

Eötvös Loránd Tudományegyetem  
Bölcsészettudományi Kar

DOKTORI DISSZERTÁCIÓ TÉZISEI

**BESZNYÁK RITA**

**GYAKORLÓBESZÉDEK GRADUÁLÁSA A TOLMÁCSKÉPZÉSBEN  
KORPUSZNYELVÉSZETI ESZKÖZÖK BEVONÁSÁVAL**

DOI: <https://doi.org/10.15476/ELTE.2023.272>

Nyelvtudományi Doktori Iskola  
Vezető: Dr. Tolcsvai Nagy Gábor akadémikus, egyetemi tanár

Fordítástudományi Doktori Program  
Vezető: Dr. Klaudy Kinga DSc, professor emeritus

**A bizottság tagjai:**

Elnök: Prof. Dr. Klaudy Kinga DSc, ELTE FTT professzor emeritus

Bírálok: Dr. Eszenyi Réka, PhD, ELTE FTT egyetemi adjunktus  
Dr. Bakti Mária, PhD, SZTE MNYKT főiskolai docens

Titkár: Dr. Zachar Viktor PhD, ELTE FTT egyetemi adjunktus

Tagok: Dr. Seidl-Pécs Olívia PhD, BMGE egyetemi docens  
Dr. Szegh Henriett, PhD  
Dr. Robin Edina, PhD, ELTE egyetemi adjunktus

**Témavezető:**  
Dr. Seresi Márta, PhD, ELTE FTT egyetemi adjunktus

Budapest  
2023

## 1. Bevezetés

Akinek nyílt már alkalma belelátni a tolmácsok munkájába – akár gyakorló nyelvi közvetítőként, akár felhasználói oldalról vagy kutatói perspektívából – az tudja, hogy igen összetett, komoly szellemi kihívást jelentő nyelvi átváltási feladatról van szó, mely magasszintű szakmai ismereteket és előképzettséget igényel. A tolmácsok felkészültségének és professzionalitásának a záloga többek között a tolmácsképző intézmények gondosan kidolgozott képzési struktúrájában rejlik. Ahhoz, hogy a hallgatók a képzőintézmény falait elhagyva felkészülten nézhessenek szembe a tolmácsolási feladatok során felmerülő nehézségekkel, a tolmácsórákon fokozatosan újabb és újabb kihívások elé állítják őket az oktatóik. Ez a fajta lépcsőzetes építkezés biztosítja, hogy az elsajátított ismeretek újabb és újabb „téglaiból” stabil tudásanyag épüljön. A tolmácsképzés keretein belül ez a fokozatosság több szinten is tetten érhető: tanmenetek szintjén abban, ahogy az egyes tolmácsolási módok oktatása egymásra épül, tananyagok szintjén pedig abban, ahogy egyre nő a tolmácsórákon feldolgozott gyakorlóanyagok nehézségi szintje.

A tolmácsképzésben használt gyakorlóbeszéd kiválasztásakor és összeállításakor az oktatóknak számos szempontot kell mérlegelniük. A fokozatos építkezés érdekében mindig azt kell szem előtt tartaniuk, hogy a tolmácsolandó beszéd megfeleljen a hallgatók aktuális tudásszintjének a képzés egyes szakaszaiban. A graduálásnak, azaz a nehézségi szint fokozatos növelésének a célja egyrészt az, hogy a gyakorlóanyagok az elsajátított tolmácsolási ismeretek és megküzdési stratégiák fényében teljesíthető kihívást jelentsenek az órákon, másrészt az, hogy felkészítsék a hallgatókat a képzést lezáró kimeneti vizsgára. A beszéd nehézségi szintjének meghatározásához és fokozatos növeléséhez az oktatóknak tisztában kell lenniük azzal, hogy milyen típusú beszéd alkalmas az egyes tolmácsolási módok gyakorlására, és ismerniük kell azokat a nyelvi és nyelven kívüli változókat, melyek megnehezíthetnek egy tolmácsolási feladatot. Be kell tudniuk azonosítani azokat a potenciális problémaforrásokat és buktatókat, melyek egy beszéden belül gondot okozhatnak a hallgatóknak. Legyen szó a tolmácsolást megnehezítő külső körülményekről, a beszélő előadásmódjából adódó nehézségekről, esetleg nehezen megfeleltethető vagy átváltható lexikai elemekről.

A tolmácsolástudományi szakirodalmi forrásokban bőven találunk arra vonatkozó útmutatást, hogy mely tényezők (ún. input változók) befolyásolják egy tolmácsolási feladat nehézségét (Gile 1995/2009, Pöchhacker 2004/2016, Mead 2015 stb.). Módszertani ajánlásokat is olvashatunk a gyakorlóanyagok graduálására vonatkozóan (G. Láng 2002, Setton–Dawrant 2016 stb.), melyeket számos tolmácsképző intézmény figyelembe vesz a képzési struktúra

kialakításakor. Ugyanakkor a napi, tolmácsórai gyakorlatban a tolmácstanárok – a vonatkozó szakirodalom ismeretében és gyakorlati tapasztalataikra támaszkodva – általában szubjektív szempontok alapján, holisztikus megközelítésben igyekeznek meghatározni az órákon feldolgozott gyakorlóanyagok nehézségi szintjét. Ezt a graduálási folyamatot jelentős mértékben megkönnyítené, ha rendelkezésre állna valamilyen objektív mérőeszköz a beszédek nehézségi szintjének összehasonlítására és „szintezésére”, illetve az esetleges tolmácsolási csapdák beazonosítására.

Érdeemes tehát megvizsgálni, hogy vajon alkalmazhatóak-e más (tudomány)területek (neveléstudomány, nyelvoktatás, lapkiadás, honlapszerkesztés) eredményei és az írott vagy hangzó szövegek feldolgozási nehézségének mérésére használt eszközei (Flesch 1974, Klare 1984, DuBay 2006) a tolmácsbeszédek nehézségi szintjének besorolására is. A beszédnehézség szempontjából releváns változók analitikus szemléletű vizsgálatához ma már számos számítógépes elemző eszköz áll rendelkezésre. A korpuszalapú tolmácsolástudomány fejlődését és kutatási eredményeit látva (Pöchhacker 1994, Shlesinger 1995, Bakti 2010, Russo 2021, Szegh 2021 stb.) joggal feltételezhetjük, hogy a beszédnehézség meghatározása és graduálása szempontjából is előremutató eredményeket hozhat egy, a gyakorlóbeszédekből összeállított korpusz számítógépes vizsgálata.

Jelen disszertáció a szakirodalmi előzményekre és a korpuszalapú tolmácsolástudomány kutatási módszereire építve vizsgálja a tolmácsórákon használt beszédek nehézségét. A kutatás relevanciáját és újszerűségét az adja, hogy egy eddig kevésbé kutatott korpuszt vizsgál számítógépes szövegelemző eszközökkel, korábban nem alkalmazott fókusszal és részletességgel. A disszertáció a számítógépes szövegelemzés segítségével mérhető objektív beszédnehézséget a hallgatók tolmácsolási teljesítményéből leszűrhető szubjektív beszédnehézség fényében is vizsgálja. Ezzel a vegyes, kvantitatív és kvalitatív elemeket is tartalmazó kutatással átfogó és reális képet kaphatunk a vizsgált beszédek tolmácsolási nehézségéről. A korpuszelemzésből nyert kutatási eredmények hozzájárulhatnak ahhoz, hogy még szisztematikusabb legyen a gyakorlóanyagok graduálása a tolmácsképzésben.

## **2. A kutatás célja**

A kutatás elsődleges célja annak feltérképezése, hogy a számítógépes szövegelemző programok alkalmazhatóak-e a tolmácsképzésben a gyakorlóbeszédek nehézségének meghatározására és graduálására. Ehhez az elsődleges kutatási célhoz kapcsolódóan az alábbi másodlagos kutatási célok határozzák meg a kutatás súlypontjait és menetét:

1. A tolmácsolási feladat nehézségét, és ezáltal a tolmácsolás minőségét befolyásoló külső és belső tényezők rögzítése a szakirodalmi források alapján. Az írott és hangzó szövegek feldolgozási nehézségének mérésére használt módszerek feltérképezése. A korpuszalapú tolmácsolástudomány eddigi kutatási céljainak és eredményeinek rövid áttekintése.
2. A tolmácsképzésben alkalmazott holisztikus szemléletű beszédnehézség-besorolási és graduálási módszerek összegzése, a vonatkozó módszertani ajánlások ismertetése. A szakirodalmi forrásokban megjelölt beszédnehézségi paraméterek összevetése a tolmácsképzésben használt gyakorlóanyagokból összeállított, graduált korpuszon. Azoknak a beszédnehézséget befolyásoló szövegen belüli tényezőknek a beazonosítása, melyeket lehet és érdemes számítógépes korpusznyelvészeti eszközökkel kutatni.
3. A tolmácsórákon használt beszédek analitikus szemléletű elemzése számítógépes szövegelemző programok segítségével. A nehézségi szint meghatározása szempontjából releváns szövegjellemzők beazonosítása és összehasonlító vizsgálata. A szövegnehézség és olvashatóság mérésére használt módszerek alkalmazhatóságának tesztelése egy graduált korpuszon. A kiválasztott gyakorlóbeszédek lexikai profiljának elemzése, a feltételezhetően tolmácsolási nehézséget jelentő lexikai elemek beazonosítása szókinccs-elemző számítógépes programokkal.
4. A konsekutív és szinkrontolmácsolás gyakorlására szánt beszédek közötti eltérések feltérképezése a tolmácsolási nehézség szempontjából releváns szövegjellemzők és lexikai elemek fényében. A két tolmácsolási mód között megfigyelhető esetleges graduálási asszimmetriák beazonosítása.
5. A feltételezett nehézségi szint és a számítógépes szövegelemzéssel beazonosított tolmácsolási buktatók validitásának tesztelése a tolmács hallgatók tolmácsolási teljesítményének fényében. Az objektív és szubjektív beszédnehézség közötti esetleges eltérések számbavétele.
6. A tolmácsképzésben használt gyakorlóbeszédek nehézségi szintjének fokozatos növelésére vonatkozó ajánlások megfogalmazása a korpusznyelvészeti elemzés alapján leszűrhető tanulságok fényében, különös tekintettel a konsekutív és szinkronbeszédek differenciált graduálására.

### 3. Kutatási kérdések

A kutatás fent megjelölt céljaihoz igazodva az alábbi kutatási kérdésekre keresem a választ a disszertációmban:

1. Tolmácsolási szempontból mi számít nehezítő tényezőnek a nyelvi átváltás során? Milyen szövegen kívüli és szövegen belüli tényezők befolyásolhatják egy tolmácsolási feladat nehézségi szintjét?
2. Milyen alapelvek mentén, mely szempontok figyelembevételével érdemes graduálni az egyes tolmácsolási módok gyakorlására szánt anyagokat? Milyen változókat kell figyelembe venni a tolmácsolási célra szánt gyakorlóbeszéd nehézségi szintjének meghatározásakor?
3. Milyen eszközökkel, milyen mutatószámok alapján mérhető egy tolmácsolási gyakorlóbeszéd nehézségi szintje? Mennyiben alkalmasak a szövegnehézségi szintek mérésére használt számítógépes korpusznyelvészeti elemzési módszerek a tolmácsolás-képzésben használt gyakorlóbeszéd nehézségének meghatározására és graduálására, illetve a szakirodalmi ajánlások besorolási rendszereinek megerősítésére?
4. Vannak-e szisztematikus eltérések a beszéd nehézségét meghatározó tényezők tekintetében az egyes tolmácsolási módok között? Melyek azok a szövegjellemzők vagy lexikai elemek, amelyek a konsekutív gyakorlóbeszédekben nagyobb nehézséget jelentenek, mint egy szinkrontolmácsolásra szánt beszédben és fordítva?
5. A tolmácsolási teljesítmény fényében melyek azok a szövegen belüli tényezők, amelyek nehézséget okoznak a hallgatóknak a gyakorlóbeszédnél? Mennyire esnek egybe a tolmácsolás-felvétel értékelésekor detektálható hibák és bizonytalanságok az előzetes kvantitatív elemzésnél beazonosított tolmácsolási csapdákkal? Milyen eltérések tapasztalhatóak a vélt és valós nehézségek tekintetében a konsekutív és szinkronbeszéd között?
6. Hogyan használhatóak a számítógépes szövegelemző programok a tolmácsolás-képzés egyes szakaszaiban a beszéd nehézségi szintjének meghatározására? Alkalmazhatóak-e a korpusznyelvészeti eszközök a két tolmácsolási mód gyakorlására szánt beszéd differenciált graduálására?

#### 4. Hipotézisek

A fent megfogalmazott kutatási kérdésekhez kapcsolódóan eddigi tapasztalataim és ismereteim alapján az alábbi előzetes feltevésekből indulok ki:

1. Az első két kutatási kérdésre a szakirodalmi forrásokat és módszertani ajánlásokat áttekintve fogunk részletes választ kapni. A korábbi, témába vágó kutatások eredményeiből feltehetően kiderül majd, hogy egy tolmácsolási feladat nehézségének megítélését nagyon sok változó körülmény befolyásolja, s mindig figyelembe kell venni a szubjektív tényezőket. Ahhoz, hogy egy ilyen fókuszú kutatás valid eredményeket hozzon, le kell szűkíteni a beszédnehézséget meghatározó változók körét. Olyan korpuszt érdemes vizsgálni, ahol a szövegen kívüli változók (hangerő, akcentus, tempó stb.) aránylag állandóak és számítógépes korpusznyelvészeti módszerekkel jól kutathatóak.
2. Egy gondosan összeállított, egységes korpusznál a számítógépes korpusznyelvészeti elemzési módszerek olyan kutatási eredményeket hoznak majd, melyek alátámasztják a tolmácsolásban alkalmazott, holisztikus megközelítésű graduálási szempontokat. A beszédek alapvető szövegjellemzőiben, szintaktikai szerkezetében, lexikai profiljában nagy valószínűséggel lehet majd azonosítani olyan tendenciákat, melyek megerősítik a fokozatosan növekvő tolmácsolási nehézséget. A tolmácsolási nehézség és a szövegszinten mérhető olvashatósági mutatók vélhetően bizonyos szintű korrelációt mutatnak majd.
3. Ugyanakkor az is valószínűsíthető, hogy találunk majd olyan kivételeket, ahol a számítógépes szövegelemző programok és az olvashatósági indexek alapján feltételezett beszédnehézség nem mutat egyértelműen növekvő tendenciát. Jó eséllyel lesznek olyan beszédek is, ahol az előzetesen korpusznyelvészeti eszközökkel beazonosított potenciális buktatók nem növelik szükségszerűen a tolmácsolási feladat nehézségét. Itt érdemes lesz megvizsgálni, hogy milyen egyéb tényezők játszanak közre a beszéd tolmácsolási nehézségében.
4. Feltételezhető, hogy korpusznyelvészeti elemzési módszerekkel is kimutatható eltérések lesznek a konsekutív és a szinkrontolmácsolási célra szánt beszédekben fellelhető nehezítő tényezők között. Miután a konsekutív tolmácsolásnál a hallgatók jegyzetből dolgoznak, és valamivel több idejük van az elhangzott információ feldolgozására, mint a

szinkrontolmácsolási feladatokban, a vizsgált szövegjellemzők tekintetében és a tolmácsolási csapdák arányában magasabb értékekre számítok a konszekutív gyakorlóanyagoknál. Emellett azt is valószínűnek tartom, hogy a konszekutív tolmácsolási feladatok során jelentkező szövegen belüli nehézségek kiszámíthatóbbak, számítógépes elemzési módszerekkel jobban megjósolhatók lesznek, mint a szinkronbeszédekénél.

5. A szövegelemzéssel előzetesen azonosított, feltételezett tolmácsolási problémák (objektív beszédnehézség) feltehetően nem esnek majd mindig egybe a hallgatók tolmácsolási teljesítménye alapján leszűrhető nehezítő tényezőkkel (szubjektív beszédnehézség). A kiszűrt tolmácsolási csapdák többsége jó eséllyel valóban kihívás elé állítja majd a hallgatókat, de ez nem mindenkinél vezet szükségszerűen hibákhoz. Ahhoz, hogy teljesebb képet kapjunk az adott beszéden belüli nehezítő tényezőkről, mindenképpen több hallgató tolmácsolási teljesítményét kell párhuzamosan megvizsgálnunk.
6. Összességében abból a hipotézisből indulok ki, hogy a tolmácsolásben is hasznos tanulságokkal szolgálhatnak, és értékes – ha nem is kizárólagos – beszédnehézség-mérő eszköznek bizonyulhatnak a számítógépes korpusznyelvészeti elemző programok. Alkalmazásukkal szisztematikusabb lehet a beszédek graduálása, az oktatók könnyebben tudják a hallgatók aktuális tudásszintjéhez igazítani és fokozatosan növelni a feldolgozott gyakorlóanyagok nehézségi szintjét.

## **5. A kutatás menete**

A kutatási célokhoz és kérdésekhez igazodva a disszertáció a vonatkozó szakirodalmi források és kutatási előzmények áttekintésével indul. Először a szakirodalmi források alapján ismertetem azokat a tényezőket – ún. input változókat –, melyek befolyásolhatják egy tolmácsolási feladat nehézségét. Röviden áttekintem a szakirodalmi forrásokban megjelenő beszédnehézség besorolási rendszereket (Gile, 1995/2009, Pöchhacker 2004/2016, Setton 2009, Mead 2015), majd összesítem a korábbi kutatásokban beazonosított lehetséges problémaforrásokat és tolmácsolási csapdákat (Gerver 1971/2002, Déjean le Féal 1982, Viaggio 1996, G. Láng 2002, Albl-Mikasa 2013). A nehezítő körülmények áttekintését követően összegzem a gyakorlóanyagok graduálására vonatkozó főbb szakmai ajánlásokat, melyeket a tolmácsolástatók a gyakorlóbeszédek nehézségi szintjének fokozatos növelésénél figyelembe vesznek (Jones 1998, Alexieva 1999, Seleskovitch–Lederer 2002, Hönig 2002, G. Láng 2002, Liu–Chiu 2009,

Setton–Dawrant 2016). Kitérek a különböző tolmácsolási módok gyakorlására szánt beszédek közötti különbségekre is.

Ezt követően rátérek arra, hogy egyéb (tudomány)területeken milyen eszközöket és mérőszámokat használnak a szövegnehézség meghatározására. Kiemelt figyelmet szentelek az olvashatóság mérésére használt formuláknak és az összeállításuknál figyelembe vett szövegjellemzőknek (Flesch 1948, 1974, Gunning 1952, Klare 1963, Darville–Hiebert 1985, DuBay 2004), valamint a hallgathatóság növelésére vonatkozó útmutatóknak (Harwood 1955, Rubin–Raforth 1986, Brindley–Slatyer 2002, Bodie 2018). A társtudományok ezirányú kutatási eredményeinek rövid áttekintése után térek ki a kutatás módszertani előzményeire, a korpuszalapú tolmácsolástudomány eddigi kutatási irányaira. Jelen kutatás fókuszának leszűkítéséhez és módszertani megalapozásához rendszerezem az eddigi korpuszalapú tolmácsolási kutatásokat, és megvizsgálom, hogy az eddigi eredmények mennyiben használhatóak a beszédnehézség meghatározására (Déjean le Féal 1978, Pöchhacker 1994, Shlesinger 1995, Kalina 1998, Monacelli 2005, Klaudy 2005, Russo et al. 2006, Bakti 2008, 2010, Robin–Seidl–Péché 2020, Szegh 2021 stb.). A korpuszalapú tolmácsolástudomány fejlődési szakaszainak bemutatása előtt röviden kitérek az ilyen típusú kutatások korlátaira is.

A jelen kutatásban vizsgált forrásnyelvi korpusz az Európai Bizottság Tolmácsolási Főigazgatóságának gyakorlási célra szánt beszédgyűjteményéből (*SCIC Speech Repository*) származik. 32 angol beszéd leírataiból áll, ezek fele a konsekutív tolmácsolás gyakorlására szánt felvétel, a másik fele szinkron gyakorlóbeszéd. A kutatás első, kvalitatív szakaszában a forrásnyelvi korpuszban szereplő beszédek tolmácsolási nehézségét holisztikus megközelítésben vizsgálom. Három forrásra támaszkodva elemzem a beszédek nehézségét: (1) a Speech Repository saját besorolási rendszere (*SCIC Speech Repository grading criteria*); (2) Setton és Dawrant beszédnehézségi indexe (*Speech Difficulty Index*) (Setton–Dawrant 2016), valamint (3) G. Láng beszédnehézségi paraméterei (G. Láng 2002) alapján.

A kutatás második, kvantitatív szakaszában azt tesztelem a vizsgált forrásnyelvi korpuszon, hogy a számítógépes korpusznyelvészeti elemzési módszerek alkalmazhatóak-e a gyakorlóbeszéd nehézségi szintjének meghatározására, illetve a fenti holisztikus megközelítésű graduálási paraméterek mérésére. A forrásnyelvi korpusz beszédeinek átiratát először a TextInspector szövegelemző programmal, majd az Oxford Text Checker szókincs-elemző eszköz segítségével elemzem. A számítógépes korpuszelemzés első fázisában a TextInspector programmal a szövegre vonatkozó alapvető statisztikai adatokat tekintem át, illetve az olvashatósági képletekkel mérhető szövegnehézségi értékeket. Ezt követi a korpuszban szereplő beszédek lexikai elemeinek szófaji elemzése, majd a szövegen belüli



logikai összefüggések feltérképezését, a szöveg értelmezését segítő ún. metadiskurzus-jelölők összehasonlító vizsgálata. A TextInspector lexikai-profil elemző funkciójával a vizsgált szövegjellemzők alapján összesítem a betáplált beszéd-leiratok feltételezhető feldolgozási nehézségét. A kvantitatív elemzés elsődleges célja valamennyi vizsgált szövegjellemzőnél annak feltárása, hogy megfigyelhetőek-e növekvő vagy csökkenő tendenciák az eltérő besorolású beszédek között. A graduált beszédek számítógépes szövegnyelvészeti elemzésénél külön vizsgálom a konszekutív és a szinkronbeszédekből álló alkorpuszokat, az esetleges szisztematikus eltérések beazonosítása érdekében.

A beszédek lexikai profiljának mélyebb feltárására szolgál a kvantitatív kutatás lezárásaként végzett szókincs-elemzés az Oxford Text Checker programmal. A számítógépes korpuszelemzés célja itt elsősorban azoknak a lexikai elemeknek a beazonosítása, melyek az szakirodalmi források alapján (vö. Setton – Dawrant 2016) tolmácsolási nehézséget jelentenek, potenciális „tolmácsolási csapdának” (*interpreting pitfalls*) számítanak. Ide tartoznak például olyan tényyszerű információk, mint a számok, tulajdonnevek (személynevek, földrajzi nevek, intézménynevek), vagy az adott szakterületen speciális jelentéstartalommal használt terminusok. Az Oxford Text Checker szövegelemző program segítségével felmérem a forrásnyelvi korpuszban szereplő beszéd-leiratok lexikai készletét, és megvizsgálom az egyes gyakorlóanyagok becsült szövegértési nehézséget. Itt is külön kitérek a konszekutív és szinkronbeszédek közötti esetleges különbségekre.

A kutatás harmadik szakaszában a szubjektív beszédnehézséget vizsgálom a tolmácsok szemszögéből, egész pontosan a hallgatók tolmácsolási teljesítménye alapján. A vizsgált forrásnyelvi korpusz beszédeiről készült hallgató hangfelvételekben elemzem, hogy beigazolódna-e a gyakorlatban a korábban analitikus módszerekkel beazonosított, feltételezhető tolmácsolási nehézségek. Összevetem 16, a forrásnyelvi korpuszból kiválasztott, fokozatosan növekvő nehézségű beszéd tolmácsolt változatait, mindkét tolmácsolási módban 8-8 beszédét. A longitudinális kutatásban a BME Tolmács- és Fordítóképző Központ hallgatóinak tolmácsolásairól készült hangfelvételeket, illetve azok leiratait elemzem (beszédenként 5-5 tolmácsolat változatot). A kutatásnak ebben a szakaszában a tolmácsolás során vétett tartalmi és formai hibák alapján vonok le következtetéseket a tolmácsolási feladat során jelentkező nehézségekre vonatkozóan. Elsősorban a korábban számítógépes eszközökkel beazonosított tolmácsolási csapdákra összpontosítok, és külön is kitérek a konszekutív és a szinkrontolmácsolás gyakorlására szánt beszédek között tapasztalható különbségekre.

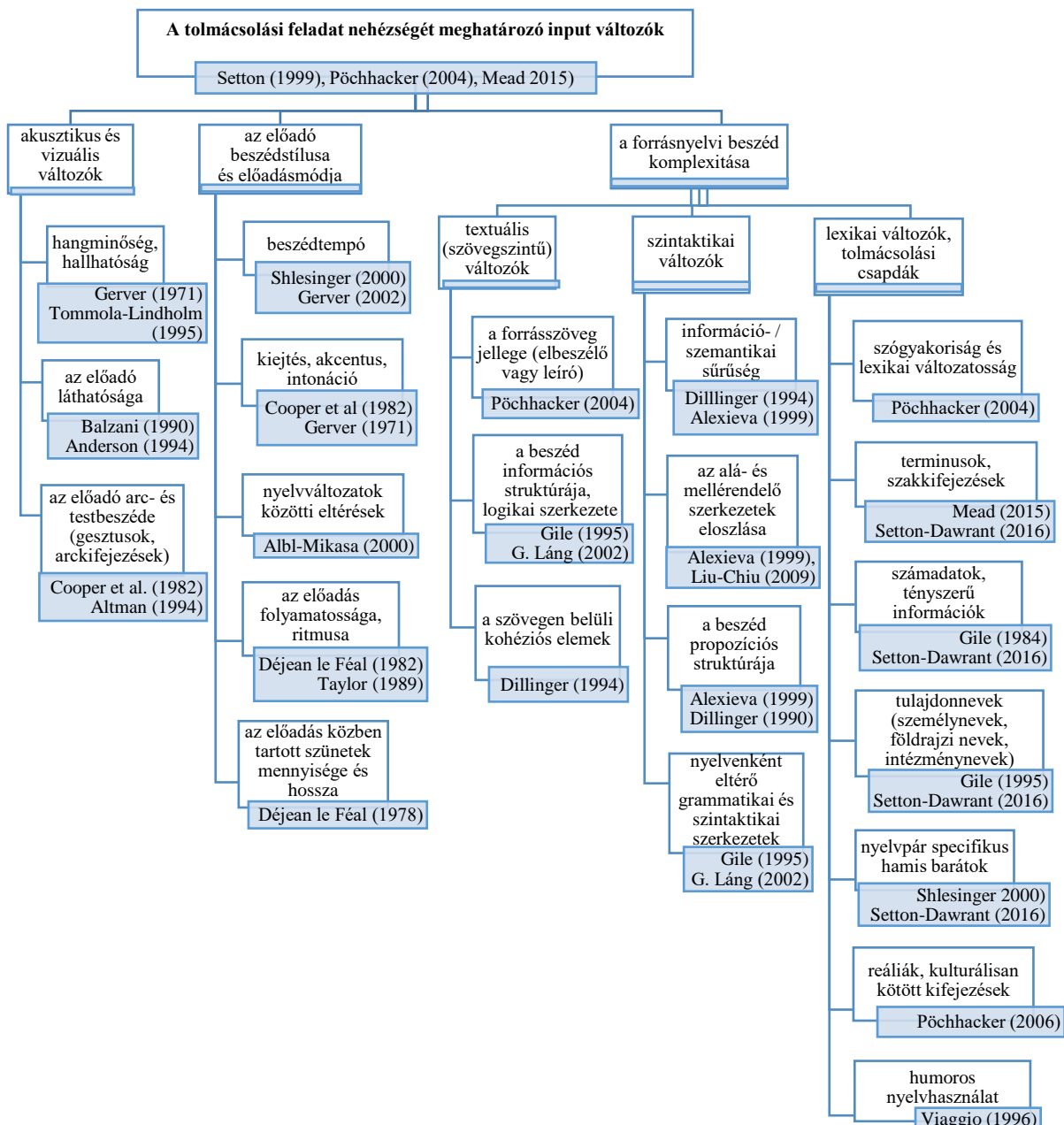
## 6. A kutatási eredmények összegzése

A kutatás eredményeit a disszertáció elején megfogalmazott kutatási kérdésekre visszacsatolva ismertetem, és kitérek az előzetesen megfogalmazott hipotézisekre is.

### 6.1. A tolmácsolási feladat nehézségét meghatározó szövegben belüli és kívüli tényezők

A tolmácsolás nehézségét befolyásoló input változókat a szakirodalmi források alapján az 1. ábrán összesítettem. A csoportosításnál Pöchacker (2004), illetve Setton (1999) és Mead (2016) rendszereit vettem alapul, és jelöltem az egyes résztémakörökkel foglalkozó kutatókat.

1. ábra. A tolmácsolási feladat nehézségét meghatározó input változók



A tolmácsolásban a beszédek nehézségét elsősorban ezeknek az input változóknak a figyelembevételével tudjuk mérni és irányítani. Úgy is mondhatnánk, a graduálás kulcsa a tolmácsoláson az input változók elemzésében és tudatosításában rejlik.

## 6.2. Graduálási alapelvek a tolmácsolásban

A gyakorlóbeszéd graduálásánál mindig szem előtt kell tartani, hogy a képzés melyik szakaszában vagyunk, és mely tolmácsolási mód gyakorlására szeretnénk használni az adott beszédet. A tolmácsoláson a graduálásnál a fent listázott szövegen belüli és kívüli tényezőket lehet és kell figyelembe venni, illetve fokozatosan módosítani.

A gyakorlóbeszéd graduálására vonatkozó szakmai ajánlások közül kiemelkedett három olyan forrás, mely a gyakorlatban is jól alkalmazható támpontokat ad a beszéd graduálásához: (1) a SCIC beszédgyűjteményének nehézség-besorolási rendszere (*Speech Repository grading criteria*); (2) Setton és Dawrant beszédnehézségi indexe (*Speech Difficulty Index*), valamint (3) G. Láng Zsuzsa beszédnehézségi paraméterei. A második kutatási kérdés megválaszolásához e három forrás szempontjait összesítettem a tolmácsolásban alkalmazható graduálási alapelvek rögzítéséhez (1. táblázat).

1. táblázat. A tolmácsolásban alkalmazható graduálási alapelvek

1. Tartalom, téma
<ul style="list-style-type: none"> <li>• az általános beszédtemáktól haladunk a szakmai tartalmak felé, egyre nő a specializáció;</li> <li>• fokozatosan csökken a kiszámíthatóság a tartalom és az üzenet szintjén;</li> <li>• a beszéd megértéséhez és feldolgozásához a beszédnehézség előrehaladtával egyre több készülésre, és egyre szélesebb körű tájékozottságra van szükség;</li> <li>• a témába vágó előzetes ismeretek megléte a beszédnehézség előrehaladtával egyre relevánsabbá válik, a téma kulturális meghatározottsága egyre erősebb;</li> <li>• a nehezebb beszédeknél jelennek meg az érzelmileg színezett, ironizáló tartalmak</li> </ul>
2. Beszédstruktúra, logikai felépítés
<ul style="list-style-type: none"> <li>• egyszerű, könnyen átlátható, lineáris szerkezetű beszédektől haladunk a logikai szempontból összetettebb felépítésű gyakorlóanyagok felé;</li> <li>• kezdetben túlsúlyban vannak a könnyen vizualizálható, leíró részek, majd egyre több az érvelés, az eltérő nézőpontokat ütköztető beszéd;</li> <li>• a beszédnehézség előrehaladtával egyre szerteágazóbb a beszéd logikai struktúrája, ugyanakkor a logikai kapcsolatok felismerését segítő pragmatikai jelölők aránya csökken</li> </ul>

3. Lexikai sűrűség, szókincs
<ul style="list-style-type: none"> <li>• fokozatosan nő a szakkifejezések és idiomatikus kifejezések száma, egyre gazdagabb és változatosabb a szókincs;</li> <li>• egyre több a tolmácsolási szempontból kihívást jelentő lexikai elem;</li> <li>• a nehézségi szint növekedésével megnő a kulturálisan kötött kifejezések és az idegen szavak aránya;</li> <li>• fokozatosan nő a formalitás szintje és a beszédek regisztere, egyre választékosabb szókincset használnak az előadók, nő azoknak a szavaknak a száma, melyek felsőfokú nyelvtudást igényelnek</li> </ul>
4. Információsűrűség
<ul style="list-style-type: none"> <li>• kezdetben még több az ismétlődő elem, tétovázás, a tolmácsolási feladatot megkönnyítő redundancia, ezek a beszédnehézség előrehaladtával fokozatosan fogynak;</li> <li>• egyre több a tényszerű információ, számadat és felsorolás; ráadásul ezek az adatok egyre koncentráltabban fordulnak elő a beszédekben;</li> <li>• a beszédek nyelvi és pragmatikai szinten is egyre kevésbé kiszámíthatóak</li> </ul>
5. Előadásmód
<ul style="list-style-type: none"> <li>• ideális esetben a tolmácképzésben használt beszédek döntő többsége oralizált, csak haladó szinten jelennek meg a részben felolvasott beszédek, ugyanakkor az előkészítettség foka fokozatosan nő;</li> <li>• a növekvő beszédtempó egyértelműen nehezítő tényezőnek minősül, a könnyű beszédek 100-120szó/perc körüliek, a 120-140 szó/perc közötti tempó közepesen nehéznek minősül, a 140 szó/perc feletti beszédek már egyértelműen kihívást jelentenek;</li> <li>• fokozatosan nő a beszédek hossza, konszekutív tolmácsolási módban 3 perctől 6-7 percig, szinkrontolmácsolásra szánt beszédeknél 5 perctől 10-12 percig;</li> <li>• a kezdő beszédeknél még standard akcentus és természetes prozódia jellemzi az előadókat, a középhaladó szinten már regionális akcentussal is találkoznak a hallgatók, nem anyanyelvi beszélők csak a haladó beszédeknél jelennek meg</li> </ul>

A kutatás első, holisztikus megközelítésű szakaszában ezt vizsgáltam, hogy a szakirodalmi forrásokban szereplő támpontok mennyiben használhatóak a gyakorlóbeszéd nehézségi szintjének meghatározására. Az eredményekből kiderül, hogy a forrásnyelvi korpusz beszédeinél a Speech Repository-ban megjelölt nehézségi szint nem mindig esik egybe azokkal az értékekkel, melyeket Setton és Dawrant (2016) a beszédnehézségi mutatók alapján az egyes képzési szakaszokra előírányoz. A konszekutív beszédek graduálásánál többé-kevésbé megfeleltethetőek egymásnak a kétféle elemzés eredményei, a beszédnehézségi mutatókban is

tetten érhető a graduálás alapját megteremtő fokozatos építkezés. Az alapozó és kezdő beszédek a Speech Difficulty Index alapján is a bevezető, alapozó és koordinációs képzési szakaszban javasoltak, a kísérletezési és konszolidációs szakasz több-kevésbé szinkronban van a SCIC beszédgyűjteményének középhaladó szintjével. A haladó beszédek a beszédnehézségi index alapján is tolmácsképzés végén javasoltak. Ugyanakkor említésre méltó, hogy a beszédtempó és a szakmai tartalmakban rejlő nehézségek megítélését illetően némi elcsúszás figyelhető meg a Speech Repository és a Speech Difficulty Index szintezési rendszere között.

A szinkronbeszédeknel nagyobb különbségeket látunk a kétféle graduálási rendszer között. A korpuszban eleve nem találunk olyan szinkronbeszédeket, amelyek megfelelnek az alapozó szakaszban javasolt beszédnehézségi index értékeinek, a kezdő beszédeknel is több szempontból (tempó, sűrűség és/vagy prozódia) meghaladják az előírányzott szintet. A középhaladó és haladó szinkronbeszédek többsége a beszédnehézségi index alapján is a képzés későbbi, konszolidációs és akklimatizációs szakaszába tartozna, de itt is látunk kivételeket. A két besorolási rendszer eredményeinek összevetéséből levonhatjuk azt a következtetést, hogy a szinkronbeszédeknel valamivel nehezebb a beszédnehézség egyértelmű meghatározása. Ez abban is megmutatkozik, hogy Setton és Dawrant (2016) módszertani ajánlásaiban tágabb értékhatárokat látunk a szinkronbeszédek szintezésében, mint a konszekutív beszédeknel.

### ***6.3. A számítógépes korpusznyelvészeti elemzési eszközök alkalmazhatósága a tolmácsolásra szánt gyakorlóbeszédek nehézségi szintjének mérésére és graduálására***

A fent ismertetett beszédnehézség-besorolási szempontokból és graduálási alapelvekből kiindulva, a holisztikus elemzés eredményeiből három olyan szempont emelkedik ki, melyeket analitikus eszközökkel is meg lehet vizsgálni. Ez a három szempont (1) az információsűrűség, (2) a lexikai / nyelvi komplexitás és (3) a beszéd felépítése, struktúrája.

A kutatás második, kvantitatív szakaszában végzett számítógépes szövegelemzés számos érdekes, helyenként nem várt eredményt hozott. A kutatás elsődleges célja szempontjából talán ez a kutatási szakasz a leginkább előremutató. A disszertációban részletesen bemutatom az összetett, többkomponensű korpusz-elemzés eredményeit, itt a 2. táblázatban összegzem a TextInspector és az Oxford Text Checker programokkal végzett szövegelemzés során tett főbb megfigyeléseket, elsősorban azokat az eredményeket és tendenciákat kiemelve, melyek a beszédek graduálása szempontjából relevánsak lehetnek és mindkét tolmácsolási módra érvényesek. A konszekutív és szinkrontolmácsolási

gyakorlóanyagok között detektált különbségeket a következő kutatási kérdésre válaszolva összesítem hasonló tagolásban.

## 2. táblázat. A tolmács gyakorlóbeszédok számítógépes elemzésének főbb eredményei

1. Szövegstatistikai adatok elemzése
<ul style="list-style-type: none"> <li>• a mondatok száma egyik tolmácsolási módnál sem nem nő szükségszerűen a beszédek nehézségének növekedésével párhuzamosan, kifejezetten magas mondatszámokat látunk az alapozó beszédek többségénél;</li> <li>• a szószámban mért átlagos mondathossz alapján az alapozó szintű beszédekben az előadók valóban rövidebb mondatokat használnak, mint a többi nehézségi szinten</li> <li>• mindkét tolmácsolási módban a középhaladó nehézségi szinten találjuk a leghosszabb mondatokat;</li> <li>• haladó szinten mindkét tolmácsolási módban „lerövidülnek” a mondatok, visszaesést látunk a mondatok átlagos szószámában</li> <li>• a szavak szótagokban mérhető hosszának növekedése egyik tolmácsolási módnál sem jelent szükségszerűen nehezedő tolmácsolási feladatot</li> </ul>
2. Olvashatósági mutatók elemzése
<ul style="list-style-type: none"> <li>• a forrásnyelvi korpusz beszédei az olvashatósági mutatók alapján az „elég könnyű”, az „átlagos” és a „meglehetősen nehéz” kategóriába sorolhatók, néhány középhaladó beszéd minősül csak nehezen olvashatónak;</li> <li>• a Flesch-féle olvashatósági értékek alapján a középhaladó szinten találjuk a legnehezebb beszédeket mindkét tolmácsolási módban;</li> <li>• a haladó szinten mindkét vizsgált alkorpuszban enyhe visszaesés figyelhető meg a szövegek „mérhető” nehézségében</li> </ul>
3. Szófaj-elemzés
<ul style="list-style-type: none"> <li>• az alárendelő kötőszavak száma mindkét tolmácsolási módban és nehézségi szinten alacsonyabb, mint a mellérendelő kötőszavak száma;</li> <li>• a nehezebb beszédekben több az alárendelő kötőszó (a konsekutív alkorpuszban a haladó szinten, a szinkronbeszédeknél már a középhaladó szinttől);</li> <li>• a tényszerű információk száma általában növekszik mindkét tolmácsolási módban a beszédnehézség növekedésével;</li> <li>• a tulajdonnevek számában is növekedés tapasztalható – ez a növekedés a szinkronbeszédeknél folyamatos, a konsekutív alkorpuszban inkább a kezdő és középhaladó szintek között figyelhető meg nagyságrendi ugrás</li> </ul>

4. Metadiskurzus-elemzés
<ul style="list-style-type: none"> <li>• a leggyakrabban előforduló metadiskurzus-jelölők a logikai kötőszavak, hangsúlyozó megerősítések és a hozzávetőlegesen megadott információkat jelölő kerekítések;</li> <li>• aránylag ritkán fordulnak elő témaváltást jelölő lexikai elemek, a beszélő szándékát beharangozó célbejelentések és összefoglaló jellegű szakaszjelölők;</li> <li>• a beszédek nehézségi szintjének növekedésével csökken a logikai kötőszavak, kerekítések, személyjelölők és attitűd-jelzők előfordulási aránya, valamint a témaváltások bejelentése;</li> <li>• enyhe emelkedés figyelhető meg az hangsúlyozó megerősítések, példák és magyarázatok arányában, a nehezebb beszédekben gyakrabban fordulnak elő a szövegen belüli endoforikus utalások és forrásmegjelölések;</li> </ul>
5. Lexikai profil-elemzés
<ul style="list-style-type: none"> <li>• a forrásnyelvi korpusz valamennyi beszédének feldolgozása B1+ feletti nyelvtudást igényel, a beszédek többsége B2+ és C1+ KER-szint közé tehető, a konzekutív haladó beszédek elérik a C2-es besorolást is;</li> <li>• a vizsgált beszédek lexikai profilja mindkét tolmácsolási módban fokozatos emelkedést mutat, az értékek jól tükrözik a növekvő beszédnehézséget;</li> <li>• a lexikai nehezítő tényezők (számok, személynevek, földrajzi nevek, intézménynevek, terminusok) elég pontosan beazonosíthatóak, számukban egyértelmű növekedés figyelhető meg a középhaladó és haladó nehézségi szintre sorolt beszédeknél.</li> </ul>

A fenti eredmények alapján azt mondhatjuk, hogy a számítógépes szövegelemző programok többnyire reális képet adnak a korpuszban szereplő beszédek nehézségi szintjéről, és számos nehezítő tényező kiszűrésére alkalmasak. A 3. kutatási kérdésre válaszolva, levonhatjuk azt a következtetést, hogy a tolmácsolási gyakorlóbeszédek számítógépes korpusznyelvészeti elemzése értékes tanulságokkal szolgálhat a beszédek graduálásra vonatkozóan. A TextInspector jól használható az alapvető szövegjellemzők összehasonlítására, a szövegkohéziót növelő metadiskurzus-jelölők beazonosítására és a beszédek lexikai profiljának felállítására. Az Oxford Text Checker-rel felmérhető a beszédek lexikai készlete, és könnyebb beazonosítani a tolmácsolási csapdákat. Emellett a szövegelemző programokkal kiszűrhetőek bizonyos különbségek a két tolmácsolási mód gyakorlására szánt beszédek között, ami továbbvezet minket a 4. kutatási kérdéshez.

#### **6.4. A konsekutív és szinkrontolmácsolási módok közötti eltérések a beszédek nehézségét meghatározó tényezők tekintetében**

A konsekutív és a szinkrontolmácsolás közti különbségekből adódóan más típusú beszédek alkalmasak a két tolmácsolási mód gyakorlására. Ezek a különbségek nemcsak a beszédek hosszában, hanem a gyakorlóanyagok logikai felépítésében, információsűrűségében és a tolmácsolást nehezítő lexikai tényezők előfordulásában is tetten érhetőek. Graduálási szempontból a számítógépes korpusznyelvészeti elemzés számos párhuzamosságot mutat a konsekutív és szinkron alkorpusz beszédei között, de eltérések is mutatkoznak. A 3. táblázatban a két tolmácsolási mód között megfigyelhetőek különbségeket fűztem egybe a vizsgált elemzési szempontok alapján.

3. táblázat. A konsekutív és szinkrontolmácsolás közötti eltérések a korpusznyelvészeti beszédnehézség-elemzés alapján

1. Szövegstatistikai adatok elemzése
<ul style="list-style-type: none"><li>• a konsekutív beszédeknel egyértelműbb az összefüggés a beszédek hosszának növekedése és a nehézségi szint emelkedése között, mint a szinkronbeszédeknel;</li><li>• a konsekutív gyakorlóanyagokban aránylag állandó a type/token arány az egyes nehézségi szintek között, a szinkronbeszédeknel viszont egyértelmű csökkenés tapasztalható a lexikai változatosság tekintetében a két középső nehézségi szinten;</li><li>• a type/token arány általában magasabb a konsekutív beszédeknel, mint a szinkron alkorpuszban, ami több ismétlődésre, nagyobb mértékű redundanciára utal;</li><li>• a konsekutív beszédekben valamivel magasabb az egy mondatra jutó szavak számának átlaga, valamivel hosszabbak a mondatokat, mint a szinkrontolmácsolásra szánt anyagoknál</li></ul>
2. Olvashatósági mutatók elemzése
<ul style="list-style-type: none"><li>• a Flesch-Kincaid olvashatósági szintek alapján a konsekutív beszédek összességében nehezebben olvashatónak tűnnek, mint a szinkronbeszédek;</li><li>• a konsekutív beszédeknel nagyobb különbség van az alapozó és a kezdő szintű beszédek nehézsége között, mint a szinkron gyakorlóanyagoknál</li></ul>
3. Szófaj-elemzés
<ul style="list-style-type: none"><li>• a konsekutív beszédeknel nő az egy mondatra jutó főnévi és igei elemek száma a beszédnehézség előrehaladtával, a szinkron alkorpuszban e tekintetben nem látunk látványos különbségeket az egyes nehézségi szintek között;</li></ul>



<ul style="list-style-type: none"> <li>• a konsekutív beszédek nehézségével párhuzamosan nő a nominális szerkezetek túlsúlya az igei szerkezetekkel szemben, a szinkronbeszédeknel ilyen növekedés egyedül a haladó szinten figyelhető meg;</li> <li>• a szinkronbeszédekben minden nehézségi szinten több szám és több tulajdonnév nehezíti a tolmácsok feladatát, mint a konsekutív gyakorlóanyagokban – ez nyilván a hosszabb beszédekből is adódik, de több tolmácsolási csapdát is jelent</li> </ul>
4. Metadiskurzus-elemzés
<ul style="list-style-type: none"> <li>• a konsekutív beszédekben összességében nagyobb arányban fordulnak elő metadiskurzus-jelölők, mint a szinkronbeszédeknel;</li> <li>• a konsekutív beszédekben a metadiskurzus-jelölők közül gyakran logikai kötőszavakból van a legtöbb, a szinkron gyakorlóanyagokban kevesebb a logikai kötőszó;</li> <li>• a szinkronbeszédeknel nagyobb arányban fordulnak elő az üzenet validitását igazoló hangsúlyozó megerősítések és a hozzávetőlegesen megadott információkat jelölő kerekítések;</li> <li>• a nehezebb konsekutív beszédeknel nő a példák, magyarázatok aránya, a szinkronbeszédeknel enyhe csökkenés figyelhető meg e tekintetben;</li> <li>• a szinkronbeszédekben szöveg-arányosan több a hallgatóságot megszólító kapcsolódás-jelölő, a konsekutív beszédeknel viszont gyakoribbak az összefoglaló szakaszjelölők;</li> <li>• a konsekutív beszédeknel nagyobb arányban fordulnak elő felsorolás-jelölők, és a beszédek nehézségének növekedésével nő a felsorolások előfordulási gyakorisága;</li> <li>• a szinkronbeszédekben gyakrabban fordulnak elő forrásmegjelölések, mint a konsekutív beszédekben, gyakrabban hivatkoznak egy-egy felmérés eredményeire</li> </ul>
5. Lexikai profil-elemzés
<ul style="list-style-type: none"> <li>• a konsekutív beszédek valamennyi nehézségi szinten nehezebbnek tűnnek a lexikai profil-elemzés fényében, mint a szinkronbeszédek;</li> <li>• a szinkron gyakorlóanyagokban – részben a hosszabb beszédekből adódóan – több nehezítő lexikai elemet találunk, mint a konsekutív alkorpuszban, helyenként kimagaslóan magas értékekkel és szintenként nagyobb szórásban;</li> <li>• a konsekutív beszédekben a kezdő és középhasadó szinteken gyakran magasabb a lexikai nehezítő elemek aránya az összes szövegszóhoz viszonyítva, mint a szinkron szövegekben</li> </ul>

A fent összesített eredmények igazolták azt a hipotézist, hogy korpusznyelvészeti elemzési módszerekkel is kimutatható eltérések vannak a konsekutív és a szinkrontolmácsolási célra szánt beszédek között. A tolmácsolási feladat nehézségét

meghatározó szövegjellemzőkben, a szövegkohéziót megteremtő metadiskurzus-jelölőben, és a potenciális tolmácsolási csapdának számító lexikai elemekben több aszimmetria is beazonosítható a két tolmácsolási mód között. Az a hipotézis is megalapozottnak tűnik, hogy a tolmácsolási feladat során jelentkező szövegen belüli nehézségek a konszekutív beszédeknel valamivel kiszámíthatóbbak, számítógépes elemzési módszerekkel jobban megjósolhatóak, mint a szinkronbeszédeknel.

A feltételezett nehézségek azonban nem vezetnek szükségszerűen tényleges minőségromláshoz. Az 5. kutatási kérdés épp arra kereste a választ, hogy az előzetesen beazonosított probléma források valóban kihívást jelentettek-e a tolmácsolóknak.

### ***6.5. A tolmácsolási teljesítmény alapján detektálható nehezítő tényezők, a vélt és valós nehézségek közötti átfedések és eltérések***

A hallgatói felvételek részletes elemzése alapján levonhatjuk a következtetést, hogy az előzetes lexikai profil-állítást, a beszédek szókincsének számítógépes elemzése sokszor valóban segíthet beazonosítani a potenciális tolmácsolási csapdákat. A szövegelemző programok elég pontosan detektálták azokat a lexikai elemeket, amelyekkel a hallgatóknak a tolmácsolási hibák (kihagyások, félretolmányok, tévesztések, pontatlanságok, téves szóválasztások, néma szünetek, hezitálások stb.) alapján nehézségei támadtak. A vizsgált célnyelvi korpuszban számos példát találtunk arra, hogy az előzetesen „kiszűrt” tényszerű információk, számadatok, tulajdonnevek, terminusok, szokatlan szókapcsolatok valóban megnehezítették a tolmácsoló dolgát.

Ugyanakkor a hallgatók tolmácsolási teljesítményét értékelve az is kiderült, hogy nem minden beszédnél esik egybe a számítógépes szövegelemzés alapján feltételezett beszédnehézség a tényleges, tolmácsolás közben érzékelt nehézségi szinttel. A vizsgált korpuszban több példát is találtunk arra, hogy az előzetesen felállított objektív nehézségi szinthez képest jóval nagyobb kihívást jelentett a beszéd tolmácsolása. Mindkét tolmácsolási módban akadt olyan beszéd, amely a számítógépes elemzés alapján nem ígérkezett különösebben nehéznek, mégis elég sok tolmácsolási hibához, tartalmi kihagyáshoz és tévesztéshez vezetett. Különösen igaz volt ez a bonyolultabb mondatszerkesztésű, vagy több idiomatikus kifejezést tartalmazó beszédekre, az ilyen típusú nehézségeket kiszűrésére a szövegelemző programok csak részben alkalmasak.

A kutatás harmadik szakaszában a két tolmácsolási módban szintenként összesítettem a hallgatói teljesítmény alapján levonható tanulságokat. Ez alapján képet kaphatunk arról, hogy

mik azok a lehetséges problémaforrások, melyek a számítógépes szövegelemző programokkal megjósolhatóak, illetve mik azok a nehézségek, melyeknek a kiszűrésére ezek az eszközök kevésbé alkalmasak.

a) Az objektív és szubjektív beszédnehézség közötti átfedések:

- az alap, illetve könnyű kezdő beszédekben kimutathatóan kevesebb a tényszerű információ (szám, személynév, földrajzi név stb.); ezek könnyen értelmezhető kontextusban szerepelnek, valóban nem jelentenek kihívást a hallgatóknak;
- a (közép)haladó besorolású beszédekben kimutathatóan több, számítógépes elemzéssel is azonosítható tolmácsolási csapdát (földrajzi neveket, személyneveket, intézményneveket, titulust) találunk – ezeknek egy része még valóban nehézséget okoz a hallgatóknak;
- a számítógépes lexikai elemzéssel kiszűrt C1-es kifejezések valóban visszatérő nehézséget okoznak (kezdő és középhasadó szinten is) – ezeknél a kifejezéseknél megszorodnak a kihagyások, néma szünetek, egyeztetési hibák;
- a metadiskurzus-jelölők segítségével beazonosítható többtagú felsorolásoknál valóban gyakrabban fordulnak elő kihagyások, grammatikai hibák;
- a számítógépes lexikai elemzést megerősítve kezdő szinten nehézséget jelentenek azok a szavak, melyek csak körülírással vagy betoldott magyarázatokkal adhatók vissza;
- a hallgatóknak többnyire nehézséget okoz – még középhasadó szinten is –, ha szakmai tartalmakat kell tolmácsolniuk, a terminusok megfeleltetése nem egyszer pontatlan, gyakoriak a tévesztések, a szakkifejezések előtti néma szünetek;

b) Az objektív és szubjektív beszédnehézség közötti eltérések:

- alap, illetve könnyű kezdő szinten a téma-specifikus köznyelvi kifejezések is gondot okozhatnak még, sok tartalmi kihagyást és általánosítást látunk ezeknél a lexikai elemeknél a hallgatók tolmácsolásaiban;
- a növekvő tolmácsolási nehézséget időnként nem maguk a potenciális csapdaként megjelölt lexikai elemek adják, hanem a kontextus, amiben szerepelnek (pl. az előadó egy komplex felmérés eredményeit ismerteti, vagy sok illusztráló példát sorakoztat fel);
- nem minden lexikai elemet tud kiszűrni a program, ahol részletesebb kifejtésre lenne szükség – pl. a kulturális utalások vagy a mögöttes jelentéstartalom miatt (ezeknél gyakran megszorodnak a kihagyások, elcsúsznak a logikai összefüggések);

- a többjelentésű vagy kulturálisan kötött köznyelvi kifejezések helyes értelmezése sokszor problematikus – ezeket a szövegelemző programok csak részben tudják kiszűrni;
- számítógépes elemzőeszközökkel csak részben szűrhetőek ki a nehezebben átváltható idiomatikus kifejezések vagy szokatlan, egyéni megfogalmazások, melyek a hallgatók tolmácsolásaiban gyakran tartalmi kihagyásokhoz, stilisztikai veszteségekhez, újrafogalmazáshoz, egyeztetési hibákhoz vezetnek.

#### c) A konszekutív és szinkronbeszédok közötti eltérések

- a szinkronbeszédok tolmácsolásaiban gyakrabban maradnak ki azok a metadiskurzus-jelölők, illetve kötőszavak, melyek a következő gondolati egységet vezetik fel; a konszekutív beszédekben ezek az összekötő elemek valamivel hangsúlyosabbak;
- a nagy számban, felsorolásszerűen előforduló földrajzi nevek és számadatok nagyobb nehézséget jelentenek a szinkronbeszédokban, sokszor hiányosan, pontatlanul vagy nem megfelelő kontextusban szerepelnek a hallgatói felvételekben; konszekutív beszédekben érezhetően sokat segít a jegyzetelés ezeknek az adatoknak a rögzítésében és felidőzésében;
- az egy szegmensben belül sűrűn előforduló szakkifejezések szinkron tolmácsolási módban komolyabb kihívást jelentenek, ami kihat a mondat szerkesztésre, a folyamatosságra, megszorodnak a hezitálások, grammatikai hibák; konszekutív tolmácsolási módban több idő van beazonosítani és szintetizálni ezeket a „halmozottan előforduló” terminusokat;
- a hallgatói felvételekben detektálható szegmentálási hibák, félrefordítások és tartalmi pontatlanságok arra utalnak, hogy szinkronbeszédokban nagyobb nehézséget okoznak a hosszú, bonyolult mondatok és a hipotetikus, feltételes módban megfogalmazott állítások; konszekutív módban ezek könnyebben átláthatóak és szegmentálhatóak;
- szinkron tolmácsolási módban visszatérő tartalmi és grammatikai hibákhoz, néma szünetekhez vezetnek a nehezen megfeleltethető kifejezések, a konszekutív tolmácsolásban több idő van ezeknek az átváltására és következetes használatára;
- a szinkron feladatokban gyakrabban találkozunk – különösen idiomatikus kifejezéseknél – magyartalan megfogalmazásokkal, az idegen szókapcsolatból átvett negatív transzferrel; konszekutív tolmácsolásnál is előfordul, de itt több idő és energia marad az átváltásra;
- a szinkronbeszédokban nagyobb nehézséget okoznak a hallgatóknak a többször ismételt, részleges átfedésben lévő gondolati egységek – ezeket konszekutív tolmácsolási módban könnyebb „szétszálazni” és pontos összefüggéseiben átadni.

## 7. Az eredmények alkalmazásának lehetőségei és korlátai a tolmácsképzésben

A disszertáció elején abból a hipotézisből indultam ki, hogy a tolmácsképzésben is hasznos tanulságokkal szolgálhatnak a számítógépes korpusznyelvészeti elemző programok. A kutatás fent bemutatott eredményei igazolták, hogy valóban igen értékes információkkal szolgálhatnak a tolmácsképzésben használt gyakorlóbeszédnek nehézségéről. Segítségükkel kiszűrhető számos olyan szövegen belüli problémaforrás, mely nehézséget jelenthet a hallgatóknak a képzés egyes szakaszaiban. Számítógépes szövegelemző eszközökkel valóban szisztematikusabbá és tudatosabbá válhat a beszédnek graduálása, a tolmácsoktatók könnyebben tudják a hallgatók aktuális tudásszintjéhez igazítani és fokozatosan növelni a feldolgozott gyakorlóanyagok nehézségi szintjét.

A kvantitatív kutatásból jól látszott, hogy mind a TextInspector, mind az Oxford Text Checker program szövegelemzése hozott olyan eredményeket, melyek az oktatási gyakorlatban is jól használhatóak. Előbbi a beszédnek szövegjellemzőin keresztül információt adhat a hallgathatóságról, az értelmezést segítő kohézív elemekről, a szövegek lexikai sokszínűségéről, és összességében a forrásnyelvi üzenet feldolgozásához szükséges nyelvtudásról. Az Oxford Text Checker segít felmérni a beszédnek szókészletét és támpontokat adhat a nehézségi szint meghatározásában is. Ráadásul rendkívül felhasználóbarát és didaktikus a színekkel kódolt szövegelemzésnek és a szókinccsfeljegyzéshez generált szólistáknak köszönhetően. Nemcsak a graduálásban, hanem a gyakorlóanyagok előkészítésében is igen hasznos eszköznek bizonyulhat.

A kvalitatív kutatási eredmények részben megerősítették a számítógépes szövegelemző programok hasznosíthatóságát, hiszen számos átfedést találtunk a feltételezett és a valós beszédnehézség között. A kiszűrt tolmácsolási csapdák jelentős része valóban nehézséget okozott a hallgatóknak, a felállított lexikai profil reálisnak bizonyult, és a hallgatóktól aktuálisan elvárható tudásszintnek megfelelt. Emellett a szövegelemző programokkal számos eltérést sikerült feltárni a konszekutív és a szinkronbeszéd közötti, így ezek az eszközök segíthetnek az oktatóknak abban is, hogy eldöntsék, az adott beszéd melyik tolmácsolási módra alkalmasabb.

A hallgatói felvételek elemzése ugyanakkor kimutatta azt is, hogy ezek az eszközök nem használhatóak kizárólagosan a gyakorlóanyagok graduálására. Részben azért, mert bizonyos típusú nehézségeket nem tudnak kiszűrni, részben azért mert a beszédnehézség megítélésénél nem hagyhatóak figyelmen kívül a szubjektív és a szövegen kívüli változók. Annak érdekében, hogy leszűkíthessem a kutatás fókuszát és valid eredményeket hozzon a

számítógépes korpusznyelvészeti elemzés, jelen kutatásban csupán a szövegen belüli változókat vizsgáltam. Egy átfogó, minden részletre kiterjedő graduálási szempontrendszer kidolgozásánál egyértelműen figyelembe kell venni a szövegen kívüli változókat is. Ezek a változók sajnos nem igazán mérhetőek korpusznyelvészeti eszközökkel, így túlmutatnak jelen kutatás fókuszán.

Ehhez kapcsolódóan meg kell jegyeznünk azt is, hogy – a legtöbb tudományos kutatáshoz hasonlóan – ennél a kutatásnál is vannak bizonyos korlátai az eredmények alkalmazhatóságának. A fent bemutatott számítógépes szövegelemző programok használatát például behatárolja, hogy csak olyan esetekben alkalmazhatóak egy az egyben, ahol elektronikus formában rendelkezésre áll a beszéd leírt változata, vagy a felhasznált forrás. Ennek megfelelően a korpusznyelvészeti elemző eszközök leghatékonyabban azoknál a beszédgyűjteményekből származó gyakorlóanyagoknál használhatóak, ahol a beszéd leírata is hozzáférhető. Ugyanakkor a graduálás szempontjából is értékes információkkal szolgálhat a készüléshez használt jegyzetek és internetes források elemzése is, különösen a potenciális tolmácsolási csapdák beazonosításában. További korlátokat szab a fenti számítógépes korpusznyelvészeti eszközök alkalmazásának, hogy a szövegelemző programokat alapvetően angol nyelvű szövegek elemzésére fejlesztették ki.

A kutatás tanulságait összegezve azt mondhatjuk, hogy a tolmácsolás szempontjából számos izgalmas lehetőség rejlik a korpusznyelvészeti eszközök használatában. Amint a fenti eredmények igazolják, jól alkalmazhatóak például a gyakorlóbeszéd tolmácsolási nehézségének mérésére, számos támpontot adnak a graduáláshoz. Tolmácsoktatóként érdemes tehát számítógépes szövegelemzéssel is bővíteni a módszertani eszköztárunkat, hiszen a technológiában rejlő lehetőségek felfedezésével és kiaknázásával új távlatok nyílnak a tolmácsolásban.

## A tézisekhez felhasznált irodalom jegyzéke

- Albl-Mikasa, M. 2013. Developing and cultivating expert interpreter competence. *Interpreters' Newsletter* 18. 17–34.
- Alexieva, B. 1999. Understanding the source language text in simultaneous interpreting. *The Interpreters' Newsletter*. No. 9. 45–59.
- Bakti M. 2008. Megakadásjelenségek a szinkrontolmácsolásban. *Fordítástudomány* X. évf. 2. szám. 22–38.
- Bakti, M. 2010. *Diszharmonias jelenségek a szinkrontolmácsolás célnyelvi beszédprodukciónakban*. Doktori disszertáció. ELTE BTK Fordítástudományi Doktori Program.
- Bodie, G. D. 2018. Listenability Style Guide (LSG). In: Worthington, D.L., Bodie, G.D. (eds). *The Sourcebook of Listening Research: Methodology and Measures*. John Wiley & Sons. <https://doi.org/10.1002/9781119102991.ch37>
- Brindley, G. and Slatyer, H. 2002. Exploring task difficulty in ESL listening assessment. *Language Testing* Vol.19. No. 4. 369–394. <https://doi.org/10.1191/0265532202lt236oa>
- Déjean le Féal, K. 1978. *Lectures et improvisations. Incidences de la forme de renonciation sur la traduction simultanée (français-allemand)*. These de doctorat de 3e cycle. Université de la Sorbonne Nouvelle
- Déjean le Féal, K. 1982. Why impromptu speech is easy to understand. In: Enkvist, N. E. (ed.) *Impromptu Speech. A Symposium*. Abo Research Institute of the Abo Academic Foundation. 221–239
- DuBay, W. H. 2006. *Smart language: Readers, Readability, and the Grading of Text*. Costa Mesa: Impact Information
- Flesch, R. 1948. A new readability yardstick. *Journal of Applied Psychology*. Vol. 32. No. 3. 221–233. <https://doi.org/10.1037/h0057532>
- Flesch, R. 1974. *The Art of Readable Writing*. New York: Harper.
- G. Láng, Zs. 2002. *Tolmácsolás felsőfokon. A hivatásos tolmácsok képzéséről*. Budapest: Scholastica.
- G. Láng, Zs. 2003. Forrásnyelvi beszédek a tolmácsképzésben. In: Klaudy K. (szerk.) 2003. *Fordítás és tolmácsolás az ezredfordulón*. Budapest: Scholastica. 38–48.
- Gerver, D. 1971/2002. A Psychological Approach to Simultaneous Interpretation. *Meta: Translators' Journal* Vol. 20. 119–128. <https://doi.org/10.7202/002885ar>

- Gile, D. 1995/2009. *Basic Concepts and Models for Interpreter and Translator Training*. Amsterdam/Philadelphia: John Benjamins. [https://doi.org/10.1075/btl.8\(1st\)](https://doi.org/10.1075/btl.8(1st))
- Gunning, R. 1952. *The Technique of Clear Writing*. New York: McGraw-Hill.
- Harwood, K.A. (1955) I. Listenability and readability. *Speech Monographs*. Vol. 22. No.1. 49–53. <https://doi.org/10.1080/03637755509375133>
- Hönig, H. G. 2002. ‘Piece of Cake – or Hard to Take? Objective Grades of Difficulty of Speeches Used in Interpreting Training. *Teaching Simultaneous Interpretation into a “B” Language*. EMCI Workshop. 20-21 September 2002.
- Jones, R. 1998/2002. *Conference Interpreting Explained*. London, New York: Routledge.
- Kalina, S. 1998. *Strategische Prozesse beim Dolmetschen. Theoretische Grundlagen, empirische Fallstudien, didaktische Konsequenzen*. Tübingen: Gunter Narr. <https://doi.org/10.7202/011017ar>
- Klare, G. R. 1963. *The measurement of readability*. Ames: Iowa State University Press.
- Klaudy K. 2005. Párhuzamos korpuszok felhasználása a fordításkutatásban. In: Lanstyák, I., Vančoné Kremmer, I. (szerk.). 2005. *Nyelvészetről – változatosan. Segédkönyvek egyetemisták és a nyelvészet iránt érdeklődők számára*. Dunaszerdahely: Gramma Nyelvi Iroda. 153–185.
- Liu, M., Chiu, Y. 2009. Assessing source material difficulty for consecutive interpreting. Quantifiable measures and holistic judgment. *Interpreting*. Vol. 11. No. 2. 244–266. <https://doi.org/10.1075/intp.11.2.07liu>
- Mead, P. 2015. Input Variables. In: *Routledge Encyclopedia of Interpreting Studies*. London, New York: Routledge. 191–192.
- Monacelli, C. 2005. *Surviving the role: A corpus-based study of self-regulation in simultaneous interpreting as perceived through participation framework and interactional politeness*. PhD. dissertation. School of Management and Languages. Heriot-Watt University. Edinburgh.
- Pöchhacker, F. 1994. *Simultandolmetschen als komplexes Handeln*. Tübingen: Narr.
- Pöchhacker, F. 2004/2016. *Introducing interpreting studies*. London: Routledge. <https://doi.org/10.4324/9780203504802>
- Robin E., Seidl-Pécs O. (szerk.) 2020. *Fókuszban a fordított és a tolmácsolt szöveg: korpuszalapú fordításkutatás Magyarországon*. Segédkönyvek a nyelvi közvetítésről I. Budapest: ELTE BTK Fordítástudományi Doktori Program, MANYE Fordítástudományi Szakosztály. 79–94. <https://doi.org/10.36252/Nyelvikozvsegedkonyv1.4>



- Rubin, D. L., Rafoth, B. A. 1986. Oral language criteria for selecting listenable materials: An update for reading teachers and specialists. *Reading Psychology: An International Quarterly* Vol.7. 137–151. <https://doi.org/10.1080/0270271860070302>
- Russo, M., Bendazzoli, C., Sandrelli, E. A. 2006. Looking for lexical patterns in a trilingual corpus of source and interpreted speeches: Extended analysis of EPIC (European Parliament Interpreting Corpus). *Forum* Vol. 4. No. 1. 221–254. <https://doi.org/10.1075/forum.4.1.10rus>
- Seleskovitch, D., Lederer M. 2002. *Pédagogie raisonnée de l'interprétation*. Paris: Didier Erudition.
- Setton, R., Dawrant, A. 2016. *Conference Interpreting. A Trainer's Guide*. Amsterdam, Philadelphia: John Benjamins Publishing Company. <https://doi.org/10.1075/btl.121>
- Shlesinger, M. 1995. Shifts in cohesion in simultaneous interpreting. *The Translator* Vol. 1. No. 2. 193–214. <https://doi.org/10.1080/13556509.1995.10798957>
- Szegh, H. 2021. *Harmadik kód a tolmácsolásban: vajon létezik tolmácsolási szöveg?* Doktori disszertáció. ELTE BTK Fordítástudományi Doktori Program. <https://doi.org/10.36252/Nyelvikozvsegedkonyv1.11>
- Viaggio, S. 1996. Elementary, my dear Colleague! Educating our students' guesses. *The Interpreters' Newsletter*. No. 7. 57–72

### **Internetes szövegelemző programok:**

- Oxford Text Checker. <https://www.oxfordlearnersdictionaries.com/text-checker/>
- TextInspector. <https://textinspector.com>

### Az értekezés témájával kapcsolatos publikációk

- Besznyák, R. 2019. A stressz hatása a szóbeli nyelvi közvetítés minőségére tolmácsvizsgákon. *Fordítástudomány*. 21. évf. 2.sz. 80-100. <https://doi.org/10.35924/forttud.21.2.6>
- Besznyák, R. 2020: Analysing lexical pitfalls in interpreter training projects. In: Besznyák, R., Fischer, M., Szabó, Cs. (eds). *Fit-for-Market Translator and Interpreter Training*. Delaware, Malaga: Vernon Press. 195–212.
- Besznyák, R. 2020: Tolmácsolási gyakorlószövegek graduálása lexikai csapdák elemzésével. In: Szabó, Cs. – Bakti, M. (eds): *Iránytű a tolmácsolás oktatásához. A kompetencia-fejlesztés új fókuszai*. Szeged: Juhász Gyula Felsőoktatási Kiadó. 109–122.
- Besznyák, R., Szabó, Cs., Fischer, M. (szerk.) 2020. *Fit-for-Market Translator and Interpreter Training in the Digital Age*. Delaware, Malaga: Vernon Press.
- Besznyák, R. 2022. Vezérfonal a tolmácsvizsgák optimalizálásához. *Porta Lingua*. SZOKOE. 2022/2. szám. 83–94. <https://doi.org/10.48040/PL.2022.1.8>
- Szabó, Cs., Besznyák, R. 2021. A BME Tolmács- és Fordítóképző Központjának online tolmácsképzése: Oktatói és hallgatói tapasztalatok. *Modern Nyelvoktatás*. XXVII. évf. 3-4. szám. 82–98. <https://doi.org/10.51139/monye.2021.3-4.82.98>
- Szabó, Cs., Besznyák, R. 2021. Perspectives of Trainers and Trainees on Online Interpreting Courses at BME's Centre for Interpreter and Translator Training. *Bridge: Trends and Traditions in Translation and Interpreting Studies*. Vol. 2. No. 2. 29–49.

### Az értekezés témájával kapcsolatos konferencia-előadások

- Besznyák R. 2018. *Analysing lexical pitfalls in interpreting classes*. „Fit-for-Market Translator and Interpreter Training in a Digital Age”. BME TFK eTransFair Conference. 2018. szeptember 28-29.
- Besznyák R. 2020. *Vezérfonal a tolmácsvizsgák optimalizálásához – avagy hogyan mérjük tolmácskézségeket és minőséget a változók erdejében*. SZOKOE XX. Szaknyelvi Konferencia. 2020. november 13. (online)
- Besznyák R. 2021. *Videók használata a tolmácsképzésben*. MANYE XIII. Miskolci Alkalmazott Nyelvészeti Konferencia. „A jövő nyelvoktatását formáló nyelvtudomány”. 2021. február 4. (online)

Szabó, Cs., Besznyák R. 2021. *Pozitív oktatói és hallgatói tapasztalatok a BME TFK online tolmácsolásáról*, TransELTE2021 – OFFI, XXIII. Fordítástudományi Konferencia „Távolléti oktatás a fordító- és tolmácsolásban. Oktatásmódszertani tapasztalatok a Covid-19 idején”. 2021. március 25. (online)

Besznyák R. 2022. *Tolmácsolási gyakorlószövegek graduálása korpusznyelvészeti eszközök segítségével*. MANYE Fordítástudományi Szakosztály. Fordítástudományi Kutatások II. konferencia. 2022. május 27. (online)