Balaskó Mária fogalmi különbségtételére támaszkodik, aki a fordításnyelv fogalmában elválasztja egymástól annak két komponensét: „A fordításnyelv két komponensből tevődik össze: egyrészt azokból a jegyekből, amelyek a célnyelvi rendszerhez tartoznak ugyan, de használatuk valamilyen téren különbözik a célnyelvi normától, a tipikus célnyelvi jegyek használatától (eltérő gyakoriság, disztribúció, nyelvi minták). A harmadik kód másik komponensét azok a jelenségek alkotják, amelyeket a leíró fordítástudomány fordítási univerzáléknak nevez (pl. az explicitáció, a szimplifikáció, a normalizáció (Balaskó, 2007, pp. 161-162).”

A fordításnyelv több oldalról megközelíthető jelenség tehát, ezért amikor a fordított szöveg strukturális sajátosságai és hatása képezi a vizsgálat tárgyát, akkor a tárgy komplexitásából adódó nehézségeket kiküszöbölendő, valamint a kutatás empirikus megalapozottságát szem előtt tartó elemzésnek mind elméleti, mind módszertani szempontból ki kell jelölnie a maga kereteit.

Seidl-Péché Olívia vizsgálódásai kiválóan példázzák a tudományos munkának ezeket az igen csak fontos paramétereit. Kutatásának tárgyaként a *fordított szöveget* (valamint ennek forrás- és a célnyelvi párhuzamos szövegeit) mint egymásra vetíthető lexikális kohéziós hálózatokat tárja elének egy olyan vizsgálati keretben, amely képes a szövegalkotók és fordítók döntéseinek szubjektivitását megfelelően kompenzálni, mégpedig egyrészt nagy mennyiségű szöveg korpusznyelvészeti feldolgozásával, másrészt a WordNet ontológiájának mint annotálási rendszernek az alkalmazásával. A kezdeményezés újszerűsége vitathatatlan, mivel a megszokott korpusznyelvészeti kollokációs kohézió kutatással szemben „a mondathatárokon túlmutató szemantikai kapcsolatok feltárásán nyugszik” (Seidl-Péché, 2024, 10), melynek notációs eszközét a WordNet ontológiai formalizmusa biztosítja.

A WordNet formalizmusával külön alfejezet (5. alfejezet) foglalkozik a kötet ötödik részében, mivel ezek a formalizmusok képesek megragadni a mondathatáron túlnyúló lexikai kohéziós kapcsolatokat is. A főnevek esetében többféle ilyen viszonyrendszert azonosíthatunk: pl. szinonímiaviszonyokat (*autó – kocs*), vagy hiponímiaviszonyokat (pl. *bútor – asztal*), amelyeket aztán automatizált gépi eljárással úgynevezett *jelentésrendekbe* (*synset*) lehet foglalni. Ezeknek a jelentésrendeknek a feltárása megfelelő nagyságú szöveggörpuszokban jelenti a vizsgálódások empirikus alapját, hiszen olyan rendezett adatsorokat kapunk, amelyek lehetőséget adnak a szövegek lexikai kohéziós mintázatainak azonosítására, a fordított valamint az autentikus szövegek lexikális eszköztára közötti eltérések feltárására, továbbá általánosítások megfogalmazására a forrásnyelvi szövegek mintázatainak célnyelvi jelenlétének súlyáról.

A korpusznyelvészeti elemzés szövegnyelvészeti támpillérét a kötet *harmadik része* biztosítja, amelyben a szerző a korábban is alkalmazott tudománytörténeti rekonstrukcióval tárja elénk a szövegkutatás főbb állomásait. A fókuszban a fordítástudomány és szövegnyelvészet kapcsolata áll, hiszen, miként Klaudy (2007, p. 139) egyik írására hivatkozva Seidl-Péché megállapítja „a fordítástudomány a fordítási feladatok természetéből adódóan, azaz a mondathatárokon túlmutató szövegegységek 'átkódolása' miatt nem nélkülözheti a szövegszintű vizsgálatokat” (Seidl-Péché, 2024, p. 39). Az első három történeti fejezet vezet el bennünket a negyedik fejezethez, amelyben a szerző Petőfi S. János és Szikszainé Nagy Irma munkáira támaszkodva határozza meg a szöveg fogalmát mint ami a „nyelvhasználat legmagasabb szintjén szerveződött legkomplexebb nyelvi jelcsoport” (Seidl-Péché, 2024, p. 41).

Mivel a szöveg mint a kutatási tárgya meghatározása nem nélkülözheti szövegszerűség kritériumait, ezért természetesen ezek is terítékre kerülnek. A szövegszerűség kritériumai közül kiemelendő a kohézió fogalma, mellyel a *negyedik rész* foglalkozik *A lexikai kohézió vizsgálata* címmel. A negyedik rész fejezetei áttekintik a kohéziókutatás legfőbb résztvevőinek kutatásait, köztük Halliday és Hasan műveit, de Hatim és Mason, Hoey, Károly Krisztina, valamint Blum-Kulka munkássága is viszonylag nagy teret kapnak.

A szövegnyelvészeti pillért kiépítő negyedik részt követi annak a korpusznyelvészeti módszernek a bemutatása, amelyet Seidl-Péché a későbbi részekben bemutatandó kis- és nagymintás korpusznyelvészeti vizsgálódásaiban alkalmaz. A szövegnyelvészeti részhez hasonlóan, a szerző az ötödik részben is tudománytörténeti kontextusba ágyazottan ismerteti a korpuszok legalapvetőbb ismérveit, majd vezeti be a saját maga fejlesztette korpusznyelvészeti módszerét. Ennek első lépése a kutatások alapját jelentő vizsgálati korpusz összeállítása (4.3. alfejezet), melyet a korpusz szövegeinek annotálásához használt modell, a WordNet formalizmusainak bemutatása, valamint alkalmazhatóságának indoklása követ.

A *hatodik rész* a kötet „laboratóriumi” része. Egy pilotkutatás keretében teszteli Seidl-Péché az általa kifejlesztett korpusznyelvészeti eljárás alkalmazhatóságát egy kézzel is annotálható kis-mintás, tematikusan egységes, ún. leíró jellegű szövegeket tartalmazó német-magyar párhuzamos szövegkorpuszra. A szövegkorpusz feldolgozásának elméleti keretét a Károly Krisztina kidolgozta lexikai ismétlődés taxonómiája képezi, amely kétféle ismétlődési típust különít el egymástól: (1) a lexikai és a (2) szövegalapú relációkat. A lexikai ismétlődés két altípusra tagolódik: (1.1) a lexikai egység változatlan formában és jelentésben ismétlődik, illetve (1.2) a lexikai egység hordozta jelentés teljesen vagy részben felidéződik egy vele valamilyen szemantikai relációban (szinonímia, ellentét, hiponímia, meronímia) álló másik lexikai egységgel. Seidl-Péché ezt a taxonómiát beépítette egy nyelvtechnológiai elemző rendszerbe (CLaRK system), és így lehetővé vált a mondathatárok jelölése, az ismétlődő lexikai elemek, valamint az ismétlés mint kohéziós típus jelölése is. A vizsgálat eredményeként beigazolódott, hogy a fordítás befolyásolja a célnyelvi lexikai kohéziót, ugyanis „a fordított német szövegekben nagyobb a lexikai elemek ismétlésének sűrűsége, mint amilyen az eredeti magyar szövegekben” (Seidl-Péché, 2024, p. 75).

Ennek a kismintás vizsgálatnak a tanulságait felhasználva végezte el a szerző azt a nagymintás kutatást, amelyet a kötet *hetedik része* mutat be *A Magyar Lexikai Kohéziós Projekt* címmel. A vizsgált korpuszt egy több mint 3 millió szövegszóból álló fordított magyar nyelvű adatbázis képezte, amelyet automatizált gépi feldolgozással vizsgált a szerző, tesztelve többek között a HuWordNet-ontológia működését is. A projekt során a szerző gondosan dokumentálja a korpuszpépítés egyes lépéseit, a szövegek kiválasztási kritériumainak rögzítésétől kezdve a nyelvek megválasztásáig. Ebben a projektben a pilotkutatástól eltérően, nem egyetlen idegen nyelvből (németből) fordított szövegek kapnak helyet, hanem összesen hat nyelvből (angol, francia, német, olasz, latin, ógörög) fordított szövegek alkotják a vizsgált tematikus alkorpuszokat (közéleti, szépirodalmi, vallási, tudományos). Az ugyanilyen tematikus tagolású autentikus magyar szövegek korpusza kicsit több mint 1 millió szövegszót fogott át. A szerzőnek az így összeállított szöveggörpusz vizsgálatával lehetősége nyílt arra is, hogy nyelvpárfüggetlen, általános következtetéseket fogalmazzon meg a lexikai kohézió eltolódására nézve fordított szövegekben, azaz, hogy empirikusan is megalapozott hipotézist állítson fel arról, hogy a fordítás során a lexikai kohézióban létrejövő eltolódás egy fordítási univerzálé.

A hetedik rész negyedik fejezete mutatja be a korpusz annotálását, amelyet nyelvtchnológiai eszközökre támaszkodva valósít meg a szerző. Az OpenOffice segítségével a szerző XML-dokumentumokat hozott létre, amelyeket egy elemző program dolgozott fel, azonosítva a mondathatárokat és a tokeneket. A HuMor-program végezte el a morfológiai annotációt, az egyértelműsítést, és keresi meg a főnevek ismétléses előfordulásait, és „azonosítja a mondathatárokon túlmutató – szinonímia, antonímia, meronímia, hiponímia vagy doménreláció alapultó – szemantikai kapcsolatokat” (Seidl-Péché, 2024, p. 87).

A Magyar Lexikai Kohéziós Projekt nagymintás vizsgálata nem pusztán a vizsgált korpusz nagysága okán innovatív vállalkozás, hanem azért is, mivel a módszertana validált, alkalmazhatósága pedig további eredményes kutatások kiindulópontja lehet. A hatodik fejezet összegzései alapján az a következtetés fogalmazható meg, hogy minden nyelv esetében megváltozik a fordított szövegek lexikai kohéziós mintázata. A szövegtípusok felől nézve szintén megerősítést nyer az a hipotézis, hogy az egyes nyelvekre jellemző lexikai kohéziós eszközhasználat eltérő módon befolyásolja a célnyelvi szövegalkotást a lexikai kohéziós eszközhasználat tekintetében (Seidl-Péché, 2024, po. 98-107).

Az egyes nyelvek és szövegtípusok mátrixában azonban nagyfokú eltérések detektálhatók a lexikai kohéziós eszközök tekintetében. Csúpn egy példát kiragadva: A tudományos szövegek esetében a lexikális ismétlések száma (R, ismétlések) az angolról és az olaszból fordított szövegekben magasabb, míg a franciából és a németből fordítottaké esetében kevesebb, mint az 1000 autentikus magyar szövegszóra vetített lexikális ismétlések száma. Ugyanakkor a szinonímia használata (S) esetében teljesen más kép rajzolódik ki: az 1000 autentikus magyar szövegszóra vetített előfordulások száma az olaszból fordított szövegekben előforduló szinonímiák számához képest alacsony, míg a többi vizsgált nyelv közül a francia és a német szövegekben ez az előfordulása közel azonos, az angol szövegekben pedig nagyon alacsony. Ha pedig csak a lexikai ismétléseket vesszük górcső alá és csak az angolból fordított szövegeket vesszük, akkor

a különböző szövegtípusok esetében nagy eltérésekkel találkozunk: az 1000 autentikus magyar szövegszóra vetített előfordulások száma irodalmi szövegek esetében közel azonos, a közéleti szövegeké kicsit alacsonyabb, viszont a tudományos szövegek esetében jóval magasabb, vallási fordított szövegekben pedig jóval alacsonyabb.

Ezek az eltérések, amelyek minden nyelv és szövegtípus esetén megfigyelhetők, egyfelől valóban arra engednek következtetni, hogy a fordított szövegek lexikai kohéziós mintázata eltér az autentikus, tehát nem fordítási céllal készült szövegek lexikai kohéziós mintázataitól. Másfelől viszont az adatokból úgy tűnik, hogy nem rajzolódnak ki általános mintázatok, amelyek nyelvtől és szövegtípustól függetlenül jellemeznének egy-egy lexikai kohéziós eszközt (az ismétlést, a szinonímiát, az antonímiát stb.). Ezért a szerző tartózkodik is attól, hogy általános, ám esetleg leegyszerűsítő magyarázatokkal álljon elő a lexikai kohéziós eltolódások eltérő intenzitásának okait firtatva. Mindazonáltal az ennyire markáns eltérések annak a hipotézisnek a megfogalmazására véleményem szerint elegendőek lennének, mely szerint a lexikai kohéziós mintázatok nagyfokú diverzitásáért nem lehetnek felelősek a nyelvi rendszerek közötti eltérések. A szerző ennél óvatosabb, az eredményeivel megnyitja az utat a további vizsgálódások felé is, amennyiben az egyes nyelvpárok és szövegtípusok esetében arra a következtetésre juttatja el az olvasót, hogy minden egyes szövegtípus és nyelvpár esetén külön megtervezett és nagymintás vizsgálódások szükségesek a lexikai kohéziós eszköztár használati mintázatainak feltárásához.

A kötetet záró összegző, *nyolcadik rész* rekapitulálja, és pontokba szedve ismerteti az eredményeket, és a lehetséges jövőbeni kutatási irányok felvázolásával zárul. A kötet gondolati, érvelési struktúrája jól követhető, az elemzések pontosak, megértésüket áttekinthető táblázatok is segítik. Seidl-Pécs könyve olyan átfogó munka a fordított szövegek lexikai kohéziós mintázatáról, amely egyúttal elméleti bevezetést és gyakorlati útmutatást nyújt a fordításszövegek kis- és nagymintás elemzésének fordítástudományi és korpusznyelvészeti háttéréről. Eszmefuttatásai érintik a fordítás minőségének értékelését, segítik a fordítóképzést, és támpontokat adnak a gépi fordítás fejlesztéséhez is. A kötet szépen szemlélteti, hogy a fordítás lényegi eleme a kohéziós hálózat célnyelvi újratерemtése, amely mindig kreatív, és kontextusfüggő folyamat.

IRODALOM

- Balaskó, M. (2007). A fordításnyelvről, avagy a flamand szőnyeg láthatatlan szálairól. In P. Heltai (Ed.), *Nyelvi modernizáció: Szaknyelv, fordítás, terminológia* (pp. 159–166). MANYE–Szent István Egyetem.
- Heltai, P. (2009). Fordítás, relevancia, feldolgozás. *Modern Nyelvoktatás*, 15(1–2), 13–28.
- Klaudy, K. (2007). *Nyelv és fordítás: Válogatott fordítástudományi tanulmányok*. Tinta Könyvkiadó.
- Szép, B. (2009). Adalékok a magyar jogi szaknyelv kialakulásának történetéhez. *Magyar Nyelvőr*, 133. (pp. 310–332.).