

ALKALMAZOTT NYELVÉSZET 3

Doktoranduszok tanulmányai
az alkalmazott nyelvészet köréből



MTA Nyelvtudományi Intézet,
Budapest

Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2019

XIII. Alkalmazott Nyelvészeti Doktoranduszkonferencia

Budapest, 2019. 02. 01.

Sorozatszerkesztő:
Váradi Tamás

Szerkesztő:
Ludányi Zsófia, Grácz Tekla Etelka

Lektorálta:

Balaskó Mária
Bóna Judit
Domonkosi Ágnes
Farkas Judit
Fenyvesi Anna
Fóris Ágota
Gósy Mária

Grácz Tekla Etelka
Hámori Ágnes
Hunyadi László
Károly Krisztina
Kohári Anna
Krepsz Valéria
Kuna Ágnes
Ludányi Zsófia

Markó Alexandra
Navracsics Judit
Prószéky Gábor
Robin Edina
Sass Bálint
Schirm Anita
Simon Eszter

Technikai szerkesztés:
Ludányi Zsófia, Grácz Tekla Etelka

A technikai szerkesztés munkatársai:
Ferenczi Zsanett, Jankovics Julianna, Kondor Kitti, Száraz Bettina

ISBN 978-963-9074-81-1

DOI: 10.18135/Alknyelvdok.2019.13

Kiadja: MTA Nyelvtudományi Intézet

A folyóiratot az MTMT indexeli és a REAL archiválja.

Budapest, 2019

Tartalomjegyzék

Asztalos Anikó: <i>A tanulói szóátadások vizsgálata</i>	5
Balogh Andrea: <i>A játéknyelvi szókincset befolyásoló tényezők</i>	21
Bartók Márton: <i>Érzelmi tartalmak észlelése a magyar beszédben</i>	34
Gyulai Livia: <i>Nem kompozicionális igekötős igék argumentumszerkezetének korpuszalapú vizsgálata</i>	45
Huszár Anna: <i>A karakterábrázolás alapfrekvencia-jellemzői</i>	61
Jankovics Julianna: <i>Kérdő funkciójú megnyilatkozások enyhe és középsúlyos értelmi fogyatékosok felolvasásában</i>	74
Kiss Anita: <i>Határon túli magyar egyetemisták nyelvhasználatának megítélése</i>	92
Klenk Márk: <i>Terminológiai változtatások az Alaptörvény különböző, angol nyelvű változataiban</i>	104
Kovács Viktória: <i>Az elérhetőségi elmélet névmási anaforafeloldásra gyakorolt hatása</i>	117
Üveges István: <i>Névelem-felismerés magyar nyelvű jogi szövegeken</i>	126
Weidl Zsófia: <i>Egyet nem értékek fonetikai jellemzői középiskolások beszédében</i>	138
XIV. AlkNyelvDok (2020. febr. 7.) felhívása	143

A tanulói szóátadások vizsgálata

Asztalos Anikó

ELTE BTK Nyelvtudományi Doktori Iskola
asztalos.aniko@btk.elte.hu

Kivonat: A tanulmány középpontjában 8 tanóra tanulói szóátadásainak az elemzése áll, valamint a tanórák fő kommunikációs jellemzőinek feltárása (pl. a tanár-diák beszéd aránya, a beszédfordulók száma stb.). A vizsgálat eredményei differenciálhatják a tanulói szóátadások típusait, valamint rámutathatnak a sajátosságokra, így például pragmatikai funkciójukra, megjelenésük gyakoriságára, terjedelmükre. A kutatási eredmények alapján a tanári és a tanulói beszédfordulók sajátosságai között összefüggések fogalmazhatók meg. A kutatási eredmények bemutathatják, hogy hogyan befolyásolják a tanítási órán alkalmazott munkaformák a tanulói beszédet, azon belül is a tanulók egymás közötti kommunikációját.

1 Bevezetés

A Nemzeti alaptanterv az anyanyelvi nevelés kiemelt feladatákként nevezi meg a nyelv mint változó rendszer megismerését és a nyelvi kompetencia fejlesztését. Ennek egyik célja, hogy a tanulók az átvitt és implicit jelentéseket fel tudják ismerni a kommunikációban, és reflektálni tudjanak rájuk (NAT 2012). Ehhez a képességhez hozzátartozik a diskurzus különböző jellemzőinek ismerete és az ezzel kapcsolatos készségek és képességek alkalmazása, fejlesztése is. Ezt a folyamatot támogathatja, ha a diákoknak a tanórákon több lehetőségük van arra, hogy egymással beszéljenek, illetve különböző szituációs gyakorlatok nyomán a diskurzusokat megfigyeljék és elemezzék. A tanulók beszédének jellemzői a tanórák felvétele és ezek több szempontú elemzése során tárhatók fel.

A tanulói beszéd kutatása számos különböző tudományág bevonásával, több megfigyelési szempont alapján történhet, így például lehetséges a tanulók egyéni beszédjellemeinek elemzése fonetikai és pszicholingvisztikai aspektusból (Bóna 2015; Horváth 2017), a tanári és a tanulói beszéd arányának megállapítása diskurzuselemzési szempontból (Herbszt 2006; Antalné 2015), a különféle munkaformák és módszerek tanulói beszédre gyakorolt hatásainak vizsgálata (Kelemenné 2012; Asztalos 2015) stb. Az egyéni tanulói beszédjellemezők az osztályterem közegében korlátozottan vizsgálhatók, ugyanis technikai akadályokba ütközhet a diákok beszédének felvétele és azonosítása. Módszertanilag a szinkódos mikrofonokkal történő felvétel előrelépést jelentett a tanulók beszédének kutatásában (Skidmore et al. 2003).

Asztalos Anikó: A tanulói szóátadások vizsgálata. In Váradi Tamás (sorozatszerkesztő), Ludányi Zsófia, Gráczki Tekla Etelka (szerkesztő): *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2019. XIII. Alkalmazott Nyelvészeti Doktoranduszkonferencia*. Budapest: MTA Nyelvtudományi Intézet. 2019. 5–20. DOI: 10.18135/Alknyelvdok.2019.13.1

Jelen kutatás célja, hogy megpróbálja felvázolni a tanulók szóátadásainak lehetséges típusait, a hozzákapcsolódó fő pragmatikai funkciókat diktafonnal és kamerával rögzített tanórafelvételek elemzése alapján, valamint, hogy rámutasson azokra a tényezőkre, amelyek a tanulói szóátadásokat befolyásolhatják. A kutatás elméleti keretét elsősorban a társalgáselemzés és módszertani kerete, különösen a nyelvészeti konverzációelemzés (Sacks et al. 1974; Boronkai 2013) és az osztálytermi diskurzuselemzés (Antalné 2006, 2010, 2015; Herbszt 2006, 2008) adja, valamint a vizsgálat támaszkodik pragmatikai szempontokra és pedagógiai szemléletre is, az itt alkalmazott módszertan is az ebben a keretben szokásos módszereket követi.

Az eddigi vizsgálatok eredményei alapján megfogalmazható, hogy az osztálytermi diskurzusban a tanár a domináns beszélő, aki általában olyan kérdéseket tesz fel, amelyekre tudja a választ. Azon beszélgetések során, ahol a pedagógus olyan kérdést fogalmaz meg, amelyre nem tudja előre a választ, vagyis a kérdés nem „tesztkérdés”, hanem valódi kérdés, megnyílik a párbeszéd lehetősége. Ez különösen érvényes akkor, ha a tanár nem jelöli ki a következő felszólalót, vagyis a diákok nem külválasztás hatására szólalnak fel, illetve lehetőséget ad arra, hogy a diákok egymással beszéljenek. A tanár általában ösztönzi azt, hogy mindenki hozzájáruljon a tanórai megbeszélésekhez, de több alkalommal is csökkenti a közvetlen diák-diák beszélőváltás valószínűségét. Az órákon a kommunikációt sokkal inkább az ún. „pedagógiai párbeszéd” uralja, amelyre az jellemző, hogy a tanár, aki ismeri és birtokolja az elvárt információt és tudásanyagot, utasít valakit a válaszadásra, aki nem tud róla, vagy hibásan tudja. A valódi párbeszédre törekvő oktatás jobb képzést eredményezne, ugyanis pozitív hatással van a diákok kognitív teljesítményére, így a tanulói beszédidő növelése kiemelt oktatási célként értelmezhető (Skidmore et al. 2003; Bentley 2007; Samei et al. 2015).

1.1 Az osztálytermi diskurzus pragmatikai és pszicholingvisztikai megközelítése

A pragmatika és a pszicholingvisztika a tanulói beszéd megközelítésében olyan szemléletmódot kínál, amely a nyelv jelenségeit a nyelvi tevékenységből kiindulva vizsgálja. Az osztálytermi diskurzus, amelybe a tanulói megnyilatkozások ágyazódnak, több, mint a mondatok összessége, a legtágabban, interakcióként értelmezett beszéd egy fajtája, amely a szereplők közötti társas viszonyt is meghatározza. A diskurzus a nyelvi tevékenység folyamatában jön létre, szükség van hozzá a résztvevők konstruktív hozzájárulására, vagyis a diákok és a tanár együttes erőfeszítéssel hozza létre az osztálytermi diskurzust is (Tátrai 2004; 2011; Pléh 2014). A diskurzus általános összetevői Hymes „SPEAKING”-modellje alapján a következő aspektusok mentén jellemezhetők (Hymes 1974; Bartha 1998; Hámori 2006; Antalné 2015):

- a beszédesemény fizikai és pszichés körülményei: az osztályterem helye, a tanóra ideje, az osztálytermi diskurzus résztvevőinek pszichés állapota stb.;
- a résztvevők egyéni jellemzői: a tanár és a diákok életkora, neme, a köztük lévő viszony, a tanár státusza az osztályteremben stb.;
- a szándékok, a célok és a kimenetek: ebbe beletartozhatnak a pedagógus és a diákok saját céljai, a tanulási-tanítási folyamat rövid és hosszú távú céljai stb.;
- a beszédesemény formája, szerkezete, beszédaktusai: azok az intézményesített jellemzők, amelyek a tanórai diskurzust jellemzik pl. tanórai jelentés, fegyelmezés, felelés stb.;

- a beszéd értelmezésére szolgáló kulcsok és jelzések: milyen verbális és nem verbális eszközöket használnak a diskurzus résztvevői;
- a csatornák és a beszédformák: milyen dialektust, regisztert és stílust választanak az osztálytermi diskurzus résztvevői;
- a beszélésre vonatkozó normák: a tanár és a diákok közös tudása, a közös értelmezések stb.;
- a műfajok: maga a tanóra, amelyben monologikus és a dialogikus részek váltakoznak stb.

Az osztálytermi diskurzus tipikus beszédhelyzete, hogy a tanár az, aki az órát koordinálja, általában ő jelöli ki a következő feladatot, ő kér számon, és segíti a tanulók munkáját. Ezzel összefüggésben, a tanári szerephez egy speciális státusz kötődik, amely általában minden résztvevő számára elfogadottnak számít. A státusz a tanár azon pozíciója a társas szerkezetben, amely magában foglalja jogait és kötelességeit (Labov–Fanshel 1997). A tanulók szerepét azonban nehezebb meghatározni, hiszen számos tényező befolyásolhatja, így például a téma, a tanórán való megszólalási esélyeik, az órán választott munkaforma vagy szervezési mód stb. A szociolingvisztika újabb, társas konstruktivista iránya másként tekint a státusz fogalmára. A diskurzusbeli státuszt és szerepeket minták és normák megvalósulásaiként értelmezi, miközben fókuszál a beszélőkre és a nyelvhasználatukra, a nyelvi viselkedésre és a jelentéslétrehozás folyamatára is. Ebben az értelmezésben a státuszt a résztvevők hozzák létre minden diskurzusban, a konkrét helyzetben folyamatosan megismételve, megerősítve vagy újra-konstruálva a szerepeket és státuszokat (Bartha–Hámori 2010).

1.2 Együttműködés az osztálytermi diskurzusban

Az osztálytermi diskurzus általában közvetlen interakciók nyomán valósul meg, a résztvevői egy térben és egy időben helyezkednek el. Ez egy olyan speciális többszereplős helyzet, amelyben a tanár és a diákok a kommunikációnak a különféle formáit gyakorolhatják. Így az osztályteremben megvalósulhat egyéni, páros és csoportos kommunikáció egyaránt (Williams–Wright 1991). Tipikus helyzetnek tekinthető, amikor a tanár megnyilatkozóként beszél a tanulóknak, akik a címzettek. A tanár megnyilatkozásai a tanulókhöz szólnak, a tanulók megnyilatkozásai szólhatnak a pedagógushoz és a diák-társakhoz egyaránt. A tanári kérdésfeltevés esetében a résztvevői szerepek változhatnak, ugyanis a tanuló válik megnyilatkozóvá, és a pedagógus lesz a befogadó.

A diskurzus létrehozásához szükség van a partnerek együttműködésére, a diskurzus közös erőfeszítés eredménye (Grice 1975/1997). A beszédet a felek úgy tervezik meg, hogy figyelembe veszik, hogy a másik mit tud, és mit nem, ettől az előtervezésétől várható az, hogy a kétértelműség nagy részét elkerülik, vagyis a beszélő törekszik arra, hogy a legnagyobb mértékben biztosítsa a hallgató számára az elhangzottak feldolgozhatóságát (Schlegloff 1984; Gósy 2003). Emellett a résztvevők folyamatosan jeleznek egymásnak arról, hogy megértették-e azt, amit a másik mond (Clark 2001). Mindezt olyan társalgási maximák segítik, mint a mennyiség, a minőség, a viszony (relevancia) és a modor (Grice 1975/1997). A maximák működésének vizsgálata nyomán megállapítható, hogy a tantermi kommunikáció aszimmetrikus jellege miatt főként a pedagógusnak áll módjában megsérteni a különböző maximákat, a tanulókra kevésbé jellemző a társalgási maximák megsértése. Ennek egyik oka, hogy az iskola szűkszavúságra neveli őket. A diákok leggyakrabban a mód maximáját sértik meg, ekkor a pedagógus javításra ösztönzi őket (Herbszt 2008).

1.3 Az osztálytermi diskurzus szerkezete, a szóátadás az osztálytermi diskurzusban

Az osztálytermi diskurzus különböző mikro- és makrostrukturális elemekből áll, alapegysége a megnyilatkozás, amely funkcionálisan, pragmatikai és paralingvisztikai szempontból is elkülöníthető. A megnyilatkozások kommunikációs funkciójuk, a hozzájuk kapcsolódó beszédaktusok alapján sokfélék lehetnek: kapcsolatteremtő, -tartó és -záró megnyilatkozások, kijelentések, felszólítások, kérések, kérdések, értékelések stb. (Walsh 2006; Antalné 2010; 2015). Egy vagy több megnyilatkozás alkotja a beszédfordulót, amely a megszólalástól addig tart, ameddig beszélőváltás nem történik. A beszélőváltás vizsgálatában kiemelt szerepe van annak, hogy kinek a kezdeményezésére történik a szó átvétele. Azokban a diskurzusokban, amelyekben több a résztvevő, a beszélő külső választással jelöli ki azt a résztvevőt, akinek át kívánja adni a szót, azonban a következő beszélő önkiválasztással is megszólalhat. A formálisabb helyzetekre, így például az osztálytermi diskurzusra is a külválasztás, míg az informálisabb helyzetekre, például a hétköznapi beszélgetésekre az önkiválasztás stratégiája jellemző (Sacks et al. 1974; Iványi 2003; Boronkai 2013). Jelen vizsgálat kifejezetten a külválasztásokra fókuszál.

A szóátadás általában a beszédforduló határán történik, lehetnek különféle nyelvi és nem nyelvi eszközei egyaránt, így az aktuális beszélő az állításoknál leviheti, a kérdéseknél felemelheti a hangját, valamint a félig formalizált szituációk során a magasabb státusú beszélő jelezheti azt is, hogy az éppen beszélő személy ideje lejárt (Rosengern 2004; Boronkai 2013; Schlegloff 2015). A korábbi kutatási anyag alapján megállapítható, hogy a szóátadás jelenségének leírásához, meghatározásához az egész diskurzust és minden tényezőjét szükséges figyelembe venni. A szóátadás alapvető pragmatikai jellemzője a figyelemirányítás, amely a megnyilatkozás tartalmától függően érzékeltetheti azt, hogy a beszélő a mondandója végére ért (pl. elhallgat), vagy információt szeretne megtudni a partnerértől (pl. kérdez). A szóátadás szándékát vagy a szó átvételére alkalmas helyet elhallgatáskor az ereszkedő intonáció jelzi, míg a kérdések esetében a teljes megnyilatkozás a szó átadásának és az információ vagy vélemény kérésének szándékát jelzi.

A szóátadáskor használt jelölők lehetnek kötőszói eredetűek, de más szófajok jellemzőit is hordozhatják, mivel a kapcsolatot a beszédforduló után a szót átvett beszélő a saját fordulója elején fogja jelezni. A szó átadása egyértelmű is lehet akkor, ha a megnyilatkozó közlése befejeződik, és a beszédpartner úgy érezheti, hogy az övé a szó (Dér 2012). Így a beszélőváltás során minden következő megnyilatkozás befejezési pontja lehetővé teheti, hogy a következő beszélő a szó átvételére törekedjen (Schlegloff 1984). Ezt a folyamatot azonban társalgási szabályok irányítják, és a beszélők általában verbálisan és/vagy nem verbálisan jelzik azt, ha át akarják adni a szót, ezt nevezzük beszélőváltásnak, ez a folyamat a beszélők összehangolt cselekvése nyomán valósul meg (Hámori 2006; Schlegloff 2015). A beszélőváltás önmagában nem elegendő szempont a beszédfordulók megállapításához, ezért a beszédjog fogalma is segítségül lehet a fordulók mennyiségi meghatározásánál. A beszédjog lehetővé teszi, hogy a társalgásnak az a résztvevője átvehesse a szót az előző beszélőtől, aki a leghamarabb reagál, vagy az éppen beszélő hozzá szólt, kérdezett tőle valamit (Dér 2012). A beszélőváltásra formális helyzetben jellemző lehet a külválasztás, amikor a beszélő külső ösztönzésre szólal meg, míg az informális helyzetben inkább az önkiválasztás stratégiája érvényesül, és ha a résztvevők jól ismerik egymást, akkor a megnyilatkozásaik átfedőbbek lehetnek (Bata 2009; Boronkai 2013).

Az osztálytermi diskurzusból az intézményspecifikus és polilogikus jelleg miatt leggyakrabban a tanár választja ki a következő forduló megszólalóját, vagyis a tanulók a nézőpontjukhoz képest külső választás hatására szólalnak fel. A tanár szóátadási módjai rendkívül nagy diverzitást mutatnak, így például: kérdő megnyilatkozással; felszólító megnyilatkozással; kijelentő megnyilatkozással; felszólító és kérdő megnyilatkozással; névvel; felszólító és kérdő megnyilatkozással, valamint névvel; indulatszóval; befejezetlen mondatokkal stb. A pedagógus beszédében több olyan szóátadási típus van, amely több csoportba is beilleszthető lenne, és gyakori a szóátadások halmozása is (Király 2015).

A tanulók szempontjából az önkiválasztással történő beszédforduló-váltás akkor lehet számottevő, ha ezt az óra témája, légköre, a munkaforma vagy a szervezési mód is lehetővé teszi. A tanári és a tanulói beszédfordulók általában egy háromelemű szomszédsági párt, ún. IRF/IRE-modellt (= IRF/IRE modell = teacher initiation, student response, and teacher feedback/evaluation) alkotnak. Ez a struktúra jellemzően a tanári kezdeményezésből, a tanulói válaszból és a tanári visszacsatolásból áll (Antalné 2015).

A jelen kutatás kérdései azok voltak, hogy i) milyen a tanulói szóátadás terjedelme (szószáma, időtartama, átlagos hossza) és tartalma, ii) milyen nyelvi és nem nyelvi sajátosságai vannak a tanulói szóátadásoknak, iii) milyen a pragmatikai funkciójuk, iv) milyen gyakori a megjelenésük, v) milyen tényezők befolyásolják a tanulói szóátadásokat. Az előzetes hipotézisek a következők: 1. A tanulói beszédidő arányát befolyásolja a frontális munkaforma aránya. 2. Vannak olyan szóátadási módok, amelyek tanulóspecifikusak. 3. A tanulói szóátadások fajtái dominánsan nem nyelvek. 4. A tanulói szóátadások módjai típusokba sorolhatók. 5. A tanulói szóátadások iránya dominánsan a tanár felé történik. 6. A tanulói szóátadások típusai az órán alkalmazott munkaformákkal összefüggésben vannak. Ez idáig viszonylag kevés magyarországi kutatás irányult a tanulói beszéd, azon belül is a szóátadások vizsgálatára, így az eredmények fontos adalékkul szolgálhatnak az osztálytermi diskurzus jellemzőinek kutatásában.

2 Anyag és módszer

A kutatás összesen 8 felső tagozatos és középiskolai tanóra (1. táblázat) tanóránként 1 kamerás és 2 diktafonos felvételeinek, valamint lejegyzett anyagának az elemzése alapján mutatja be a tanulói szóátadások fő jellemzőit az osztálytermi diskurzusból.

A vizsgálat a tanulói szóátadásokat a következő paraméterek alapján jellemzi: a tanulói és a tanári beszédidő aránya a frontális munka arányához viszonyítva (%), a tanulói szóátadások típusainak aránya (%), a tanulói szóátadások irányának aránya (%); a tanulói szóátadások mennyisége (db), a szóátadások időtartama (s), a szóátadások átlagos hossza (s), a tanulói szóátadások egyes típusainak MLU-száma (= Mean Length of Utterance; a megnyilatkozások átlagos hosszának mérője, a szavak száma elosztva a megnyilatkozások számával Mabel et al. 2010; Neuberger 2014), a szóátadások pragmatikai funkciója, a tanulói szóátadások statisztikai elemzése. A kutatás eredményei az ELAN 5.1. transzkripció programmal (ELAN 2017) és az SPSS 17.0 (SPSS Inc. 2008) szoftverrel készültek. A Kolmogorov–Smirnov normalitáspróba igazolta, hogy az adatok minden esetben normál eloszlásúak. A Kendall's Tau-b korrelációs vizsgálat segítségével feltárható volt, hogy a szóátadás típusát milyen tényezők befolyásolják.

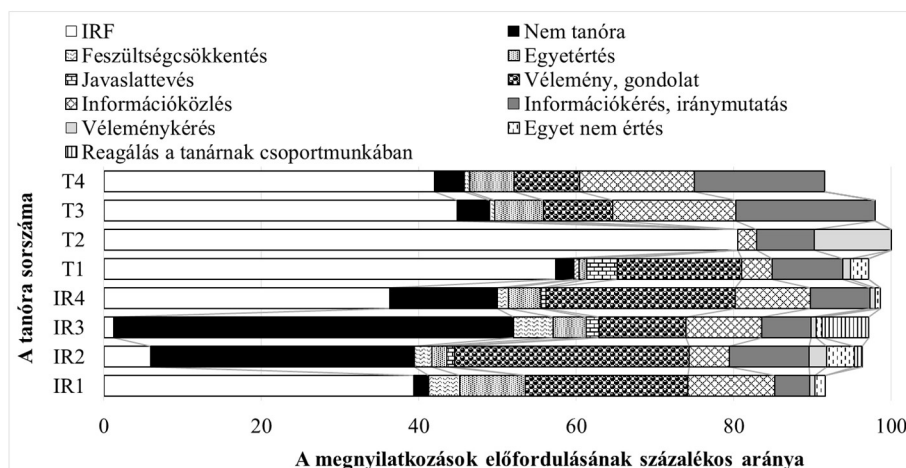
A jelen kutatást megelőzte egy – ugyanezen a korpuszon végzett – interakcionális kommunikációs szempontú elemzés (Bales 1951), amely a tanulók összes, a tanórához kapcsolódó megnyilatkozásait beszédaktusok szerint csoportosította (1. ábra) (Asztalos

2019). A feltárt megnyilatkozástípusok közül az IRF-szekvenciariész 6 tanórán is a legnagyobb arányú tanulói megnyilatkozás. A vélemény és a gondolatok megfogalmazása az IR2 számú irodalomórán domináns, 33%-ban fordult elő. Ez az előzetes vizsgálat hozzájárult ahhoz, hogy a tanulói szóátadások megállapítása a teljes kontextus figyelembevételével történjen.

A típusok megállapításának alapja az előzetes interakcionális kommunikációs elemzés és a kontextuális tényezők megfigyelése, valamint Király (2015) tanári szóátadásokra vonatkozó tipologizálása. A tanulói szóátadásokat – egy korábbi kutatás nyomán – nem vizsgáltam azokban a megnyilatkozásokban, amelyek nem kapcsolódtak a tanóra témájához, valamint a szóátadások körébe nem vettem fel azokat a megnyilatkozásokat, amelyek különféle háttér-csatorna-jelzéseknek feleltethetők meg. Ezekkel a megnyilatkozásokkal valódi szóátadás nem történik, általában párhuzamosan jelentek meg az egyéb megnyilatkozásokkal, fő pragmatikai funkcióik a figyelem megerősítése, a visszajelzés vagy az egyetértés kifejezése volt (Dér 2012).

A tanóra sorszáma	A pedagógus neme	A pedagógus életkora	A felvétel készítésének helye	Tantárgy	Tanított korosztály
IR1	Férfi	35–40	Budapest	irodalom	középiskolai
IR2	Férfi	30–35	Budapest	irodalom	felső
IR3	Férfi	30–35	Budapest	irodalom	középiskolai
IR4	Férfi	35–40	Budapest	irodalom	középiskolai
T1	Férfi	40–45	Budapest	történelem	középiskolai
T2	Nő	40–45	Budapest	történelem	középiskolai
T3	Nő	40–45	Budapest	történelem	középiskolai
T4	Férfi	25–30	Budapest	történelem	felső

1. táblázat. A vizsgált tanórák jellemzői



1. ábra. A tanulói megnyilatkozások előfordulásának százalékos aránya

A korpuszban azokat a megnyilatkozásokat, teljes beszédfordulókat vagy beszédfordulórészeket tekintettem szóátadásnak, amelyek után ön- vagy külválasztás

hatására közvetlenül megszólal a következő beszélő, továbbá azokat az eseteket is szóátadásként értelmeztem, amelyeket követően a beszédpartner direkt (pl. felszólítás, közbevágás) vagy indirekt (pl. a beszédpartner megvárja, míg a másik befejezi a közlést) módon megszólal. A tanulói szóátadások típusait a vizsgálatban aszerint is differenciáltam, hogy a pedagógus vagy a tanulótársak irányába történik-e, mindezek alapján a következő fajtákat különítettem el:

- a) Elhallgatás a tanár vagy a diák felé irányulva: a diák a saját beszédfordulója végén elhallgat, ezt követően általában a tanár vagy egy másik diák magához veszi a szót.
 - [1] Tanár: *És ő perzsa volt. Miért? Mi volt a lényege ennek a rendszernek, és mi volt az eredménye ennek a rendszernek? Mondjad!*
Diák: *Képes ellenőrzés alatt tartani, és hogyha az egyik fel akarna lázadni, akkor a másik nem tud.*
- b) Kérdés a tanár vagy egy másik diák felé irányulva: azok a tanulói kérdések, amelyeknek alapvető célja a tájékoztatás vagy a véleménykérés mellett, hogy a tanárnak vagy egy másik diáknak átadják a szót.
 - [2] Diák: *Mindenkinek le kell írni?*
Diák: *Aha.*
- c) Kérdés egy másik diák vagy a tanár felé irányulva megszólítással: külön típusként értelmeztem azokat a kérdéseket, amelyeknek része volt a másik diák vagy a tanár név szerinti megszólítása is.
 - [3] Diák: *Na, írod, Laci?*
Diák: *Írom.*
- d) Tanári vagy tanulói közbevágás: azok a tanulói beszédfordulórészek, amelyek során a pedagógus vagy egy másik diák belevág az éppen beszélő diák szavába. Előfordulnak olyan esetek is, amelyek során a két beszélő egy rövid ideig párhuzamosan beszél.
 - [4] Diák: *És (ööö) úgy gondoltuk, hogy (ööö) ez azért lenne jó, mert így az embereknek, tehát hogy nem kéne szűrnie a hírforrásokat, hanem mindegyikben ugyanúgy tudna hinni, és az, hogy mondjuk melyik adott, vagy hogy melyik pártnak az oldalán áll, az nem befolyásolná a híreket. És...*
Tanár: *Hát hogy milyen nézetet vall, vagy nem tudom, aha.*
- e) Tanulói válasz kérdő formában: ez a fajta tanulói szóátadás az IRF-szekvencia részeként megjelenő tanulói válaszadás egy lehetséges altípusa, amely a tanár kezdeményezésére hangzik el, megfigyelhető azonban, hogy a diákok kérdő formában fogalmazzák meg a válaszaikat a pedagógus irányába, visszaadva ezzel a szót a tanárnak. Ennek a tanulói választípusnak a meghatározása korábbi szakirodalmakban nem olvasható. A jelenség egyrészt a tanulók saját tudásukban való bizonytalanságát sugallja, másrészt azt az implicit szándékot is közvetítheti, ahogy a diákok az IRF-szekvencia monotonitását megtörve próbálnak valódi kommunikációs helyzetet kialakítani a tanárral. Ez a szóátadás a szekvenciában elfoglalt helyét tekintve csak a diákok beszédét jellemzi.
 - [5] [5] Tanár: *Hm. (...) Na, mennyire érzed? Barnabás? Hány sztratégosz volt Athénban?*

Diák: *Kilenc?*

Tanár: *Kilenc arkhón volt. Tíz, tíz sztratégosz volt, tíz sztratégosz volt, és ők napi váltásban, megvolt, hogy ki éppen az első sztratégosz. Na, most ezen a jeles napon, amikor meg kellett volna ütöz, ütközni a perzsákkal, és hát nagyrészt úgy gondolták, hogy, hogy ez egy eléggé vesztes játszma lesz, hiszen több, mint kétszeres túlerőben voltak a perzsák, valahogy mindenki lemondott az aznapi vezetőségről. Miltiádész nem.*

- f) A másik diák vagy a tanár közvetlen és egy másik diák közvetett felszólítása: közvetlen felszólításként értelmeztem azokat a beszédfordulókat vagy megnyilatkozásokat, melyekben a tanulók egymást vagy a tanárt szólítják fel a szóátadás szándékával. Közvetettnek tekintettem azokat az eseteket, amikor az alapvető cél ugyancsak a diáktársak felszólítása, de ezt nem egymásnak fogalmazták meg a tanulók, hanem a tanárnak egy szituációs játék során.

[6] Diák: *Kati, ne izgasd fel magad!*

Diák: *Csak nem szeretem, ha ide nem illő szavakat dobálunk, na mindegy.*

- g) Felkiáltás: azok a megnyilatkozások, amelyek az egész kontextust figyelembevéve felkiáltásként értelmezhetők, a szóátadás szándékát is hordozzák azonban.

[7] Tanár: *Volt egy ilyen szó, már teljesen elfelejtődött, a grépfrútra mondták, azt hiszem, a citrancsot, ami egyékként szerintem jó szó egyébként.*

Diák: *Nagyon jó!*

Tanár: *De mindenki azt mondja, hogy milyen jó, és (ööö), és mégsem terjedt el. Citrancs, szerintem jó az (b). Eleve a narancs is szerintem egy nagyon jó szó, de citrancs az végképp szuper. (ööö) Sári!*

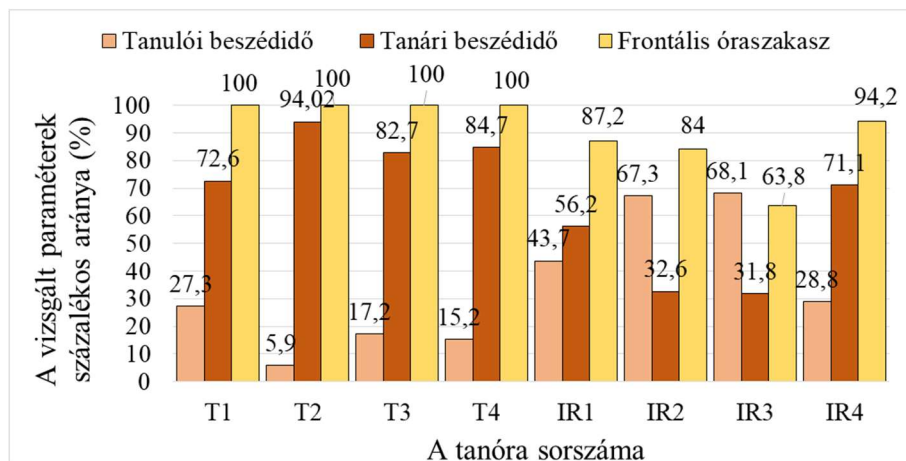
- h) A másik diák vagy a tanár megszólítása: azok a megnyilatkozások, amelyek a többi megnyilatkozástól elkülönülten jelennek meg a megszólítás, a név szerinti megnevezés és a szóátadás kifejezett szándékával.

[8] Diák: *Köszönöm, tanár úr.*

3 Eredmények

A 2. ábra a tanulói és a tanári beszédidő százalékos arányát mutatja a frontális munka tanórai százalékos arányához viszonyítva. A vizsgált történelemórák mindegyikén 100%-ban frontális munka volt, míg az irodalomórákon alacsonyabb arányban jelent meg. Az IR3 számú tanórán volt a legalacsonyabb, 63,8%. A 8 vizsgált tanórából 6 tanórán a tanár beszédidőjének aránya dominált, a T2 történelemórán ez az érték 94,02%. Az IR2 és IR3 irodalomórákon a tanulói beszédidő százalékos aránya magasabb volt a pedagógusétól, az IR2. tanórán ez a százalékos arány 67,3%, míg az IR3 tanórán 68,1% volt. A Kendall's Tau-b korrelációs vizsgálat – a korábbi kutatásokkal egybehangzóan – alátámasztja, hogy a frontális óraszakasz százalékos aránya és a tanulói beszédidő százalékos aránya között negatív irányú korreláció van ($T = -0,886$; $p = 0,003$). Eszerint a frontális munkaforma aránya negatívan befolyásolja a tanulói beszédidő százalékos megoszlását, vagyis minél nagyobb a frontális munkaforma, annál kisebb a tanulói beszédidő aránya.

A korpuszban azonosított tanulói szóátadások típusait és százalékos arányát mutatja a 2. táblázat. A tanulói szóátadások közül a legmagasabb arányban a tanár (46,6%), majd a tanuló felé irányuló elhallgatás (20,7%) jelent meg. A diákok egymás irányába megfogalmazott kérdéseinek feltevése 11,4%, míg a tanár irányába megfogalmazott kérdések aránya 5,1% volt. A kérdő formában megfogalmazott tanulói válaszok aránya 3,5% volt.



2. ábra. A tanulói és a tanári beszéidő százalékos aránya a frontális munkaforma százalékos arányához viszonyítva

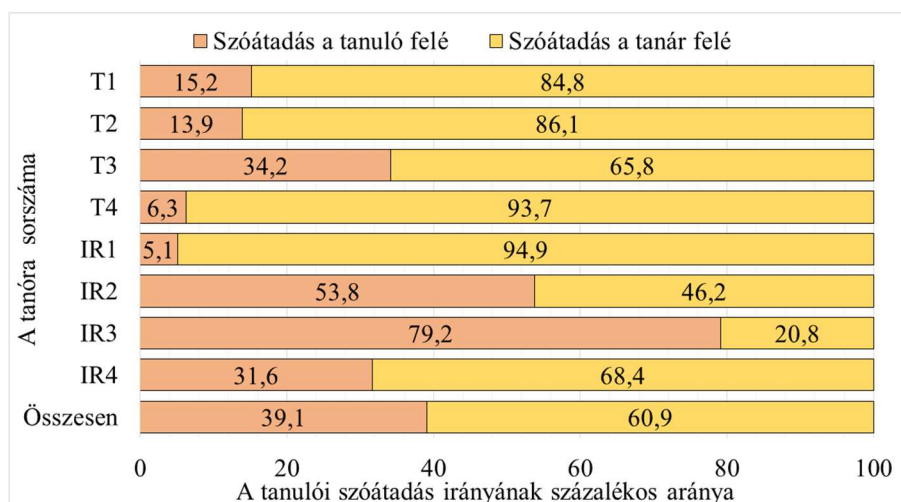
A típus neve	Százalékos arány
Elhallgatás a tanár irányába	46,6
Elhallgatás a diák irányába	20,7
Kérdés a diák irányába	11,4
Kérdés a tanár irányába	5,1
Tanári közbevágás	3,9
Tanulói válasz kérdő formában	3,5
A másik diák közvetlen felszólítása	2,7
Tanulói közbevágás	2,5
A másik diák közvetett felszólítása	1,3
Kérdés egy másik diák irányába megszólítással	1,0
Felkiáltás	0,6
A másik diák megszólítása	0,2
A tanár megszólítása	0,2
Kérdés a tanár irányába megszólítással	0,2
A tanár felszólítása	0,2

2. táblázat. A tanulói szóátadások azonosított típusai és százalékos arányuk

A tanulói szóátadások irányának százalékos megoszlását mutatja a 3. ábra. Az összes vizsgált tanórát tekintve a diákok a szót 60,9%-ban a pedagógusnak, míg 39,1%-ban a többi tanulónak adták át. Az eredmények alapján megállapítható, hogy két tanóra kivételével minden vizsgált órán a tanárnak történő tanulói szóátadás százalékos aránya volt

nagyobb. A tanulói szóátadás a diáktársaknak az IR1 tanórán jelent meg a legalacsonyabb százalékban (5,1%). A korrelációs vizsgálat alátámasztja, hogy a tanulók egymásnak történő szóátadásának százalékos arányát negatívan befolyásolja a tanári beszédidő százalékos aránya ($\tau = -0,429$; $p = 0,138$), valamint a frontális munkaforma százalékos aránya is negatívan hat rá ($\tau = -0,403$; $p = 0,184$).

A további ábrák azoknak a tanulói szóátadásoknak az eredményeit mutatják, amelyeknek a százalékos aránya az összes vizsgált tanórát tekintve nagyobb vagy egyenlő volt, mint 1%. A típusok közül az elhallgatás volt a legtöbbször előforduló szóátadás. A T1 történelemórán a tanár irányába történő elhallgatások száma 141 db, az IR2 tanórán 114 db volt. A diákok irányába történő elhallgatások száma az IR3 tanórán volt a legtöbb, 127 db. A T1 történelemórán volt a legnagyobb a tanár irányába elhangzó kérdések (21 db) és a „tanulói válaszok kérdő formában” típus száma (17 db). Ezen a tanórán olyan tanulói válasz is elhangzott, amely egy szituációs játékban jelent meg, a tanár kérdésére válaszolva, a jelentése azonban arra irányult, hogy mit csináljanak a diáktársak – ez a szóátadási típus a diákok közvetett felszólítása (14 db). A tanár közbevágására történő tanulói szóátadás egy tanóra (IR4) kivételével minden tanórán megjelent, a legnagyobb számban az IR3 tanórán 21 alkalommal. A tanulók egymás szavába főként azokon a tanórákon vágtak, ahol a munkaforma lehetővé tette ezt. Az IR2 tanórán 15 db és az IR3 tanórán 11 db ilyen szóátadás figyelhető meg (4. ábra). A Kendall's Tau-b korrelációs vizsgálat alátámasztja, hogy a tanulók egymásnak megfogalmazott kérdéseinek darabszámát negatívan befolyásolja a tanár felé irányuló szóátadások százalékos aránya ($\tau = -0,964$; $p = 0,001$).

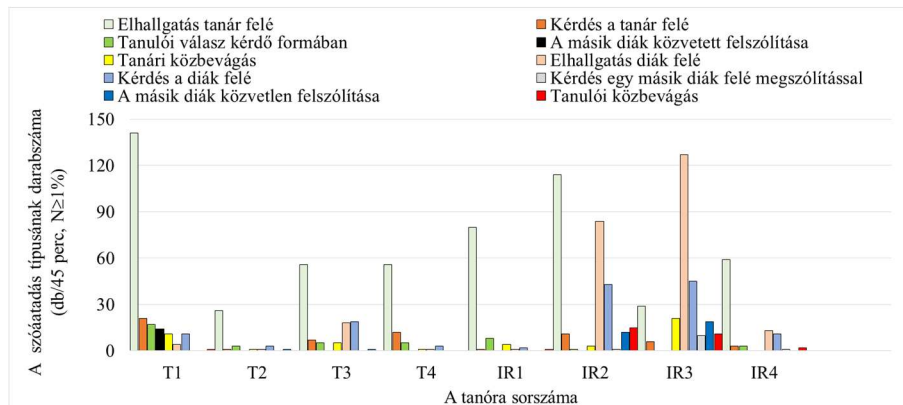


3. ábra. A tanulói szóátadás irányának százalékos aránya a vizsgált tanórákon

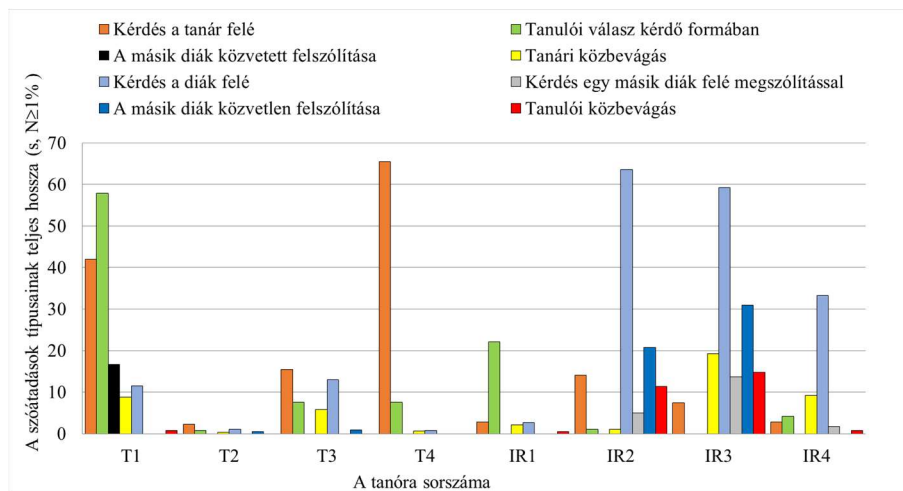
A tanulói szóátadások teljes hosszát mutatja az 5. ábra. A szóátadások közül a tanárnak megfogalmazott kérdések voltak a leghosszabbak, a T4. számú történelemórán 65,5 s. A diákoknak megfogalmazott kérdések hossza az IR2 tanórán 63,6 s, míg az IR3 tanórán 59,25 s volt. A tanulók válaszainak kérdő formában való megfogalmazása a T1 történelemórán volt a leghosszabb (57,9 s). A diákok felszólítása azokon a tanórákon

jelent meg, ahol a munkaforma lehetővé tette az egymással való változatosabb kommunikációt. A korrelációs vizsgálat alátámasztja, hogy a tanulók egymás felé irányuló kérdéseinek teljes hosszát negatívan befolyásolja a tanári beszédidő százalékos aránya ($\tau = -0,571$; $p = 0,048$), valamint a tanár felé irányuló szóátadások százalékos aránya is ($\tau = -0,714$; $p = 0,013$).

A tanulói szóátadások átlagos hosszát mutatja a 6. ábra. Az IR4 számú irodalomórán a szóátadások típusa közül a tanár közbevágása miatti tanulói szóátadás volt átlagosan a leghosszabb, 9,2 s. A tanulók kérdő formában megfogalmazott válaszai a T1 történelemórán voltak a leghosszabbak (3,8 s), míg az IR1 irodalomórán a legrövidebbek (2,9 s). A korpuszban csak az IR3 tanórán nem jelent meg ez a típusú szóátadás. A tanárnak megfogalmazott kérdések típusa a T1 történelemórán volt a leghosszabb időtartamú, átlagosan 41,9 s, míg a legrövidebb az IR4 számú irodalomórán volt, 0,8 s.



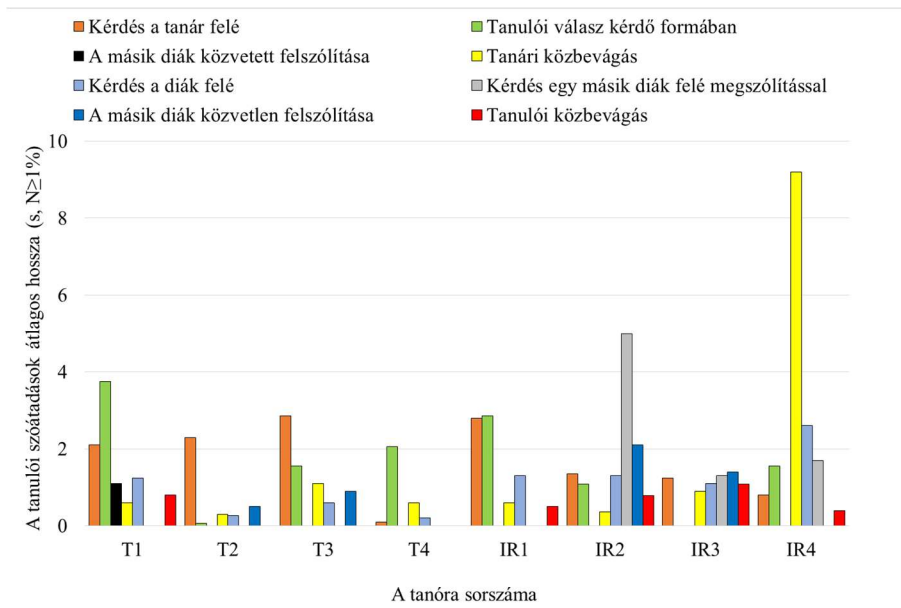
4. ábra. A tanulói szóátadások eloszlása (db)



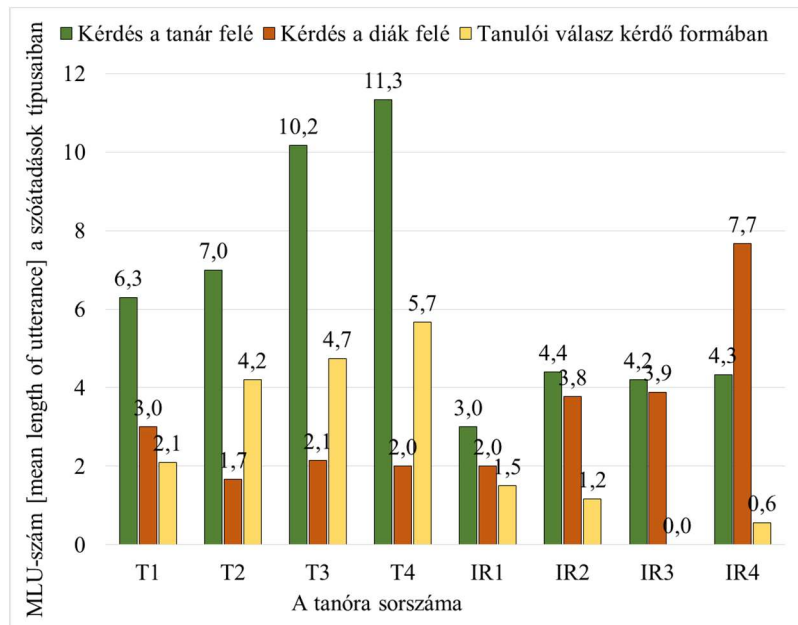
5. ábra. A tanulói szóátadások teljes hossza

Három, a százalékos arányokat tekintve gyakran előforduló szóátadási típus MLU-számát mutatja a 7. ábra. Az ábrán jól látható, hogy a tanárnak megfogalmazott

kérdések MLU-száma azokon a tanórákon volt több, amelyeken frontális munka zajlott, a legtöbb a T4 számú történelemórán (11,3). A kérdő formában megfogalmazott válszok MLU-száma a T1 számú történelemórán volt a legkimagaslóbb. A diákok egymás felé irányuló, szóátadást célzó kérdéseinek MLU-száma az IR4 számú irodalomórán volt a legnagyobb (7,7). A statisztikai elemzés alátámasztja, hogy a diákok beszédidejének százalékos arányának növelése pozitív hatással van az egymás számára megfogalmazott kérdéseik MLU-számára ($\tau = -0,618$; $p = 0,034$), valamint a frontális munkaforma százalékos aránya pozitívan befolyásolja a tanár irányába megfogalmazott kérdések MLU-számát ($\tau = -0,645$; $p = 0,034$).



6. ábra. A tanulói szóátadások átlagos hossza



7. ábra. A szóátadások MLU-száma (Mean Length of Utterance)

4 Összegzés és következtetések

A jelen kutatás eredményei a korábbi kutatások eredményeivel egybehangzóan alátámasztják, hogy a tanulói beszédidő százalékos arányát befolyásolja a frontális munkaforma százalékos aránya. Ezáltal az eredmények alátámasztják az első hipotézist, hiszen a vizsgált tanórákból 6 tanórán a tanár beszéddidejének aránya volt nagyobb. A statisztikai próba megerősíti, hogy a frontális munkaforma aránya negatívan befolyásolja a tanulói beszédidő százalékos megoszlását.

A kutatás alátámasztja azt az előfeltevést is, miszerint vannak olyan szóátadási módok, amelyek tanulóspecifikusak. A tanulói válasz kérdő formában már az IRF-szekvenciában való helye alapján is csak a tanulói beszédet jellemezheti, és a korpuszban főként a tanulók bizonytalanságát, valamint a kétoldalú kommunikáció ösztönzését támasztja alá. Nagyobb korpuszban való elemzés során feltételezhető, hogy összefüggést mutatna a frontális munkaforma százalékos arányával.

A vizsgálat annak ellenére, hogy kamerás és diktafonos felvételek komplex elemzése és lejegyzése nyomán készült, nem támasztja alá azt a hipotézist, hogy a tanulói szóátadások fajtái dominánsan nem nyelvi. A korpusz alapján az azonban megállapítható, hogy számos nem nyelvi elem kíséri a tanulói szóátadásokat, így például: a szemkontaktus felvétele, gesztusok, a testtartás megváltoztatása stb.

A kutatás alátámasztja a negyedik hipotézist azzal, hogy differenciálta a tanulói szóátadások lehetséges típusait. 15 tanulói szóátadástípus elkülönítésével kísérletet tett arra, hogy a szóátadás fogalmának meghatározása árnyaltabbá váljon. A szóátadásokat aszerint is csoportosította, hogy a tanulók a tanár vagy a diáktársak felé irányulva adják át a szót. Az eredmények megerősítették az ötödik hipotézist, a nyolc vizsgált tanórából

hat tanórán a tanárnak történő tanulói szóátadás százalékos aránya volt nagyobb. A statisztikai próba alátámasztja, hogy a tanulók egymás felé irányuló szóátadására negatív hatással van a tanári beszédidő százalékos aránya, valamint a frontális munkaforma százalékos aránya is. A vizsgálat ezzel megerősíti a hatodik előfeltevést is, mely szerint a tanulói szóátadások típusai az órán alkalmazott munkaformákkal összefüggésben vannak. A kutatás fő célja volt, hogy rámutasson arra, hogy a tanulói szóátadásokat milyen tényezők befolyásolhatják. A statisztikai vizsgálat alátámasztja, hogy a tanulók egymás felé történő szóátadásának százalékos arányára negatívan hat a tanári beszédidő százalékos aránya, valamint a tanulók kérdéseinek darabszámát negatívan befolyásolja a tanár felé irányuló szóátadások százalékos aránya. Továbbá a tanulók egymás felé irányuló kérdéseinek teljes hosszát negatívan befolyásolja a tanári beszédidő százalékos aránya, valamint a tanár felé irányuló szóátadások százalékos aránya is. A statisztikai elemzés alátámasztja, hogy a diákok beszédidőjének növelése pozitív hatással van az egymás számára megfogalmazott kérdések MLU-számára, valamint a frontális munkaforma százalékos aránya pozitívan befolyásolja a tanár irányába megfogalmazott kérdések MLU-számát is.

Összegzésként elmondható, hogy bár a vizsgálat korpusza mindösszesen 8 tanóra felvételének lejegyzett és elemzett anyagát tartalmazza, az eredmények és a következtetések mégis hasznosak lehetnek a tanulói beszéddel és az osztálytermi diskurzusokkal kapcsolatos kutatásokban. A jelen vizsgálat emellett olyan irányokat is kijelölhet (például a beszédjog eloszlásának kutatását vagy a diskurzusjelölők szerepének feltárását a szóátvétel és szóátadás során stb.), amelyeket nagyobb mintán érdemes elemezni annak céljából, hogy teljesebb képet kaphassunk a tanulók beszédéről és azokról a tényezőkről, amelyek befolyásolják.

Köszönetnyilvánítás

A jelen vizsgálat az Emberi Erőforrások Minisztériuma ÚNKP-18-3 kódszámú Új Nemzeti Kiválóság Programjának támogatásával készült.



EMBERI ERŐFORRÁSOK
MINISZTERIUMA

Irodalom

- Antalné Szabó Á. 2006. *A tanári beszéd empirikus kutatások tükrében*. A Magyar Nyelvtudományi Társaság Kiadványai. 225. Budapest: Magyar Nyelvtudományi Társaság.
- Antalné Szabó Á. 2010. A tanári megnyilatkozások típusai. In: Kozmács I. – Vanconé Kremmer I. (szerk.) *Közös jövőnk a nyelv I.* Nyitra: Konstantin Filozófus Egyetem. 5–18.
- Antalné Szabó Á. 2015. Osztálytermi kommunikáció. In: Antalné Szabó Á. – Major É. (szerk.) *Mozaikok a magyar nyelvről és nyelvhasználatról*. Budapest: Eötvös Loránd Tudományegyetem. 57–70.

- Asztalos A. 2015. A tanórai kifejezőképesség fejlesztésének eredményei egy empirikus kutatás tükrében. In: Antalné Szabó A. – Major E. (szerk.) *Tanóratervezés és tanóratudományok*. Budapest: Eötvös Loránd Tudományegyetem. 103–127.
- Asztalos A. 2019. A csoportkommunikációs stratégia vizsgálata a tanulói beszédben magyar nyelvű tanórák alapján. In: Karlovitz János Tibor – Torgyik Judit (szerk.) *Szaktudományok és más emberközpontú tanulmányok*. Komárom: International Research Institute. 71–84.
- Bales, R. F. 1951. *Interaction process analysis; a method for the study of small groups*. Addison – Wesley Press. Cambridge.
- Bartha Cs. 1998. *A szociolingvisztika alapjai*. Budapest: Eötvös Loránd Tudományegyetem Bölcsészettudományi Kar.
- Bartha Cs. – Hámori Á. 2010. Stílus a szociolingvisztikában, stílus a diskurzusban. Nyelvi variabilitás és társas jelentések konstruálása a szociolingvisztika „harmadik hullámában”. *Magyar Nyelvőr* 134(3): 298–321.
- Bata S. 2009. Beszélőváltások a beszédpartneres személyes kapcsolatának függvényében. *Beszédkutatás* 2009: 107–120.
- Bentley, K. 2007. STT: Student talking time. How can teachers develop learners' communication skills in a secondary school CLIL programme. In: Francisco, L. – Sonia Casal, M. – Virginia de Alba, Q. – Pat, M. (szerk.) *Revista Española de Lingüística Aplicada. Models and Practise in CLIL*. Logroño: Reproestudio. 129–139.
- Bóna J. 2015. 6–13 éves iskolások megakadása különböző beszéd típusokban. *Anyanyelv-pedagógia*. 2. Elérhető: <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=569> (letöltés ideje: 2019. 02. 10.)
- Boronkai D. 2013. A beszélőváltások és a szociokulturális tényezők összefüggései társalgási szövegekben. *Magyar Nyelvőr* 137(1): 68–87.
- Clark, H. H. 2001. Conversational Linguistic Aspects. In: Neil, J. S. – Paul, B. B. (szerk.) *International Encyclopedia of the Social & Behavioral Sciences*. Elsevier Science Ltd. Berlin: Palo Alto. 2744–2747.
- Dér Cs. I. 2012. Beszélőváltások során használt diskurzusjelölők a magyar spontán beszédben. *Beszédkutatás* 2012: 132–143.
- ELAN 2017. Linguistic Annotator Version: 5.2 Max Planck Institute for Psycholinguistics Nijmegen, The Netherlands.
- Gósy M. 2003. Virtuális mondatok a spontán beszédben. *Beszédkutatás* 2003: 19–44.
- Grice, H. 1975/1997. A társalgás logikája. In: Pléh Cs. – Terestyéni T. – Síklaki I. (szerk.) *Nyelv-kommunikáció – cselekvés*. Budapest: Osiris Kiadó. 213–227.
- Hámori Á. 2006. A társalgási műfajokról. In: Tolcsvai Nagy G. (szerk.) *Szöveg és típus*. Budapest: Tinta Könyvkiadó. 157–181.
- Herbst M. 2006. A megszólalások esélyei a tanítási órán. In: Cseh B. (szerk.) *Nyelv és nyelvtudomány*. Baja: Eötvös József Főiskola.
- Herbst M. 2008. Grice társalgási maximáinak érvényesülése a tantermi interakciókban. A tanári kérdéses stratégiák egyéni sajátosságai. *Anyanyelv-pedagógia* 3–4. Elérhető: <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=88> (letöltés ideje: 2018. 09. 10.)
- Horváth V. 2017. Közlélek grammatikai szerkesztettsége 6–9 éves gyermekek narratíváiban. *Anyanyelv-pedagógia* 4. Elérhető: <http://anyanyelv-pedagogia.hu/cikkek.php?id=703> (letöltés ideje: 2018. 12. 15.)
- Hymes, D. 1974. A nyelv és a társadalmi élet kölcsönhatásának vizsgálata. In: Pléh Cs. – Síklaki I. – Terestyéni T. (szerk.) 1997. *Nyelv – kommunikáció – cselekvés*. Budapest: Osiris Kiadó. 458–495.
- Iványi Zsuzsanna 2003. A „nyelvemen van”-jelenség német és magyar nyelvű beszélgetésekben. Nyelvészeti vizsgálatok a konverzációelemzés módszereivel. *Magyar Nyelvőr* 127: 76–91. Elérhető: http://real-j.mtak.hu/6079/1/MagyarNyelvor_2003.pdf (letöltés ideje: 2019. 06. 29.)
- Kelemen Széll Zs. 2012. Kooperatív tanulás – tantárgyi attitűdök. *Anyanyelv-pedagógia* 2. Elérhető: <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=381> (letöltés ideje: 2018. 12. 15.)
- Király F. 2015. A beszédfordulók és a szóátadások vizsgálata osztálytermi kontextusban. *Anyanyelv-pedagógia* 1. 19–31. Elérhető: <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=552> (letöltés ideje: 2018. 11. 12.)
- Labov, W. – Fanshel, D. 1997. Beszélgetési szabályok. In: Pléh Cs. – Síklaki I. – Terestyéni T. (szerk.) *Nyelv – kommunikáció – cselekvés*. Budapest: Osiris Kiadó. 395–435.
- Mabel, L. R. – Smolik, F. – Perpich, D. – Thompson, T. – Rytting, N. – Blossom, M. 2010. Mean length of utterance levels in 6-month intervals for children 3 to 9 years with and without language impairments. *Journal of Speech, Language, and Hearing Research* 53: 333–349.
- Nemzeti alaptanterv 2012. Oktatási és Kulturális Minisztérium. Budapest.

- Neuberger T. 2014. *A spontán beszéd sajátosságai gyermekkorban*. Budapest: ELTE Eötvös Kiadó.
- Pléh Cs. 2014. A társalgás pszichológiája. Pléh Cs. – Lukács Á. (szerk.) *Pszicholingvisztika 1 – 2. Magyar pszicholingvisztikai kézikönyv*. Budapest: Akadémiai Kiadó. 1383–1404.
- Rosengern K. E. 2004. *Kommunikáció*. Budapest: Typotex Kiadó.
- Sacks, H. – Schegloff, E. – Jefferson, G. 1974. A simplest systematics for the organisation of turn taking for conversation. In: Scheinken, J. (szerk.) *Studies in the Organisation of conversational interaction*. New York: Academic Press. 7–55.
- Samei, B. – Olney, A. – Kelly, S. – Nystrand, M. – D’Mello, S. – Blanchard, N. – Graesser, A. 2015. Modeling classroom discourse: Do models of predicting dialogic instruction properties generalize across populations? In: Santos C., O. – Boticario G. J. – Romero, C. – Pechenizkiy, M. – Merceron, A. – Mitros, P. – Luna, J. M. – Mihaescu, C. – Moreno, P. – Hershkovitz, A. – Ventura, S. – Desmarais, M. (szerk.) *Proceedings of the 8th International Conference on Educational Data Mining*. Madrid: International Educational Data Mining Society. 444–447.
- Schlegloff, E. A. 2015. Conversational Interaction. The Embodiment of Human Society In: Schrifrin D. – Tannen, D. – Hamilton, Heidi E. (szerk.) *The Handbook of Discourse Analysis*. Oxford: Wiley Blackwell. 346–366.
- Skidmore, D. – Perez-Parent, M. – Arnfiel, S. 2003. *The quality of teacher-pupil dialogue in guided reading*. Paper presented at the British Educational Research Association Annual Conference. Edinburgh: Herio-Watt University. 47–53.
- SPSS Inc. Released 2008. SPSS Statistics for Windows, Version 17.0. Chicago: SPSS Inc.
- Tátrai Sz. 2004. A kontextus fogalmáról. *Magyar Nyelvőr* 128(4): 479–494.
- Tátrai Sz. 2011. *Bevezetés a pragmatikába*. Budapest: Tinta Könyvkiadó.
- Walsh, S. 2006. *Investigating Classroom Discourse*. New York: Taylor & Francis e-Library.
- Williams, M. – Wright, T. 1991. Classroom interaction. In: Ur, P. 1991. *A Course in Language Teaching*. Cambridge: Cambridge University Press. 227–241.

A játéknyelvi szókincset befolyásoló tényezők

Balogh Andrea

ELTE BTK Nyelvtudományi Doktori Iskola
balogh.andrea.elte@gmail.com

Kivonat: A videójátékok szókincse, bár egyre gyakrabban kerül be a köznyelvi formák közé (például *avatár*, *színlépés*), még döntően egy zárt csoportnyelvhez tartozik, amelyen belül a különböző játékok eltérő szókészlettel rendelkeznek. A hosszú távú célom egy olyan alapkutatás összeállítása, ami segít bemutatni és előkészíteni a további nyelvi kutatókhoz ezt a nyelvileg izgalmas és sokszínű szubkultúrát. A videójátékok nyelvi környezetének sajátosságait figyelembe véve azt a kutatási kérdést tettem fel ebben a tanulmányban, hogy tetten érhetőek-e szókincset befolyásoló tényezők a játékosok és a tartalomgyártók írott szövegeiben. Ennek érdekében a gamer kultúrához kapcsolódó online cikkeket és Youtube-kommenteket vizsgáltam egyrészt egy szólista alapján, másrészt statisztikai módszerekkel. A külső befolyásoló tényezők közül a nyelvválasztást és a játéktípust emeltem ki, mivel ez a kettő megjelenhet áttételesen a játékosok más területen zajló kommunikációjában is.

1 Bevezetés

A videójátékok egyre nagyobb teret hódítanak a mindennapokban, részben a mobil eszközök és az internet elterjedésének következtében. Egy 2018-as kutatás a magyarországi videójátékosokat 3,7 millió főre becsüli (eNET 2019). A megkérdezés alapján a férfiak közül valamennyivel többen játszanak (62% a nők 52%-ához képest), és bár dominánsan a 18–25 év közöttiek foglalkoznak ezzel a hobbival (66%), az idősebb korosztályokban is 50% feletti az érték. A legtöbben az internetes játékok közül a kártyajátékokat választják, majd a stratégiai játékokat, végül következnek a szó-, a kirakós és a kalandjátékok. A legtöbben okostelefonon játszanak (65%), ezután következik a PC és a laptop (43 és 38%). A 18–65 év közöttiek 47%-a hallott már az e-sportról is, amely a leggyorsabban fejlődő és a mainstream médiában is megjelenő területe a videójátékoknak (eNET 2019).

A videójátékokkal kapcsolatos tudományos kutatások során több tényezőt is figyelembe kell vennünk az elemzéskor, így például a hazai játékosközösség jellemzőit, a játéknyelvi kommunikáció általános jellegzetességeit, valamint az angol és a magyar nyelv viszonyát a gamerszövegekben.

1.1 A játéknyelvi kommunikáció

A játéknyelvi szókincs elemeinek kategorizálására több szinten és több nézőpontból is lehetőség van. A közös pont, hogy az általános játékelemek legkevésbé speciális kifejezéseihez férnek hozzá a legtöbben (a *casual* 'alkalmi' és a *hardcore* 'komoly, a játékba nagy erőfeszítést ölt' játékosok egyaránt), és ahogy haladunk az egyedi játékok irányába, egyre kisebb lesz az a beszélőközösség, amely tisztában van a kifejezés pontos jelentésével (Balogh 2017).

A korábbi modellemben a nyelvhasználók oldaláról közelítettem meg a kérdést: ebből az volt látható, hogy a jelentős különbségek a játékosok és a nem játékosok mellett a különböző játéktípusokkal foglalkozó gamerek között voltak (bővebben lásd Balogh 2017: 171–172).

Ensslin a 2012-es *The language of gaming* című munkájában foglalta össze először, hogy milyen csoportokra bontható a számítógépes játékokkal kapcsolatos szókincs. Az összegzés alapját több, sajnos ma már nem működő, kifejezetten játéknyelvre specializálódott honlap gyűjteménye és a hackernyelvvel való párhuzam alapján állította össze a szerző. Az elmélet alapját a számítógépes játékokhoz kapcsolódó, különböző szociolektusok összehangolása adja. Ensslin komplexen, egy modellen belül jeleníti meg a játékokkal kapcsolatos szakmai, a játékos és a kívülálló, de kapcsolódó közösségek szóhasználatát, összekapcsolva az elméletet azzal a másik nagy kérdéssel is, hogy hol, milyen környezetben találkozhat valaki a játékokban és a játékokkal kapcsolatban használt speciális szókincsel. Ennek megfelelően a különböző nyelvhasználói csoportok alapján öt fokozatra osztotta a kifejezéseket: a fejlesztői zsargonra, a ludológiai, játékkutatási zsargonra, a ludolektusokra, vagyis a játéknyelvekre, a játékok reklámozásával kapcsolatos kifejezésekre és az általános, köznyelvi médiadiskurzus körébe tartozó szavakra (Ensslin 2012: 67). Az osztályozás közepén elhelyezkedő ludolectus-nek (játékosnyelveknek) nevezett csoport nagyon vegyes terület, átmenetet képez a zsargon és a köznyelv között: egy része közelebb áll a szakmai szöveghez (osztottnak például különböző technikai kifejezéseken és kommunikációs formákon), ugyanakkor a köznyelvhez közelebb álló kifejezések ebből a körből kerülnek ki, tehát a köznyelvi réteghez is kapcsolódik. Nagyban befolyásolja ennek a csoportnak a tartalmát, hogy melyik játéktípusról és játékról szól a diskurzus (Ensslin 2012: 66). Ensslin ezt a csoportot további három területre bontotta: a játék egyedi szókincsén túl hozzátartozik a játéktípusról szóló ismeretek köre és az általános játéknyelv ismerete is (az osztályozás részletes bemutatását és értékelését lásd Balogh 2017).

A videójátékok nyelve a digilektus körébe sorolható be (Veszelszki 2017), ennek egy speciálisan elkülöníthető, tematikus csoportja a gamer szubkultúrára jellemző nyelvváltozat. A digilektus lexikai-grammatikai jellemzői közül az informatikai szókincs átvétele, az idegen nyelvi hatás, illetve a nagyszámú neologizmus egyaránt megfigyelhető a gamerszövegeken belül (Balogh 2013; Veszelszki 2017: 90).

A játéknyelvet elsősorban hobbinyelvként, társadalmi nyelvváltozatként tartják számon. Használata eleinte erőteljes szubkulturális körben zajlott, később a számítógépek és egyéb informatikai eszközök, valamint természetesen az internet elterjedésével elemei egyre szélesebb körben is ismertté váltak, bár a mai napig elég zárt csoportnyelvnek tekinthető (ezt részben az is indokolja, hogy a különböző játéktípusoknak, játékoknak más az alapszókincse) (Balogh 2014: 350).

A játéknyelv vizsgálatánál két, eltérő jellegzetességekkel rendelkező kutatási irányt szükséges megkülönböztetni. Egyrészt a játékról szóló kommunikációt, ami hagyomá-

nyos keretek között történik, viszont a tárgya egy videójáték vagy egy azon belüli cselekvés, a célja ennek játékkörnyezeten kívüli megvitatása (például egy beszélgetés, egy játékosmagazinbeli cikk). Másrészt a játékon belüli kommunikációt, amely magába foglalja a játékosok ingame (játékmeneten belüli) kommunikációját egymással és a játékkal, és a célja a játékban való előrejutáshoz nélkülözhetetlen információk megszerzése, továbbítása (például egy játékon belüli chatbeszélgetés).

Carrillo Masso 2009-es tanulmányában külön kitért egy, kifejezetten a számítógépes játékok vizsgálatával kapcsolatos gyakori hibára, miszerint a kutatók úgy vonnak le következtetéseket, és állítanak fel a játékokkal kapcsolatos modelleket, hogy a kutatók tárgyaül szolgáló játékokkal (vagy akár a játéktípuson belül hasonló darabbal) egyszer sem játszottak (Carrillo Masso 2009: 150), így nem ismerik a játék és a játékosok kapcsolódó szókincset, kommunikációs szokásait. Ez már önmagában hibalehetőséget vet fel, ám kiküszöbölhető a kutatási metódusba előre betervezett játékorákkal. Carrillo Masso mindezek mellett felhívja a figyelmet arra is, hogy a játékok elemzésekor figyelembe kell venni, hogy a kutató egy már ismert vagy éppen először játszott videójátékot dolgoz fel. Ő kifejezetten általa ismeretlen játékokat választott, de olyan játéktípusokból, amelyeknek az irányítását ismerte, hogy a friss, de folyamatos játékélményt tudja feldolgozni (Carrillo Masso 2009: 150).

1.2 A szókincset befolyásoló tényezők

A számítógépes játékokban zajló és a róluk szóló kommunikációt több tényező is befolyásolta, befolyásolja a mai napig. Három csoportba osztottam őket annak alapján, hogy a játékosoknak mekkora ráhatásuk van az egyes okokra: a külső tényezők esetén nincs lehetőségünk érdemben beleszólni, a belső tényezők a játékosokon és a játékosok közösségein alapulnak, a kettő közötti blended területen pedig mindkettő hatása érvényesül (lásd 1. táblázat). Ennek a tanulmánynak a célja, hogy részletesebben vizsgálja a szókincs esetében a külső hatásokat.

Külső tényezők	Blended tényezők	Belső tényezők
A technikai fejlődés (konzol, pc, chat)	A játékról szóló cikkek (e-sport, esport)	A szerver, a csapat befolyása (csírek, sutyi, kek)
A játék típusa, témája (tank, split, karakter, kaszt, azeroth)	A játékközvetítések	A szociokulturális jellemzők
A játék nyelve (ward, őrszem; lanere)	A közösségi oldalakon megjelenő tartalom (wololo)	A játékos céljai

1. táblázat. A játéknyelvi szókincset befolyásoló tényezők

A külső tényezők olyan okok, amelyeket a játékosok nem vagy csak a fejlesztők által megadott keretek között tudnak megváltoztatni. Ide tartoznak a technikai fejlődéshez kapcsolódó szavak, a játék típusa és témája által behatárolt kifejezések és a játék nyelve. Az első kategória leírja, hogy milyen platformon lehet használni a játékot, meghatározva az egyéb technikai paraméterek lehetőségeit (például az okostelefonoknak érintőképernyőjük van, a pc-hez tartozik egér és billentyűzet). Ahogy arról is a játékos dönt, hogy milyen eszközön játszik, úgy arról is ő határoz, hogy milyen játéktípust és témát helyez előtérbe vagy kerül el mindenképpen. Viszont egy adott témához (például

western) adott kifejezések (például *cowboy*) tartoznak, és ez a játéktípusokra is hasonlóképpen jellemző (a *platform*erekben – *ugrálós játékok* – különböző *akadály*okon kell keresztül menni). A videójátékok alapvetően angol nyelvűek, ehhez hivatalos fordítás ritkán, rajongók által előállított szöveg gyakrabban készül. A játékos arról dönthet, hogy a rendelkezésére álló nyelvek közül melyiken szeretné futtatni a programot.

A technikai, informatikai szókincs hosszú ideig együtt fejlődött a játékosok szókincsével, mivel a számítógép-programozás és a számítógépes játék témakörök hagyományosan egy érdeklődési körbe tartoztak. A különböző technikai újítások bevezetése, és ezzel egy időben a régebbi technológiák és eszközök kikopásának háttérbe szorulása erősen hat a szókincsre. Vannak olyan eszközök, amelyek bevezetésük óta ugyanazon a néven szerepelnek, ilyen például a (számítógépes) *egér*, amelyet 1963-as megjelenése óta így hívunk, bár eleinte *bogárnak* nevezték, de végül maradtak az alaki hasonlóság miatt a ma is elterjedt változatnál (W1). Egyes kifejezések elavultak, és ma már szinte nosztalgiával emlegetik őket, ilyen például a *C64* rövidítés, a Commodore 64-es számítógép elnevezése, amely az 1980-as évek kultikus személyi számítógépe volt, és 1982-ben megkezdett és 1993-ban befejezett gyártásáig 17–25 millió eladott darabszámmal minden idők legtöbbet eladott számítógépmo­dellje (W2). Nemcsak új kifejezések jelennek meg, és már nem használatos eszközök, programok elnevezései avulnak el, kopnak ki a szókincsből, hanem a szavakat jelentésváltozás is érintheti: az 1998-ban alapított *Google* cégnévről ma már mindenki elsősorban a keresésre, az általuk működtetett keresőre asszociál (W3). Az új eszközöknek és alkalmazásoknak a szókincsre tett hatása bár különböző jelentőségű, de nem kikerülhető, ha a számítógépes játékok nyelv­változatát vizsgáljuk.

A „számítógépes játék” gyűjtőnév szinte beláthatatlanul sok játéktípust jelöl. Mivel folyamatosan fejlődő területről van szó, nem lehetséges egyetlen átfogó taxonómiát kijelölni, mindig a célnak megfelelően más és más szempontot kell figyelembe venni. „A számítógépes játékok tervezésének állapota gyorsan változik. Ezért várható, hogy az (...) előadott rendszerezés rövid időn belül elavult és nem megfelelő lesz. Az elkövetkező új taxonómiáknak reagálniuk kell majd a következő években megjelenő piaci változásokra is” (Crawford 1984: 39, ford. a szerző). Ezt is figyelembe véve megállapítható, hogy ha nehezen megszámlálható játéktípust tudunk megkülönböztetni változó, nem akadémikusan meghatározható szempontok alapján, akkor a játékok nem lehetnek működőképese­k egy statikus szókinccsel, a típusok változásával a hozzájuk kapcsolódó szó­készlet is variálódik.

A videójátékok nyelve alapvetően az angol. A játékosok ezen a nyelven szocializálódnak a játék közben, viszont a saját anyanyelvi kompetenciájukra támaszkodva tovább is képzik, alakítják a gyakran használt kifejezéseket. Egy korábbi kutatás (Balogh 2013) alapján elmondható, hogy a jelenlegi magyar játéknyelv is dominánsan az angolra (mint a játékok alapnyelvére, a külföldi játékosokkal való kapcsolattartás közvetítőnyelvére) támaszkodik, azonban ez a vizsgálat is dokumentálja, hogy megjelennek jellegzetesen magyar kifejezések is, sőt, a játékosok nagyon kreatívan használják fel, változtatják meg a játékban elérhető nyelvi elemeket a saját kommunikációjuk megkönnyítésére. Ez utóbbi megállapítás a *MOBA*-k (*multiplayer online battle arena* 'sokszereplős online csatater') területén végzett nyelvészeti vizsgálat eredménye (Balogh 2013), de a későbbiekben szükséges lenne több játéktípusból, magyar anyanyelvű beszélőktől származó elemek vizsgálata is, ahogy arra már angol nyelven lehet mintát találni (Herring et al. 2009).

Ezzel szemben a legutóbbi kutatási részeredmények alapján (Balogh 2018a) valószínűsíthető, hogy meghatározó különbség van a nyelv­választás terén a játékon belüli

és a játékról szóló kommunikáció között. Utóbbi esetben a magyar kifejezések vannak túlsúlyban, amellyel mind a cikkek szerzői, mind az olvasók is egyetértettek. A jelenlegi eredmények alapján úgy tűnik, hogy az írott sajtóban a szubkultúrán kívül is megszokott magyar nyelvet preferálják az olvasók. Egy területen alakul speciálisabban a nyelvválasztás: az e-sporttal kapcsolatos cikkekben megjelenik a játékok, játéktípusok befolyása a szövegben használt játéknyelvi elemek nyelvére (lásd Balogh 2018a).

1.3 A kutatás során vizsgált online felületek, játékok

A kutatás során két olyan online felület szókincsét vizsgáltam meg, amelyek célközönsége egyértelműen a játékosok, valamint a tartalmat előállítók is magyar anyanyelvű gamerek: több tematikus Youtube-oldal kommentjeit, valamint egy magyar nyelvű online újságot.

A Youtube elsődleges tartalomgyártó felület, külön aloldala foglalkozik videójáték-közvetítéssel, amelyen a regisztrált felhasználóknak lehetőségük van élőben közvetíteni saját játékmeneteiket és e-sport-mérkőzéseket, versenyeket a szabályok betartásával. Ingyenesen használható, de a játékosok komoly támogatást is kaphatnak a reklámbevételekből. Közösségi jellegét a videó mellett futó chatablak adja, amit a streamer mellett a közvetítés összes nézője lát, és regisztráció után tud is használni. Az élő közvetítés után, ha felkerül a videó a streamer oldalára, akkor lehetőség van visszánézni és a kommentekben további megjegyzéseket tenni, kérdezni.

Az Esport1 nevű tematikus online gamer magazin több e-sport játékkal foglalkozik, hazai és nemzetközi bajnokságok követésével, újdonságok és stratégiák bemutatásával, interjúkkal és véleménycikkekkel. A cikkek alatt nem lehet kommentelni, ehhez a Facebook-oldalra csatornázzák be a megjelenéseket. A cikkírók mind játékosok, a célközönségük is elsősorban a különböző játékok iránt érdeklődő gamerek (Balogh 2018b).

A kutatás során három, különböző játéktípusba tartozó játékot vizsgáltam meg. A *League of Legends* (LoL) egy 2009-ben kiadott *MOBA*, amely 2014 óta elérhető az angol mellett hivatalosan magyar szinkronnal is. A játék során két csapat igyekszik elpusztítani az ellenfél legvédehetőbb épületét, ehhez minden játékos választ egy hőst a játék idejére. A *Counter-Strike* (CS) egy 1999-ben indult *FPS* (*first person shooter* 'belső nézetű lövöldözős játék'), amelynek 2012-ben kiadott *Global Offensive* változata a legfrissebb. A játékban szintén két csapat vesz részt, akik az adott pályára jellemző feladatot hajtanak végre. A játék angol nyelvű, de a megjelenése óta erős hazai játékosközössége van, sajátos szókincssel. Az eddig említett két játék hagyományosan az elektronikus sportok közé tartozik, rendszeresen szerveznek magas összdíjazású versenyeket, bajnokságokat belőlük. A harmadik játék a 2017-ben indult *Fortnite Battle Royal* a nevében viseli a játéktípusát (*battle royal* 'olyan játéktípus, amelyben az utolsóknak talpon maradó nyer'). Egy meccs alatt 100 játékos igyekszik az utolsó túlélő lenni az egyre kisebbedő térképen egyedül, párban vagy hármas csapatokban harcolva. A játék angolul érhető el, nagyon népszerű a különböző *streamerek* ('élő videót, jelen esetben játékot közvetítő személy') között, ami hozzájárult a gyors elterjedéséhez. Mivel a külső befolyásoló tényezők kutatása volt a kitűzött rész célja, így a szókincsvizsgálat kiegészítéseként megvizsgáltam az érintett játékok nyelvi körülményeit befolyásoló tényezőket is.

1.4 Hipotézisek, kutatási célok

A kutatás célja egy adatbázis építésén keresztül annak a vizsgálata, hogy a különböző befolyásoló tényezők hogyan érhetők tetten a magyar játékosok által generált online, írott szövegekben. Ezt több részvizsgálaton keresztül elemeztem, jelen tanulmányban a külső befolyásoló tényezőkre, azon belül is a nyelvválasztás és a játéktípusok kérdé- sére fókuszálva.

A korpuszépítés mellett a kutatás másik fontos, a folytatást is befolyásoló kérdése, hogy a korábbi, önbevallás alapján készített szólista adatai validak-e. Feltételezésem szerint igen, a gyűjtés többsége szerepelni fog a korpuszban.

Hipotézisem szerint a legspecifikusabb szókincsrész, vagyis az egyedi játékok szint- jén mutatkoznak meg a legjobban a vizsgált külső befolyásoló tényezők. Előzetesen azt feltételezem, hogy ha egy adott játékhoz van hivatalos fordítás, az támogatja a magyar nyelvű alakokat, és így nagyobb számban fordulnak elő a korpuszban.

Az előzetes kutatások alapján a nyelvválasztás során a cikkekben a magyar játék- nyelvű szavak túlsúlyát várom. Mivel a Youtube-kommentek közelebb állnak a szóbeli kommunikációhoz, így ezen a területen az angol játéknyelvi szavak dominanciájára számítok, akár önmagukban, akár magyar todalékolással.

2 Módszertan

A vizsgálat a doktori kutatásom részét képezi, amelynek jelen fázisában a játéknyelvi szókincs tanulmányozásához szükséges korpusz építését végzem. A most elemzett adatbázis 18 000 online cikket tartalmaz egy többszerzős, tematikus videójátékos új- ságból (a 2016. július 6. és a 2019. január 29. közötti időszakból), valamint 102 000, a vizsgált cikkekben is érintett játékok alapján szűrte, gamer tematikájú Youtube-videók alá érkezett kommentet (2016–2019). Az adatbázis építésekor a cél az volt, hogy ma- gyar tartalomkészítőktől magyar játékosoknak szóló online szövegekből épüljön fel. A gyűjtést még nem fejeztem be, a korpusz folyamatosan bővül.

Két típusú szűrést végeztem a kapott adatokon. Egyrészt kiemeltem három, mindkét részkorpuszban (cikk és komment) egyaránt szereplő játékot, és megvizsgáltam a sza- vak és a szóelőfordulások számát, mind a teljes, mind a játéknyelvi kifejezések eseté- ben, külön-külön a cikkek és a kommentek esetében. A kutatás során minden releváns szóelőfordulást figyelembe vettem (előfordulás 10-nél több), az egyértelmű elgépelé- seket a legközelebbi alakhoz soroltam. A játéknyelvi kifejezéseknél külön kategorizál- tam a szavak nyelvét angol, magyar, valamint angol és magyar típusba (prototipikus esete az angol szó magyar todalékkal), majd vizsgáltam az elemek számát és előfordu- lási arányát. A játékokkal kapcsolatos szavak listáját korrigáltam, a ragozott vagy elgé- pelt alakokat egységesítettem, és az eredményt újból végigfuttattam a játék korpuszán.

A korpuszon való szűrést egy dinamikusan bővülő szólista alapján is elvégeztem, amely korábbi, önbevalláson alapuló gyűjtések eredményeiből állt össze (778 szó a duplikátumok nélkül). Az első a 2012-ben megjelent *Netszótár* (Veszelszki 2012) című kiadvány, amelynek alapjául egy országos szópályázat szolgált. Veszelszki Ágnes, a kötet szerkesztője arra kérte a pályázókat, hogy küldjenek netes kommunikáció során használt kifejezéseket példamondatokkal és magyarázatokkal, így egyértelműsítve a használat helyét és módját. Ebből kiemeltem a kifejezetten videójátékokkal kapcsolatos szavakat. A projekt folytatásaként 2015 (erről lásd Veszelszki 2015) őszén új gyűjtés indult, ezúttal viszont csak játéknyelvi kifejezéseket vártunk a pályázóktól, ennek az

eredménye is része a vizsgált korpusznak. A harmadik elem az *Egymillióan a magyar e-sportért* egyesület honlapján megtalálható szótár (W4), valamint az annak bővítésére kiírt gyűjtés eredménye. A szólistát mindkét részkorpuszon és minden játékon külön végigfuttattam, hogy ellenőrizni tudjam a szavak érvényességét a játéknyelvi közegben. Az eredmények után a másik, gyakorisági vizsgálatban jól szereplő kifejezésekkel bővítettem a szólistát és újból elvégeztem a számításokat.

3 Eredmények

3.1 Az előzetes gyűjtések anyagának kontrollja

A duplikátumok nélküli 778 elemből 299-re nem szerepelt adat, vagyis a kiindulási anyag 38%-a nem volt megtalálható a mostani korpuszban, így a szóelőfordulásoknak csak 2,8%-a tekinthető előzetesen játéknyelvi elemnek. Azok közül a szavak közül, amelyek nem szerepeltek a mostani korpuszban 9 olyan, amelyik korábban két gyűjtésből érkezett, 2 pedig mindhárom előzetes listában szerepelt.

A két részkorpusz elemzése alapján 134 további játéknyelvi kifejezést vettem fel első körben a szólistára, ennek eredményeként nemcsak a három kiemelt játékban, hanem mindkét részkorpusz játécai között növekedést tapasztaltam.

3.2 A vizsgált játékok statisztikai eredményei

Az online cikkekben a legtöbb önállóan adatolt szó a League of Legends (LoL) alkorpuszában volt (140 563 szó, 2 445 134 előfordulással), ezt követte a Counter Strike (CS) 104 124 szóval, amihez 1 647 857 előfordulás tartozott, végül a Fortnite-hoz kapcsolódóan 43 313 szót adatoltam 364 394 előfordulással (2. táblázat). A játéknyelvi kifejezéseket külön vizsgálva ugyanez a sorrend: a LoL alkorpuszában 696 szó, a CS-vel kapcsolatban 436 szó, a Fortnite témájában pedig 49 kifejezés tartozott ebbe a témakörbe, a teljes korpuszhoz viszonyított előfordulásukat lásd az 2. táblázatban. Az eredmények korrigálása után az önálló játéknyelvi elemek száma átlagosan 71,5%-ra csökkent, viszont az így lefedett előfordulási arány átlagosan megháromszorozódott.

A cikkekben vizsgált angol és magyar gamer kifejezések számát és arányát a 3. táblázat mutatja. Az eredményeket a minimum 100 előfordulást elérő kifejezések elemzése alapján számoltam. Mindhárom esetben a magyar kifejezések voltak a leggyakoribbak, ezután következtek az angol variánsok, majd az angol tőhöz magyar toldalékot kapcsoló szavak.

A Youtube-kommentek korpuszában a LoL 68 875 szava után a Fortnite 44018 eleme következett, és a CS-hez kapcsolódott a legkevesebb kifejezés, 3940. Arányait tekintve nagyon kis részét tették ki a játéknyelvi kifejezések az adatoknak (átlagosan 1,7%, részletesen lásd 4. táblázat). A CS esetében kivételesen a minimum 10 előfordulás esetéig is megvizsgáltam a szavakat, mivel nagyon kevés volt az elemszám. A korrigálás utáni előfordulási arányok ismét pozitív változást mutattak, átlagosan ismét megháromszorozódtak a teljes szóelőforduláshoz viszonyított értékek.

A játéknyelvi elemek nyelvi adataiban a kommentek között az angol és a magyar töredéket tartalmazó szavak továbbra is a legkisebb számban voltak adatolhatóak.

A LoL-ban és a Fortnite-ban az angol kifejezések gyakrabban fordultak elő, mint a magyarok. A CS-ben maradt a cikkekhez hasonló arány a magyar kifejezések többségével (5. táblázat).

Az angol és a magyar nyelven is megjelenő kifejezések vizsgálatához a MOBA játéktípusból választottam példákat (6. táblázat). A ward vagy őrszem egy tárgy, amellyel láthatóvá lehet tenni a pálya egy bizonyos részét, a skin vagy hőskinézet a választott karakter grafikai megjelenését változtatja meg, a lane vagy ösvény pedig a pálya egy kijelölt sávja. Az adatokból látszik, hogy a kommentek között egyértelműen az angol változatok szerepeltek gyakrabban, a cikkekben pedig egy kivétellel szintén az idegen nyelvű kifejezést alkalmazták többször.

	League of Legends	Counter-Strike	Fortnite
Szóelőfordulás	2 445 134	1 647 857	364 394
Önálló szó	140 563	104 124	43 313
Játéknyelvi elem szó/előfordulás (arány a játék teljes szóelőfordulásához viszonyítva)	696 / 207 188 (8,4%)	436 / 141 758 (8,5%)	49 / 12 997 (3,5%)
Játéknyelvi elem: korrigált (arány a játék teljes szóelőfordulásához viszonyítva)	494 / 792 981 (32,4%)	295 / 392 219 (23,8%)	38 / 35 838 (9%)

2. táblázat. Az online cikkek korpuszának statisztikai adatai

	League of Legends	Counter-Strike	Fortnite
Angol (arány az összes játéknyelvi elemhez viszonyítva)	296 (42%)	199 (45%)	21 (42,8%)
Magyar (arány az összes játéknyelvi elemhez viszonyítva)	352 (50%)	211 (48%)	26 (53%)
Angol és magyar (arány az összes játéknyelvi elemhez viszonyítva)	48 (6%)	26 (5%)	2 (4%)

3. táblázat. Az online cikkek korpuszában lévő játéknyelvi szavak nyelvi megoszlása

	League of Legends	Counter-Strike	Fortnite
Szóelőfordulás	497 612	13 404	280 016
Önálló szó	68 875	3940	44 018
Játéknyelvi elem: szó/előfordulás (arány a játék teljes szóelőfordulásához viszonyítva)	75 / 14 822 (2,9%)	29 / 807 (0,6%)	22 / 4320 (1,5%)
Játéknyelvi elem: korrigált (arány a játék teljes szóelőfordulásához viszonyítva)	68 / 64 322 (12%)	22 / 2577 (19%)	19 / 8004 (2,8%)

4. táblázat. A Youtube-kommentek korpuszának statisztikai adatai

	League of Legends	Counter-Strike	Fortnite
Angol (arány az összes játéknyelvi elemhez viszonyítva)	58 (77%)	13 (44%)	12 (54%)
Magyar (arány az összes játéknyelvi elemhez viszonyítva)	16 (21,3%)	15 (51%)	9 (40%)
Angol és magyar (arány az összes játéknyelvi elemhez viszonyítva)	3 (4%)	1 (3%)	1 (4%)

5. táblázat. A Youtube-kommentek korpuszában lévő játéknyelvi szavak nyelvi megoszlása

	A cikkekben		A kommentekben	
	db	%	db	%
<i>ward</i>	524	61%	80	83%
<i>ország</i>	335	39%	16	16%
<i>skin</i>	1458	60%	1159	99%
<i>hőskinézet</i>	954	39%	8	0,006%
<i>lane</i>	678	17%	243	83%
<i>ösvény</i>	3314	83%	48	16%

6. táblázat. Angol és magyar szópárok előfordulásának az adatai

4 Következtetések

4.1 A korábbi szólista érvényessége

A korábbi önbevallás alapján felépített (Balogh 2017, 2018a) szólista 62%-a szerepelt a most vizsgált korpuszban, így elmondható, hogy kiindulási pontként jól használható, valamint a gyűjtési módszer sikeres volt. A további szövegtípusokon való tesztelés során (például a közösségi médiában megjelent posztok) a játéknyelvi elemek jelzésének arányán lenne szükséges javítani. Ennek megoldása lehet, hogy a mostani és a folyamatban lévő többi részkutatás során a gyakorisági vizsgálaton gyakran szereplő szavakkal bővítsem a listát. Nehézséget jelent azonban a szavak eltérő írásváltozatainak gyakorisága, például a *l33t* forma nem szerepelt a korpuszban, de ennek a latin betűs átírása, a *leet* már igen. A több játékon és játéktípuson futtatott szógyűjtemény alátámasztotta azt a feltételezést, hogy vannak kifejezetten egyéni, csak egy játékhoz tartozó kifejezések, de emellett vannak olyan szavak is, amelyeken játéktípusok osztoznak. A továbbiakban ennek a részletesebb vizsgálatát tervezem elvégezni.

4.2 A cikkek és a kommentek korpuszán vizsgált játékok

4.2.1 A játéknyelvi elemek előfordulásának vizsgálata

A cikkek vizsgálatakor feltűnő, hogy a leggyakoribb, 1000 feletti szóelőfordulások között volt a legkevesebb gamerszó, de már az 500-as gyakorisággal vizsgálva is háromszoros növekedést tapasztaltunk, és ez a kilencszeresére nőtt az 500-as és a 100-as gyakoriság között. Ennek az egyik oka a variánsok száma: nagyon sokszor toldalékolt formában kerültek szavak a korpuszba, így megosztva az egy szóra eső gyakoriságot.

Ezért tartom nagyon fontos elemnek a korrigálás beemelését, mivel így a gyakori kifejezések összes alakja megjelenik az előfordulásban, ugyanakkor nem torzítja a jelentős különbséggel rendelkező szóalakok felismerését.

A kommentek vizsgálatakor feltűnő volt a játéknyelvi elemek alacsony száma. 1000-es gyakoriság felett nem is volt található gamerszó, a CS-alkorpuszban 500 felett sem, ismét 500 és 100 között volt jelentősebb a játéknyelvi elemek megjelenése. Ennek oka egyrészt a kommentek beszélgetéshez hasonló kommunikációs helyzete, mivel nagyon sok névmást, utalósót találtam a gyakori szavak között, így valószínű, hogy a közvetítésre reagálva születtek a megjegyzések. Másrészt az elgépelések nagy száma is befolyással lehet a számra, mivel még jobban megszorja a toldalékolás miatt amúgy is kisebb részletekben megjelenő gyakori szavak előfordulási számait is. Ugyanakkor megfigyelhető, hogy a legfiatalabb játéknál, a Fortnite-nál, hiába volt magas a szóelőfordulások száma, a játéknyelvi kifejezések száma alacsony volt. Ennek hátterében az állhat, hogy bár gyakori a közvetítések során, hogy valaki ezzel játszik, de a nézők elsősorban közösségi élményként írnak megjegyzéseket, és a kommentek között folytatott diskurzusok témája nem a játék, hanem például a streamer által felvetett egyéb kérdés megvitatása. Ezek a nézők nagy valószínűséggel értik a játékot, mivel jelen vannak nézőként, ugyanakkor ezt csak ezen a közvetett módon tudom alátámasztani. Ez a tanulság felveti a kommunikációs helyzet befolyásoló szerepét a szubkultúrához kapcsolódó egyedi szókincs használatában is.

4.2.2 A játéknyelvi elemek nyelvi megoszlásának vizsgálata

A nyelvválasztással kapcsolatos eredmények azt mutatják, hogy a cikkekben a magyar kifejezések voltak többségben, a kommentek esetében pedig, egy játékot kivéve, az angol szavak. A CS-ben lévő, bár kis mintán mutató, magyar dominanciát magyarázhatja, hogy a játék már 1999 óta elérhető Magyarországon, és régóta aktív közösség szerveződik körülötte. A két szövegtípus közötti nyelvi különbségre egyrészt választ adnak a korábbi kutatások, miszerint a hagyományos szövegtípusok, például az újságcikkek esetében a játékosok a magyar nyelvet részesítik előnyben, így a cikkírók is ennek alapján fogalmazzák meg a szöveget. A kommentek angol szavainak fölényét magyarázhatja az, hogy a gyorsabb, szóbeli, játékon belüli kommunikációhoz hasonló kommunikációs helyzetben használják a kifejezéseket a játékosok, és ezen a területen még mindig az angol dominanciája a jellemző.

Bár a cikkekben a magyar nyelvű szavak voltak többségben, de nem sokkal maradtak el az angol nyelvű kifejezések sem. Ennek oka az angol kifejezések csoportjaiban kereshető: a szorosan a játékhoz kapcsolódó szavak, például a nevek mindenképpen ezen a nyelven kerülnek be a korpuszba, ezen kívül az e-sporttal kapcsolatos kifejezések közül a játékosok nicknevei (választott felhasználónevei), a csapatok elnevezései és a taktikák is angol nyelven fordulnak elő.

A korrigálás a nyelvválasztás vizsgálatára is befolyással van. Ahogy a mostani eredményekből is látszik, kiemelkednek az egyértelműen angol vagy magyar nyelvű elemek a *hunglish* (magyar és angol elemeket is tartalmazó) szavak rovására. A korrigálás után ez az arány feltételezhetően csak romlik, viszont a leggyakoribb angol szavak szócsaládjainak vizsgálatával pontosabb eredményt lehet elérni a magyar toldalékolás elterjedtségének a vizsgálatában. Ez azért is lenne fontos, mivel az angol kifejezések használatakor két lehetséges alternatívája van a játékosoknak: kapnak vagy kialakítanak egy működőképes magyar fordítást, vagy az angol szót magyarul toldalékolva használják. A mostani eredmény azt mutatja, hogy az előbbi a többiek által követett nyelvhasználati irány, ugyanakkor a korrigálás utáni gyakorisági szám emelkedése az utóbbi stratégiát is támogatja. További kutatásokra van még szükség ennek a kérdésnek az eldöntéséhez,

mivel ez hosszabb távon is fontos következtetésekre mutat rá a nyelvválasztás témakörében.

A 6. táblázatban kiemelt kifejezések tovább árnyalják a nyelvválasztás kérdését. Három olyan kifejezést vizsgáltam, amelyek angol és magyar nyelven is szerepelnek a korpuszban, és a MOBA játéktípusán belül fontosak. A kapott eredmények nincsenek teljesen összhangban a korpuszban kapott számokkal, mivel a háromból két esetben a cikkek részkorpuszában is az angol kifejezés volt a gyakoribb a magyarral szemben. Ez a kettősség rámutat arra, hogy az egymás mellett élő kifejezések között nem döntötte még el egyértelműen a beszélőközösség, hogy melyik alakváltozat kanonizálódik. A játéktípusokon belül több játékot is vizsgáló kutatás során ennek a részletesebb szabályszerűségeire is fény derülhet. Ha az informatikai szakkifejezések honosodását veszem példaként, akkor se egyértelmű, hogy melyik alak szilárdulhat meg, mivel vannak olyan kifejezéseink (például a *számítógép*), amelyek a magyar változatot tartották meg, és vannak az angol szavaknak is meggyökeresedett változataik (*fájl, e-mail*).

Érdemes kitérni az azonos alakú szavakra, amelyeknek a besorolása emiatt sok esetben kérdéses (játéknyelvi vagy sem). A fejlesztők is több esetben élnek ezzel az eszközzel a játékok elnevezésekor: a CS rövidítés a játék neve mellett jelent a MOBA-kban creep statisztikát is, azaz hogy hány semleges egységet ölt a játékos karaktere, a LoL és a WoW (*World of Warcraft*) nevek pedig netes akronimaként is elterjedtek. Ennek vizsgálatára a szavak környezetét is be kell vonni, ami egy új típusú adatfeldolgozási modell kidolgozását igényli.

4.3 A hipotézisekre kapott válaszok

A kutatás jelenlegi fázisára vonatkozóan több részterülettel kapcsolatban voltak előzetes hipotéziseim, ezekkel kapcsolatban az alábbi eredményekre jutottam. Az volt a feltételezésem, hogy a korábbi, önbevallás alapján készített szólista (Balogh 2017, 2018a) adatai többségükben szerepelni fognak a korpuszban. Ezt az eredmények alátámasztották, 62%-ban megjelentek a kifejezések a mostani gyűjtésben is. Ez egyrészt megerősíti, hogy valós adatokat kaptunk az önbevallás alapján készült gyűjtések során is, azonban jelzi, hogy szükség van további bővítésre is a nagyobb lefedettség érdekében.

Hipotézisem szerint a legspecifikusabb szókincsrész, vagyis az egyedi játékok szintjén mutatkoznak meg a legjobban a vizsgált külső befolyásoló tényezők. Ezt a feltételezést az eredmények részben alátámasztották: a nyelvválasztás kérdésében tényleg ezen a szinten lehet releváns válaszokat kapni, ugyanakkor az előzetesen szerkesztett szólista eredményei alapján a különböző játéktípusok és témák befolyása egy szinttel feljebb, a közös típusba tartozó játékok szókincsének együttes elemzésével valósulhat meg effektíven.

Az előzetes kutatások alapján a nyelvválasztás során a cikkek részkorpuszában a magyar játéknyelvi szavak túlsúlyát vártam. Ez a hipotézisem beigazolódott, bár jelentős mértékben szerepelnek mellettük az angol nyelvű kifejezések is.

Mivel a Youtube-kommentek közelebb állnak a szóbeli kommunikációhoz, így ezen a területen hipotetikusan az angol játéknyelvi szavak dominanciájára számítottam, ami két játék esetén teljesült. Ez a két, nyelvválasztással kapcsolatos eredmény rámutat a magyar játékosok tartalomfogyasztási szokásaira és a nyelvi szocializálódására is. A hivatalosabb kommunikációs formák esetén a magyar nyelv használatát preferálják, mivel ezt sajátítják el a korábbi példák (jelen esetben más témájú cikkek) esetén. Érde-

mes a későbbiekben megvizsgálni, hogy a digitális kommunikáció más elemeihez hasonlóan (például a szmájlik) a játéknyelvi elemek is bekerülnek-e a játékosok más típusú írásos kommunikációjába, és ha igen, akkor ott melyik tapasztalat érvényesül: a kommunikációs helyzeté vagy a játéktapasztalaté.

Azt feltételeztem, hogy ha van hivatalos fordítása a játéknak (a korpuszban a LoL), az támogatja a magyar nyelvű játéknyelvi alakokat, és nagyobb számban fordulnak elő a vizsgált szövegekben. A hipotézisemet az eredmények csak részben támasztották alá, mivel ezt a tényezőt felülírja az, hogy milyen típusú szövegben vizsgáljuk a szavak előfordulását: a cikkek részkorpuszán igaznak bizonyult, azonban a kommentek között nem. Ez azt bizonyítja, hogy a különböző külső befolyásoló tényezők között is jelentős hierarchia van.

4.4 Összegzés

A videójátékok nyelvi környezetének sajátosságait figyelembe véve azt a kérdést tettem fel, hogy tetten érhetőek-e szóincset befolyásoló tényezők a játékosok és a tartalomgyártók írott szövegeiben. A kutatás alapján azt a választ kaptam, hogy igen, az angol és a magyar nyelv közötti választás a játéknyelvi szavak között tetten érhető, de a vizsgálat tanulsága, hogy a szabályszerűségeket biztosabb levonásához és megerősítéséhez újabb, ezúttal a játéktípus szintjét célzó kutatásokra van szükség. A magyar játékosok szubkultúráján belüli írásos nyelvhasználatában ezekben a játékokról szóló szövegekben megoszlik aszerint, hogy a tanult minták alapján milyen nyelvet várnak el (cikkeknel a magyart, kommenteknél inkább az angolt). Tehát nem minden esetben egyértelmű, hogy milyen nyelven célszerű megszólítani a célközönséget, mivel történelmi és fejlesztési okokból kifolyólag lehet befolyásoló egy játéktípus vagy egy konkrét játék a nyelv-választás folyamatában. Emellett jelentős befolyásoló tényezőként fel kell venni a kommunikációs helyzetet is a korábbi külső okok közé, mivel a szóhasználatra és a nyelvválasztásra is egyaránt erőteljes kihat.

A hosszú távú céloom egy olyan alapkutatás összeállítása, amely segít bemutatni és előkészíteni a további nyelvi kutatásokhoz ezt a nyelvileg izgalmas és sokszínű szubkultúrát.

Források

A cikkek forrása: <https://esport1.hu/>

A Youtube-kommentek forrásai:

League of Legends Hungary: <https://www.youtube.com/user/LeagueofLegendsHU>

Magyar Fortnite Tippek: <https://www.youtube.com/channel/UCHbWYRTA4XrR-vO9XIaJqoA>

AlexGaming: <https://www.youtube.com/channel/UChLf4L6Gi0MX-me7NAk5PGg>

Irodalom

Balogh A. 2013. *A számítógépes játékok és az oktatás*. Szakdolgozat, kézirat. Budapest: ELTE BTK.

Balogh A. 2014. Geeknyelvtan? A számítógépes játékok nyelvéről. In: Veszelszki Á. – Lengyel K. (szerk.) 2014. *Tudomány, technolektus, terminológia. A tudományok, szakmák nyelve*. Budapest: Éghajlat Kiadó. 345–352.

- Balogh A. 2017. A számítógépes játékok nyelvi környezetének vizsgálata. In: Gyói R. – Kocsis Zs. – Krizsai F. (szerk.) 2017. *Kérdések a Félúton műhelyéből, utak a megoldásokhoz. Az Eötvös Loránd Tudományegyetem Nyelvtudományi Doktori Iskolájának 11. Félúton konferenciáján elhangzott előadások tanulmánykötete*. Budapest: Kalota Művészeti Alapítvány. 161–177.
- Balogh A. 2018a. Videójátékokról szóló online diskurzusok vizsgálata. In: Fazekas B. – Kaposi D. – P. Kocsis R. (szerk.) 2018. *Csomópontok – Újabb kérdések a Félúton műhelyéből. Az Eötvös Loránd Tudományegyetem Nyelvtudományi Doktori Iskolájának 12. Félúton konferenciáján elhangzott előadások tanulmánykötete*. Budapest: Kalota Művészeti Alapítvány. 203–214.
- Balogh A. 2018b. *Az e-sport-kommunikáció médiafelületei*. Kézirat.
- Carrillo Masso, I. 2009. Developing a Methodology for Corpus-based Computer Game Studies. *Journal of Gaming and Virtual Worlds* 1(2): 145–172.
- Crawford, C. 1984. *The Art of Computer Game Design*. Berkeley: McGraw-Hill/Osborne Media.
- eNET 2019. *Egyre többet költünk videójátéokra*. Elérhető: <https://enet.hu/hirek/egyre-tobbet-koltunk-videojatekra/> (letöltés ideje: 2019. 04. 07.)
- Ensslin, A. 2012. *The Language of Gaming*. Basingstoke: Palgrave Macmillan.
- Herring, S. C. – Kutz, D. O. – Paolillo, J. C. – Zelenkauskaitė, A. 2009. Fast Talking, Fast Shooting: Text Chat in an Online First-Person Game. *Proceedings of the Forty-Second Hawai'i International Conference on System Sciences (HICSS-42)*. Los Alamitos, CA: IEEE Press.
- Veszelszki Á. 2015. *Gamernyelv*. Elérhető: <https://www.veszelszki.hu/gamernyelv.html> (letöltés ideje: 2019. 07.07.)
- Veszelszki Á. 2017. *Netnyelvészet. Bevezetés az internet nyelvhasználatába*. Budapest: L'Harmattan Kiadó.
- Veszelszki Á. (szerk) 2012. *Netszótár. @-tól a Zukbergnetig*. Budapest: ELTE Eötvös Kiadó W1: Wikipédia szócikk: egér. Elérhető: [https://hu.wikipedia.org/wiki/Eg%C3%A9r_\(sz%C3%A1l%C3%ADt%C3%A1stechnika\)](https://hu.wikipedia.org/wiki/Eg%C3%A9r_(sz%C3%A1l%C3%ADt%C3%A1stechnika)) (letöltés ideje: 2019. 04. 14.)
- W2: Wikipédia szócikk: C64. Elérhető: http://hu.wikipedia.org/wiki/commodore_64 (letöltés ideje: 2019. 04. 14.)
- W3: Wikipédia szócikk: Google. Elérhető: https://hu.wikipedia.org/wiki/Google_keres%C5%91 (letöltés ideje: 2019. 04. 14.)
- W4: Egymillióan a magyar esportért: Esport szótár. Elérhető: <https://esportmilla.hu/2015/02/esport-szotar/> (letöltés ideje: 2019. 04. 14.)

Érzelmi tartalmak észlelése a magyar beszédben

Bartók Márton

ELTE BTK Nyelvtudományi Doktori Iskola,
MTA-ELTE Lendület Lingvális Artikuláció Kutatócsoport
bartokmarton@gmail.com

Kivonat: Érzelmi állapotaink kifejezésében nagy szerepe van a szupraszegmentumoknak. Az itt bemutatott kísérlet az érzelmi tartalmak ilyen módon történő, csak a beszéd nemnyelvi elemeit hasznosító közvetítésének észlelését vizsgálja. A kísérlet során eljátszott érzelmi állapotokat kifejező beszédhangokat játszottunk le 57 adatközlőnek, akik azokat az érzelmeket jellemző három, sematikus illusztrációkkal megjelenített, az érzelmi állapotok jellemzésében gyakran használt skálán (aktiváció, valencia, dominancia) értékelték. Az eredmények azt mutatják, hogy az adatközlők képesek a különböző aktivációjú, valenciájú és dominanciájú érzelmi tartalmak elkülönítésére kizárólag az adatközlő beszéde alapján az alkalmazott vizuális skála segítségével, azonban ennek mértékét több tényező is befolyásolja: mind a vizsgált érzelmi dimenzió, mind a stimulust produkáló beszélő személye befolyásolta az érzelmi állapotok elkülönítésének mértékét.

1 Bevezetés

Az emberi beszéd a nyelvi információn kívül számos nyelven kívüli, azaz paranyelvi tartalmat is hordoz. Ilyen nemnyelvi információ többek között a beszélő által kifejezett érzelem, melyet jellemzően a beszéd időzítési mintázatai, az intonáció és hangszín, azon belül is a hangszalagrezgés egyéb jellegzetességeit magába tömörítő fonációtípus milyensége segítségével vagyunk képesek kifejezni (Murray–Arnott 1993). Azzal, hogy ezek az érzelmi tartalmak milyen artikulációs-akusztikai változásokkal jelenítődnek meg, számos tanulmány foglalkozik, az itt bemutatott kutatás célja azonban a beszédben megjelenő érzelmi tartalmak észlelésének vizsgálata – az itt bemutatott kísérlet azt a kérdést járja körbe, hogy a beszédben kifejeződő érzelmi tartalmak különböző komponenseit hogyan dolgozzuk fel, és hogyan befolyásolják ezt a feldolgozási folyamatot különböző tényezők.

Az emberi érzelmek kategorizációját leíró tudományos modelleket kétfelé oszthatjuk: az alapérzelem-modellek (pl. Ekman 1992) kötött számú diszkrét, univerzális alapérzelmet feltételeznek (pl. düh, félelem, szomorúság, boldogság, undor és félelem), melyek egymással kombinálódva alkothatnak más érzelmeket. A dimenzionális érzelemelméleti modellek ezzel szemben az érzelmek kategorizációját és elrendeződését bizonyos dimenziók mentén gondolják el: azt feltételezik, hogy az emberi érzelmi

Bartók Márton: Érzelmi tartalmak észlelése a magyar beszédben. In Várad Tamás (szerkesztő), Ludányi Zsófia, Grácsi Tekla Etelka (szerkesztő): *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2019. XIII. Alkalmazott Nyelvészeti Doktoranduszkonferencia*. Budapest: MTA Nyelvtudományi Intézet. 2019. 34–44. DOI: 10.18135/Alknyelvdok.2019.13.3

állapotok sokféleségét két, három, vagy esetleg több kontinuum segítségével elégségesen jellemezhetjük. Russell (1980) szerint az érzelmek az érzelmi állapotban jellemző aktiváció és a valencia, tehát az érzelem pozitív-negatív jellege alapján jellemezhetők. Magas aktivációval jellemezhető például a félelem vagy a meglepődés, alacsony aktivációval az unalom és a nyugalom, ezeket az érzelempárokat pedig megkülönbözteti egymástól a valencia dimenziójában elfoglalt helyük (ti. az unalom negatív, a nyugalom pozitív valenciájú). Emellett több modell is tartalmazza az egyén adott állapotban érzett dominanciájának mértékét (pl. Mehrabian–Russell 1974), tehát lényegében azt az érzetet, hogy az egyén mennyire érzi képesnek magát környezete kontrollálására. Az itt bemutatott kutatásban azt vizsgálom, hogy az érzelmi állapotok beszédalapú kategorizációja hogyan működik ebben a három dimenzióban (aktiváció, valencia, dominancia). Azt vizsgálom tehát, hogy az érzelmi tartalmak észlelése során, amennyiben csak a beszédjel áll rendelkezésünkre, hogyan vagyunk képesek megkülönböztetni egymástól a különböző aktivációval, valenciával vagy éppen dominanciával jellemezhető érzelmi állapotokat. Az egyes érzelmeknek az ebben a háromdimenziós térben elfoglalt pozícióját Fontaine és mtsai. (2007) eredményei alapján határozom meg: ez a kutatás az érzelmi állapotok hat komponensét (az érzelmet kiváltó események értékelése, pszichofiziológiai változások, motoros kifejezések, cselekvési tendenciák, szubjektív érzések, érzelemszabályozás) megragadó 144 Likert-skálás kérdés varianciaanalízise alapján, közel 500 adatközlő részvételével helyezte el a különböző érzelmi állapotokat ebben a térben.

Az érzelmek észlelését vizsgáló kutatások általában eljátszott érzelmek észlelését vizsgálják, és arra fókuszálnak, hogy milyen arányban azonosítják helyesen az adatközlők a felvételen hallott érzelmi állapotokat – tehát kötött számú, a stimulusok között valóban megjelenő érzelmi állapot közül milyen arányban választják ki az adott stimuluson valóban kifejezni szándékozott érzelmi állapotot. Ezek alapján a kutatások alapján a megjelenített érzelmi tartalmat egészséges beszélők esetén a véletlen szintje fölött képesek vagyunk beazonosítani. Scherer (1989) harminc, jellemzően alapérzelmek észlelését vizsgáló tanulmány eredményei alapján átlagosan 60 százalékosra teszi az érzelmi állapotok helyes azonosításának arányát, ami valamivel kisebb pontosságra utal, mint amilyen arányban az érzelmeket közvetítő másik fontos tényező, az arckifejezések alapján vagyunk képesek azok azonosítására. Bár a legtöbb kutatásban profi színészek bemondásai alapján vizsgálják az érzelmi tartalmak észlelését, Jürgens és mtsai (2015) eredményei alapján az azonosítás pontosságára nem volt hatása annak, hogy színészek vagy laikus beszélők hangját használták stimulusként.

Bár a jelen kutatás során is eljátszott érzelmek beszédalapú észlelését vizsgáltam, a fentiekkel ellentétben azt kíséreltem meg leírni, hogy az érzelmi állapotokat jellemző dimenziók mentén hogyan jellemzik az adatközlők az észlelt érzelmi állapotot. Vizsgáltam egyrészt, hogy a hallott stimulusokat valóban úgy jellemzik-e az adatközlők a három vizsgált érzelmi dimenzióban, mint ahogy az adott stimulus produkciójakor eljátszott érzelem Fontaine és mtsai. (2007) eredményei alapján jellemezhető. A stimulusként használt beszédanyagon végzett akusztikai mérések eredményeit figyelembe véve azt is vizsgáltam továbbá, hogy mutat-e eltérést az érzelmi állapotok észlelése annak függvényében, hogy a beszélő mekkora változatossággal valósítja meg az érzelmkifejezést. Azt feltételeztem, hogy amennyiben a beszélő hangmintáin mért, a szakirodalom alapján az érzelmek kifejezésében relevánsnak tekinthető akusztikai paraméterek nagy változatosságot mutatnak a különböző érzelmek között, abban az esetben az egyén érzelmi állapotait nagyobb mértékben képesek vagyunk elkülöníteni, mint kisebb akusztikai változatossággal jellemezhető beszélő hangmintái esetén.

2 Módszertan

A kutatás tehát egy olyan percepciókísérletre épült, mely során az adatközlők különböző érzelmi töltetű beszédanyagokat minősítettek a kutatás során alkalmazott érzelmelméleti modell dimenziói (aktiváció, valencia, dominancia) szerint.

Mivel a kísérlet célja az érzelmi tartalmak észlelésének vizsgálata volt, így stimulusként olyan hanganyagokat használtam, melyek kizárólag a kifejezett érzellem tekintetében különböznek, a nyelvi forma minden esetben állandó volt. Ehhez 17 magyar anyanyelvű, 18–34 év közötti (Geifman et al. 2013 orvostudományilag elhatárolt korcsoportjai alapján meghatározott intervallum) nőt kértem arra, hogy produkáljanak kitarított [a:] beszédhangokat, különböző érzelmeket kifejezve. Az eljátszott érzelmeket a GRID projekt (Fontaine et al. 2013) 34 nyelven elérhető, visszafordítással validált érzelmlistájából, a Fontaine és mtsai. (2007) által feltárt dimenziótér minél teljesebb lefedésével állítottam össze. Az érzelmelek a következők voltak: lenézés, undor, düh, ingerültség, stressz, meglepődés, érdeklődés, boldogság, szeretet, csalódás, együttérzés, büntudat, szégyen, szomorúság, illetve egy érzelmileg semleges állapot (1. táblázat). A beszélők öt ismétlést produkáltak minden érzelemből, olyan módon, hogy az öt ismétlés egymást követte, azonban az érzelmelek sorrendje randomizálva volt. A vizsgált érzelmelek megoszlását a Fontaine és mtsai. (2007) által kialakított dimenziótérben az alábbi táblázat szemlélteti.

érzelem	aktiváció	valencia	dominancia
boldogság	magas	pozitív	domináns
csalódás	alacsony	negatív	nem domináns
düh	magas	negatív	domináns
együttérzés	alacsony	pozitív	domináns
érdeklődés	alacsony	pozitív	domináns
ingerültség	magas	negatív	domináns
lenézés	alacsony	negatív	domináns
meglepődés	magas	pozitív	nem domináns
stressz	magas	negatív	nem domináns
szégyen	magas	negatív	nem domináns
szeretet	magas	pozitív	nem domináns
szomorúság	alacsony	negatív	nem domináns
undor	alacsony	negatív	domináns

1. táblázat. Az aktiváció, valencia és dominancia értékei a vizsgált érzelmekekre Fontaine és mtsai. (2007) eredményei alapján

Ebből a 17 beszélőtől származó anyagból választottam ki a kísérletben használt stimulusnak három beszélőt arra törekedve, hogy ez a három beszélő eltérő akusztikai változatossággal valósítsa meg a különböző érzelmi tartalmakat: az volt tehát a cél, hogy az egyik beszélő csak kis mértékben, a második közepesen, a harmadik pedig nagyon változatosan fejezze ki a különböző érzelmeket. A megfelelő beszélők kiválasztása érdekében az egyénre jellemző szisztematikus, érzelmemfüggő változatosságot az alábbi módon jellemeztem. Minden beszédhangon megmértem 12 akusztikai paramétert, melyek a hangszalagok működésének akusztikai következményét, a zöngét jellemzik, és amelyek nagymértékben különböznek egyes érzelmelek kifejezésekor (Patel et al. 2011). Ezek a paraméterek a következők voltak: jitter, shimmer, HNR, f_0 , alpha

ratio, L_{eq} , az automatikus inverz szűréssel kinyert forrásjel amplitúdója, MFDR, NAQ, CQ, valamint H1-H2 az akusztikai és a forrásjel alapján. Az akusztikus mérések a Praat szoftver (Boersma–Weenink 2018), illetve a COVAREP szkriptcsomag (Degottex et al. 2014) használatával történtek. A beszédhangok teljes idején mért átlagértékekből az érzelmenként öt ismétléshez tartozó öt adatpontot beszélőnként átlagoltam, és az adott beszélő semleges állapotban produkált ejtésein mért értékek átlagához képesti százalékos eltérés formájában normalizáltam. Végül ennek a beszélőnként 168 értéknek (12 paraméter \times 14 érzelem) vettem az átlagát, ami tehát egy, a beszélőre jellemző, a semlegestől való átlagos relatív eltérésként értelmezhető mérőszámot eredményez – ez a szám tehát az adott beszélőre jellemző általános érzelemfüggő akusztikai változatosságot számszerűsíti. A beszélőket sorba rendeztem e szerint a mérőszám szerint, és az akusztikai változatosság alsó kvartilisénél, mediánjánál és felső kvartilisénél található beszélőktől használtam beszédmintákat a percepciók kísérletben. Beszélőnként egy érzelemhez egyetlen megvalósulást, az ismétlések sorrendjében a középsőt választottam stimulusnak. Így végeredményben a semleges beszédmintákat is felhasználva 45 hangmintát használtam ((14 + 1 érzelem) \times 3 beszélő). Az adatközlők minden stimulust csupán egy alkalommal hallottak, mivel az érzelmi tartalmak észlelése általában nagy igénybevételt jelent az adatközlők számára. Az ismétlések számának növelésével a jövőben fokozható lenne az eredmények megbízhatósága, azonban ezt feltehetőleg a felismerendő érzelmek számának csökkentésével kellene ellensúlyozni.

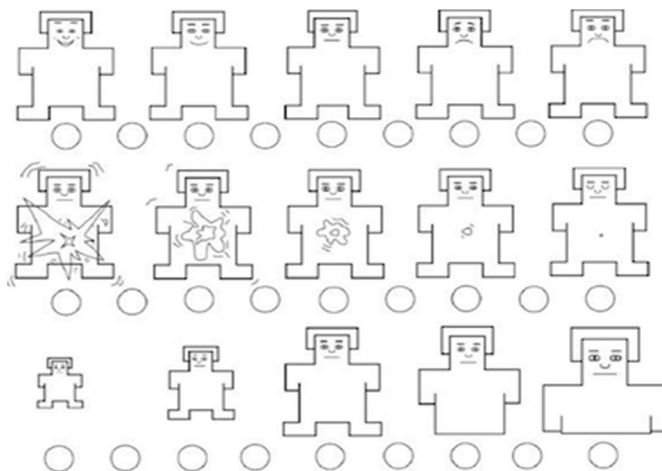
Mivel a kutatás során az érzelmi állapotok különböző dimenziók mentén történő kategorizációját vizsgáltam, olyan percepciók teszt használata volt szükséges, melyben az adatközlők valamilyen módon a vizsgált dimenziók mentén minősítik a hallott stimulus érzelmi tartalmát. Jelen kutatásban az úgynevezett Self-Assessment Manikin (Bradley–Lang 1994, a továbbiakban SAM) módszert alkalmaztam. A SAM egy három skálából felépülő vizuális értékelési rendszer (vö. 1. ábra), mely az általam vizsgált három dimenzió pontjait sematikus emberfigurákkal jeleníti meg. A valencia értékelése egy olyan skálán történik, melynek egyik végpontján a sematikus figura arckifejezése negatív érzelmeket tükröz (lefelé görbülő száj), míg a másik végponton pozitív érzelem jelenik meg (mosoly), a skála közbeeső pontjaihoz pedig a két arckifejezés közötti átmenetet megjelenítő illusztrációk tartoznak. Az aktivációhoz tartozó skálát egy relaxált (inaktív) és egy dinamikus, aktív figura közötti átmenet alkotja, míg a dominanciát mérő skálán a figura mérete változik: a nagyobb méret az egyén dominanciáját jeleníti meg a kisebb figurákkal szemben. A SAM előnye, hogy nem kötődik kultúraspecifikus verbális kategóriákhoz, ezért széles körben használható az érzelmi állapotok egységes jellemzésére. Adatközlőimet arra kértem, hogy a hallott ingereket 9 fokú SAM-skálán jellemezzék.

A percepciók tesztet online bonyolítottam le, a jsPsych könyvtár (Leeuw 2015) testreszabásával. Bár ezzel a módszertani megoldással kapcsolatban gyakran felmerül, hogy a hallgatás körülményei nem kontrollálhatók, Horton és mtsai. (2011), valamint Germine és mtsai. (2012) eredményei alapján az, hogy egy auditív percepciók kísérlet laboratóriumi vagy online lebonyolítással történik, nincs szignifikáns hatással a teszt eredményeire. A kísérlet elején egy szélsőségesen halk és egy szélsőségesen hangos megnyilatkozás alapján az adatközlőket arra kértem, hogy állítsák be készülékük hangerejét olyan módon, hogy kényelmesen hallják mindkét hangot, és ezen a beállításon a kísérlet során ne módosítsanak. Ezután néhány személyes adatra kérdeztem rá.

Ezt követően a SAM alsóskálájának bemutatása következett, a képi skálák mellett minimális verbális leírással magyarázva azok használatát (Bradley–Lang 1994 alapján). Az adatközlők a következő leírást kapták: „A következőkben arra foglak kérni, hogy

a hangminták alapján három képi skála segítségével jellemezd a felvételen hallható ember lelkiállapotát. Az első skála arra kérdez rá, mennyire pozitív (balra) vagy negatív (jobbra) a kifejezett érzelm. A második skálán azt jelöld, hogy mennyire érzed aktívnak (balra) vagy passzívnak (jobbra) a személyt a hangja alapján. Végül a harmadik skálán azt kell megjelölni, mennyire tűnik dominánsnak, irányítósnak (jobbra) vagy éppen irányítottknak (balra) az épp hallott ember”. Ezután következett a 45 hangminta értékelése, randomizált sorrendben. Az adatközlők egy hangmintát többször is meghallgathattak, mielőtt jellemezték volna azokat. A kísérletben összesen 57 adatközlő vett részt, 49 nő és 8 férfi. Átlagéletkoruk 29 év volt.

A SAM alskáláin kapott értékeket adatközlőnként z-transzformáltam, hogy az egyéni válaszadási tendenciák hatását minimalizáljam. A Fontaine és mtsai. (2007) által meghatározott aktiváció, valencia és dominancia nominálissá alakított változók és a beszélőre jellemző akusztikai variabilitás szintje (szintén nominális, háromszintű változó) közötti, az adott dimenzióhoz tartozó SAM-alskála z-transzformált értékeire kifejtett interakciós hatások tesztelésére lineáris kevert modelleket alkalmaztam (Kuznetsova et al. 2015) az R szoftver segítségével (R 2018). A modellek az interakciós hatások mellett adatközlőnként random interceptet tartalmaztak.



1. ábra. 9 fokú SAM-skála. Az alskálák sorrendje felülről lefelé: valencia, aktiváció, dominancia

3 Eredmények

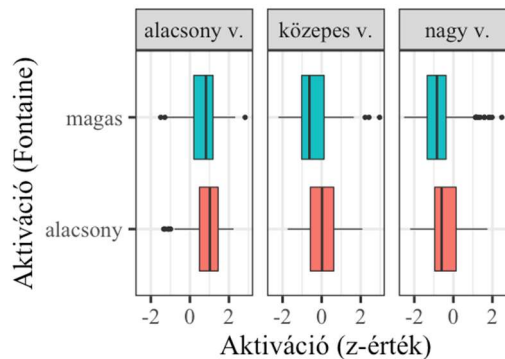
Az alábbiakban az egyes alskálákra vonatkozó eredményeket egyesével mutatom be. A SAM-skálák (z-transzformált) értékeinek értelmezésekor fontos figyelembe venni, hogy azok a vízszintesen tájolt vizuális skálák elrendezéséből kiindulva értelmezendők, melyeken az értékek mindig balról jobbra növekednek (1-től 9-ig), azonban az adott skálán megjelenített dimenzió tartalmilag ennek ellentmondhat – így például a magas aktivációt reprezentáló figura éppen a skála bal oldalán helyezkedik el. A könnyebb értelmezhetőség érdekében az eredményeket a vizuális skálákkal megegyező tájolású

dobozdiagrammokon ábrázoltam, minden esetben részletezve, hogy a skála mely végpontja milyen sematikus emberfigurával volt jelölve.

3.1 Aktiváció

Szignifikáns interakciót találtam a beszélőre jellemző akusztikai variabilitás és a fontaine-i aktiváció között ($F[2, 2217] = 8,08, p < 0,001$) az aktiváció alskála z-transzformált értékeire (2. ábra). Az aktiváció vizuális SAM-alskáláján a magas aktivációs szintű, dinamikusabbnak ábrázolt figura a bal oldalon helyezkedik el, ennek megfelelően azt vártam, hogy ezekre az érzelmekre az adatközlők alacsonyabb értékeket adnak a Likert-skálán – ez így is történt. Ez a különbség továbbá várakozásaimnak megfelelően változik az adatközlőkre jellemző akusztikai variabilitás függvényében – a legkisebb eltérést a kis variabilitású beszélő esetén találtam, ugyanakkor a közepes és a nagy variabilitású beszélő stimulusain közel hasonló mértékben különböztették meg az alacsony és magas aktivációjú érzelmeket az adatközlők.

Ezen túl az is látszik, hogy nem csak a különböző aktivációjú érzelmek elkülöníthetősége változik beszélőnként, hanem az is, hogy általánosságban milyenek minősítik az adott beszélőt ebben a dimenzióban. Az alacsony variabilitású beszélő esetén az értékek az érzelmek alacsony vagy magas aktivációjától függetlenül általánosan magasabbak, tehát alacsonyabb aktivációjúnak (átlagosan 0,79, $sd = 0,74$) vannak minősítve, mint a közepes (átlagosan $-0,2$, $sd = 0,86$), valamint a kifejezett érzelmi tartalomtól függetlenül általánosan is nagyobb aktivációval (kisebb értékekkel) jellemzett nagy variabilitású beszélő (átlagosan $-0,58$, $sd = 0,81$).

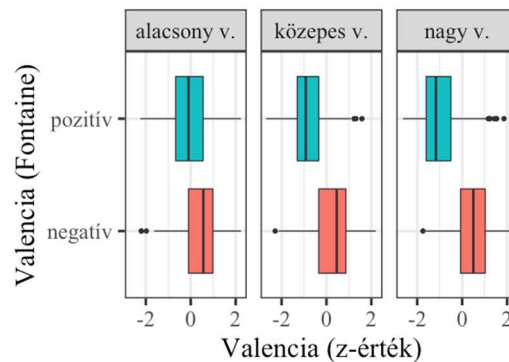


2. ábra. Az aktiváció alskála értékei az érzelmek aktivációjának Fontaine és mtsai. (2007) szerinti alacsony/magas kategorizációja és a beszélőre jellemző akusztikai variabilitás mértéke szerint. Az alacsonyabb értékek a vizuális skálán bal oldalon reprezentált magas aktivációjú állapotot jelentik.

3.2 Valencia

Szintén szignifikáns interakciós hatást találtam a beszélőre jellemző variabilitás és a fontaine-i valencia között ($F[2, 2217] = 57,17, p < 0,001$) a valencia SAM-alskálájának z-transzformált értékeire (3. ábra). A vizuális skálán itt balra található a

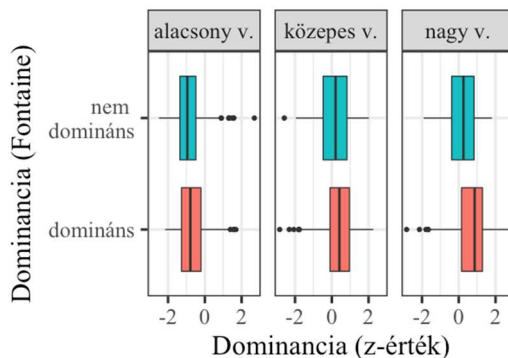
pozitív valenciát reprezentáló, mosolygós figura, tehát a pozitív érzelmek esetén alacsonyabb értékeket vártunk – és az adatok valóban ezt a tendenciát mutatják. Az eltérés a pozitív és a negatív érzelmek között fokozatosan nő a beszélőkre jellemző akusztikai variabilitás növekedésével. Az adatközlőre általánosan jellemző valencia nem mutat olyan fokozatos változást, mint amit az aktiváció esetében tapasztalhattunk, inkább az alacsony variabilitású beszélőre jellemző, inkább negatív valencia (nagyobb értékek, átlagosan 0,22, $sd = 0,86$) áll szemben a közepes (átlagosan $-0,12$, $sd = 1$) és a nagy variabilitású beszélőre jellemző pozitívabb valenciájú minősítéseivel (átlagosan $-0,1$, $sd = 1,05$), azonban ezek a különbségek közel sem akkorák, mint az aktiváció esetén.



3. ábra. A valencia alskála értékei az érzelmek valenciájának Fontaine és mtsai. (2007) szerinti pozitív/negatív kategorizációja és a beszélőre jellemző akusztikai variabilitás mértéke szerint. Az alacsonyabb értékek a vizuális skálán bal oldalon reprezentált pozitív valenciájú állapotot jelentik.

3.3 Dominancia

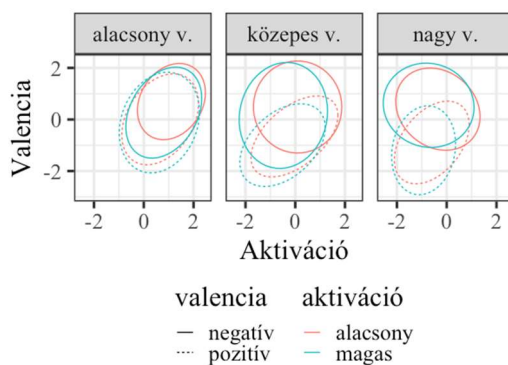
Szignifikáns interakciós hatást találtam a beszélőre jellemző akusztikai variabilitás és a dominancia mértéke között a dominancia alskálájának z -transzformált értékeire ($F[2, 2217] = 9,48, p < 0,001$) (4. ábra). Ezen az alskálán a kisméretű figurával reprezentált nem domináns végpont balra helyezkedett el, így az ebbe a kategóriába tartozó érzelmek esetén várunk kisebb értékeket – és valóban, a Fontaine és mtsai. (2007) által nem dominánsnak kategorizált érzelmeket jellemzően kevésbé dominánsnak értékelték az adatközlők, mint a domináns kategóriába tartozókat. Ez az eltérés a kis akusztikai variabilitással jellemzett beszélő esetén még minimális, és fokozatosan nő a nagyobb variabilitású beszélő felé. Az aktiváció alskálájához hasonlóan az interakciós hatás nem csak annak eredménye, hogy az érzelmek két csoportja (domináns/nem domináns) közötti különbség növekszik a beszélőre jellemző variabilitás hatására, hanem a kifejezett érzelmetől függetlenül az adatközlők szerint a három beszélőre általánosan jellemző általános dominancia mértékében is találunk különbséget, mégpedig ismét a két kategória közötti eltérés nagyságával párhuzamosan: az alacsony variabilitású beszélő esetén a dominancia alacsony, tehát kevésbé dominánsnak ítélik a beszélőt (átlagosan $-0,77$, $sd = 0,76$), a közepes beszélőt ennél magasabb dominanciájúnak találják (átlagosan $0,29$, $sd = 0,85$), a legdominánsabbnak pedig a kifejezett érzelmetől függetlenül a legnagyobb akusztikai variabilitású beszélőt jellemezték (átlagosan $0,48$, $sd = 0,85$).



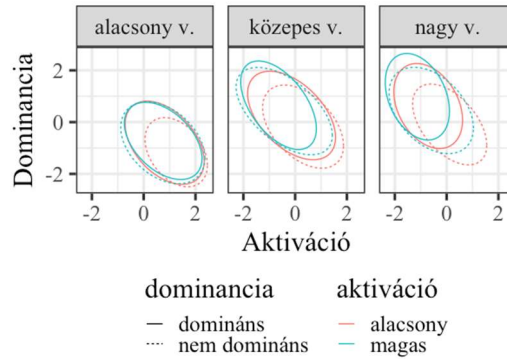
4. ábra. A dominancia alskála értékei az érzelmek dominanciájának Fontaine és mtsai. (2007) szerinti domináns/nem domináns kategorizációja és a beszélőre jellemző akusztikai variabilitás mértéke szerint. Az alacsonyabb értékek a vizuális skálán bal oldalon reprezentált nem domináns állapotot jelentik.

3.4 Összefoglalás

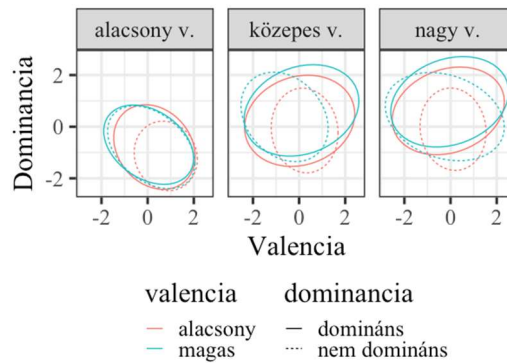
Eredményeink összességében azt mutatják, hogy adatközlőink a SAM mindhárom skáláján képesek voltak elkülöníteni a Fontaine és mtsai. (2007) alapján különböző aktivációjú, valenciájú és dominanciájú érzelmeket. Az alskálák együttes ábrázolásával jól látszik, hogy a legjobban az adatközlők a valencia alskálája mentén képesek differenciálni a különböző kategóriákba tartozó érzelmi állapotokat (5. és 7. ábra). Ezzel szemben a dominancia és az aktiváció alskáláin jóval kisebb az értékek eltérése a fontaine-i kategorizáció alapján elkülönített érzelmek csoportjai között, különösen a kis akusztikai variabilitással jellemzett beszélő esetén (6. ábra).



5. ábra. A valencia és aktiváció alskálák értékei 95%-os konfidenciaellipszisekkel ábrázolva, Fontaine és mtsai. (2007) alapján csoportosítva az érzelmeket



6. ábra. Az aktiváció és dominancia alszkálák értékei 95%-os konfidenciaellipszisekkel ábrázolva, Fontaine és mtsai. (2007) alapján csoportosítva az érzelmeket



7. ábra. A valencia és dominancia alszkálák értékei 95%-os konfidenciaellipszisekkel ábrázolva, Fontaine és mtsai. (2007) alapján csoportosítva az érzelmeket

4 Következtetések

Kutatásom célja az érzelmi tartalmak beszédalapú észlelésének vizsgálata volt. Kutatásomban arra fókuszáltam, hogy három olyan dimenzió mentén, mely segítségével az érzelmeket gyakran jellemzi a szakirodalom (aktiváció, valencia, dominancia), mennyire vagyunk képesek elkülöníteni egymástól az e tekintetben különböző érzelmi állapotokat akkor, ha kizárólag a beszélő hangjával találkozunk, míg a nyelvi forma állandó. Emellett a kérdéssel együtt arra is kerestem a választ, hogy mennyiben befolyásolja az érzelmi tartalmak észlelését azoknak az akusztikai paramétereknek az érzelemfüggő változatossága, melyek a szakirodalom alapján akusztikai szempontból megbízhatóan különítik el egymástól a különböző érzelmi állapotokat – jelenti ez vajon azt is, hogy az észlelésben is nagyobb különbséget teszünk ezen hangminták között? 57 adatközlő vett részt abban a kísérletben, mely során 14 különböző érzelmi állapot eljátszását rögzítő beszédhangokat játszottam le, hogy azt a vizsgált három dimenzió-

ban jellemezzék. Minden érzelmet három beszélő megvalósításában hallottak az adatközlők. Ezt a három beszélőt egy nagyobb, összesen tizenhét beszélőt tartalmazó anyagból választottam ki úgy, hogy az akusztikai variabilitás eltérő szintjeit mutassák (Patel et al. 2011) alapján. Az adatközlők a stimulusokat a Self-Assessment Manikin (SAM, Bradley–Lang 1994) vizuális skáláin értékelték a három fent említett érzelmi dimenzió szempontjából.

Az eredmények azt mutatják, hogy az adatközlők képesek a különböző aktivációjú, valenciájú és dominanciájú érzelmi tartalmak elkülönítésére kizárólag az adatközlő beszéde alapján a SAM vizuális skáláinak segítségével, azonban ennek mértékét több tényező is befolyásolja. Egyrészt nagy különbség van az egyes alskálák között: leginkább az érzelmek valenciáját, tehát pozitív/negatív jellegét voltak képesek megkülönböztetni az adatközlők, míg az érzelem aktivációja és dominanciája jóval kevésbé különült el a válaszok alapján.

Nagyban befolyásolta az eredményeket az is, hogy milyen mértékű akusztikai variabilitást mutat az adatközlő – várakozásaimat alátámasztva a nagyobb akusztikai variabilitás esetén mindhárom alskálán jobban elkülönülnek egymástól a különböző kategóriákba sorolható érzelmek, azonban az akusztikai változatosság csökkenésével ezek a szembenállások egyre redukálódtak. A kismértékű akusztikai változatosságot mutató beszélő esetén már főként csak az érzelmi állapotok valenciája mentén voltak képesek azokat szétválasztani az adatközlők.

A beszélőre jellemző akusztikai variabilitás mértékével nem csak az változik, hogy milyen mértékben vagyunk képesek elkülöníteni egymástól a különböző érzelmi tartalmakat, hanem az is, hogy általánosságban (a kifejezett érzelemtől függetlenül) hogyan minősítjük az adott beszélő érzelmi állapotát a vizsgált dimenziókban. Az eredményekből az a tendencia rajzolódik ki, hogy a magasabb akusztikai variabilitást mutató adatközlőt általánosan aktívabbnak és nagyobb dominanciájúként jellemezték, mint a stimulusokat kisebb akusztikai variabilitással produkáló beszélőket. Ez az összefüggés arra utalhat, hogy az érzelmi tartalmak változatosabb kifejezése a beszédben a nagyobb (érzelmi) aktivitás, az extrovertáltság percepcióját is eredményezi, ami a társas viszonyrendszerekben esetleg nagyobb dominanciához is vezethet.

Irodalom

- Boersma, P. – Weenink, D. 2018. PRAAT: Doing phonetics by computer. v 6.1. Elérhető: <http://www.fon.hum.uva.nl/praat/> (letöltés ideje: 2019. 02. 15.)
- Bradley, M. M. – Lang, P. J. 1994. Measuring emotion: the Self-Assessment Manikin and the Semantic Differential. *Journal of Behavior Therapy and Experimental Psychiatry* 25(1): 49–59.
- Degottex, G. – Kane, J. – Drugman, T. – Raitio, T. – Scherer, S. 2014. COVAREP – A collaborative voice analysis repository for speech technologies. In: *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 960–964.
- Ekman, P. 1992. An argument for basic emotions. *Cognition and Emotion* 6(3-4): 169–200.
- Fontaine, J. J. R. – Scherer, K. R. – Soriano, C. (szerk.) 2013. *Components of emotional meaning: A sourcebook*. New York, NY, US: Oxford University Press.
- Fontaine, J. J. R. – Scherer, K. R. – Roesch, E. B. – Ellsworth, P. C. 2007. The world of emotions is not two-dimensional. *Psychological Science* 18(12): 1050–1057.
- Geifman, N. – Cohen, R. – Rubin, E. 2013. Redefining meaningful age groups in the context of disease. *Age* 35(6): 2357–2366.
- Germine, L. – Nakayama, K. – Duchaine, B. C. – Chabris, C. F. – Chatterjee, G. – Wilmer, J. B. 2012. Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin & Review* 19(5): 847–857.

- Horton, J. J. – Rand, D. G. – Zeckhauser, R. J. 2011. The online laboratory: conducting experiments in a real labor market. *Experimental Economics* 14: 399–425.
- Jürgens, R. – Grass, A. – Drolet, M. – Fischer, J. 2015. Effect of Acting Experience on Emotion Expression and Recognition in Voice: Non-Actors Provide Better Stimuli than Expected. *Journal of Nonverbal Behavior* 39(3): 195–214.
- Kuznetsova, A. – Brockhoff, P. – H. B. Christensen, R. 2015. LmerTest: Tests in linear mixed effects models. *R Package Version 2*.
- Leeuw, J. R. de 2015. jsPsych: a JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods* 47(1): 1–12.
- Mehrabian, A. – Russell, J. A. 1974. *An approach to environmental psychology*. Cambridge, MA, US: The MIT Press.
- Murray, I. R. – Arnott, J. L. 1993. Toward the simulation of emotion in synthetic speech: a review of the literature on human vocal emotion. *The Journal of the Acoustical Society of America* 93(2): 1097–1108.
- Patel, S. – Scherer, K. R. – Björkner, E. – Sundberg, J. 2011. Mapping emotions into acoustic space: the role of voice production. *Biological Psychology* 87(1): 93–98.
- R, Core Team 2018. R: A language and environment for statistical computing. v. 3.5.3. Elérhető: <https://www.r-project.org>.
- Russell, J. 1980. A Circumplex Model of Affect. *Journal of Personality and Social Psychology* 39(6): 1161–1178.
- Scherer, K. R. 1989. Vocal correlates of emotional arousal and affective disturbance. In: H. Wagner – A. Manstead (szerk.) *Handbook of social psychophysiology*. Wiley handbooks of psychophysiology. Oxford, England: John Wiley & Sons. 165–197.

Nem kompozicionális igekötős igék argumentumszerkezetének korpuszalapú vizsgálata

Gyulai Livia

SZTE BTK Nyelvtudományi Doktori Iskola
liviagyulai95@gmail.com

Kivonat: A jelen cikkben egy korpuszalapú vizsgálatot mutatok be, melyben kísérletet teszek a nem kompozicionális igekötős igék azonosítására pusztán szintaktikai módszerek segítségével. Ehhez elméleti háttérként a klasszikus kompozicionalitásfogalom egy általánosított, aszemantikus verzióját használom fel egy skaláris argumentumszerkezeti modellel (Szécsényi 2019) kombinálva. Az igekötők által az igei argumentumszerkezetben okozott változás lehet szisztematikus és nem szisztematikus. Hipotézisem szerint a nem szisztematikus, egyedi változásokból arra a következtetésre lehet jutni, hogy az adott igekötő-ige pár rendelkezik nem kompozicionális olvasattal. Az említett változások tanulmányozásához az igei bővítménymintázatok eseteit mint könnyen hozzáférhető grammatikai jellemzőt használtam. A kutatásban bemutatott módszer egy a későbbiekben automatikusan működő eszköz alapjaként szolgálhat.

1 Bevezetés

A nem kompozicionális kifejezések összegyűjtése igen nehéz feladat, hiszen ezen kifejezések lexikai reprezentációjához, jelentéséhez nem férhetünk közvetlenül hozzá a nyelvi adatokból. A tanulmányban egy, a nem kompozicionális igekötős igék argumentumszerkezetére irányuló kutatást mutatok be, melyhez korpuszból nyert adatokat és a kompozicionalitás elvének egy aszemantikus verzióját használtam.

A kompozicionalitás a szemantikai vizsgálódások alapja, ha azt szeretnénk megmagyarázni, hogy hogyan rendelhetünk jelentéseket a szintaxis által elvben létrehozható végtelen sok mondat mindegyikéhez. Már a problémafeltevésből is látszódik, hogy a szemantika és a szintaxis között, ha a kompozicionalitást tekintjük, szoros kapcsolat van. Maga a Frege-elv is erre a párhuzamosságra utal (lásd pl. Partee 2004): bármely nyelvi kifejezés jelentése meghatározható az alkotórészeinek jelentéséből és azok kapcsolódási módjából.

A tanulmányban a kompozicionalitás fogalmának egy általánosított, aszemantikus változatát veszem alapul, így az összetett kifejezések jelentésének kompozicionalitását azok formai jegyei és formális viselkedése, azaz szintaxisának nézőpontjából vizsgálom. Fő kérdésem tehát az, hogy pusztán szintaktikai eszközökkel meg lehet-e jósolni,

Gyulai Livia: Nem kompozicionális igekötős igék argumentumszerkezetének korpuszalapú vizsgálata. In Várad Tamás (sorozatszerkesztő), Ludányi Zsófia, Grácsi Tekla Etelka (szerkesztő): *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2019. XIII. Alkalmazott Nyelvészeti Doktoranduszkonferencia*. Budapest: MTA Nyelvtudományi Intézet. 2019. 45–60. DOI: 10.18135/Alknyelvdok.2019.13.4

hogy egy igekötő-ige párnak van-e nem kompozicionális olvasata. Ennek tanulmányozásához korpuszból származó adatokat használtam fel. Ahhoz, hogy megállapítsam, hogy egy igekötő-ige párnak van-e nem kompozicionális olvasata, azt vizsgáltam meg, hogy az igekötő megjelenése szisztematikus vagy nem szisztematikus módon változtatja-e meg az igei bővítményszerkezetet (az igekötőkről lásd bővebben például Komlósy 1992). Példaként említhető, hogy a *be* igekötő mellett rendszeresen jelenik meg *-bA* ragos bővítmény, azonban egyedi hatása, hogy a *rüg* igével összekapcsolódva megjelenhet mellette *-tÓl* ragos bővítmény (pl. *Peti berügött a sok pálinkától*).

Vizsgálatomat egy olyan módszer segítségével végeztem el, amely kiindulópontja lehet a későbbiekben egy automatikusan működő eszköznek, mellyel kísérletet tehetünk a nem kompozicionális olvasattal rendelkező igekötős igék meghatározására.

A tanulmány a következőképpen épül fel: először a kompozicionalitás általánosított fogalmát mutatom be, majd ismertetem azt a skaláris argumentumszerkezet-fogalmat, amelyet a kutatás során használtam. Ezután tárgyalom a kutatás során kitűzött célokat, illetve a hipotéziseket. Bemutatom a korpuszt, az annotálási folyamatot és a módszert, amivel dolgoztam. Végül ismertetem az eredményeket és a jövőbeli célkitűzéseimet.

1.1 Kompozicionalitás

1.1.1 Klasszikus kompozicionalitásfogalom

A fregei kompozicionalitásfogalom (lásd korábban a Bevezetésben) alapján az a kijelentés tehető, hogy az adott összetett kifejezés szemantikai és szintaktikai szerkezete homomorf, azok leírásának homomorfizmusát állítja ez az elv. A természetes nyelvekben azonban nem csupán kompozicionális összetett kifejezések vannak. A klasszikus kompozicionalitási elv szerint ezek a szerkezetek vagy egyáltalán nem tekinthetők összetett kifejezésnek, vagy rögzített hangalakkal rendelkező összetett kifejezések. Amennyiben a második eset áll fenn, egyedi összetételi módról van szó, melyben a részkiefejezések egyedi jelentése nincs hatással az összetett kifejezés jelentésére. A két megközelítés azonos nézetet képvisel abból a szempontból, hogy a nem kompozicionális konstrukciók önálló lexikai elemnek tekintendők.

1.1.2 Aszemantikus kompozicionalitásfogalom

Ahhoz, hogy a nem kompozicionális olvasattal rendelkező kifejezések azon részét meghatározzuk, amelyeknek nincs kompozicionális olvasatuk, szükség van egy olyan kompozicionalitásfogalomra, amely elsősorban nem a kifejezések jelentésére támaszkodik. Ehhez módosítani, általánosítani kell a szemantikában használatos klasszikus fregei definíciót (Szécsényi–Gyulai 2019) a következőképpen: egy összetett kifejezés viselkedése kiszámítható a mondatban szereplő kifejezések grammatikai tulajdonságai és azok kapcsolódási módja alapján.

A definícióban szereplő grammatikai tulajdonság lehet egy könnyebben hozzáférhető tulajdonság is.

Kutatásom során elsősorban az igekötős igék bővítményszerkezetében megjelenő bővítmények esetét vagy névutóját használtam a kompozicionalitás tanulmányozásához, azonban célszerűnek tűnt bizonyos eseteket összevontan kezelni, ezért a dependencia elemzés bizonyos címkéit is figyelembe vettem (pl. TO, FROM, LOCY, TFROM, TLOCY, TTO). Ezeknek a címkéknek a használata hasznos például az olyan szerkezetek vizsgálatánál, mint *születik*, *lakik*. A Mazsola korpuszlekérdező eszköz is a bővítmények esetét használta (Sass 2009). A klasszikus kompozicionalitásfogalommal szemben az aszemantikus megközelítés csak részben követeli meg az anyanyelvi

intuíció használatát, mivel pusztán szintaktikai tulajdonságokat vesz figyelembe. A kompozicionalitás ilyen vizsgálata azért hasznos, mert a nyelvi adatok értelmezését minimalizálva, azok viselkedésének tanulmányozásával is elkülöníthetjük a kompozicionális és nem kompozicionális szerkezeteket.

1.1.3 Igekötős igék kompozicionalitása

Az előző szakaszban bemutatott aszemantikus kompozicionalitásfogalmat alkalmazva tanulmányozható az igekötős igék kompozicionalitása a következőképpen: ha az igekötő és az ige független egyedi tulajdonságaiból megjósolható az igekötős ige viselkedése (bővítményeinek esete), akkor a szerkezete kompozicionálisnak tekinthető, ha azonban az összetevők egyedi tulajdonságai nem indokolják az ige mellett megjelenő bővítmények esetét, akkor nem kompozicionálisnak értelmezendők. Például az [1] mondatban az *el* és a *számol* együttes megjelenése egy *valameddig* típusú argumentumszerkezettel jelenik meg kompozicionálisan, és ez a bővítmenytípus az *el* igekötő esetében rendszeresen/szisztematikusan megjelenik több ige esetében, ezért az *elszámol* ige ezen előfordulását aszemantikusan kompozicionálisnak nevezem.

[1] Péter elszámolt százig.

[2] Péter mindig elszámol Marival, ha tartozik neki.

Azonban a [2] mondat nem kompozicionális, mivel sem az *el*, sem a *számol* nem vonz tipikusan *-val* ragos bővítmenyt. Ennek a bővítmenytípusnak a megjelenését ezért nem magyarázhatjuk másként, csak az *elszámol* ige egyedi lexikai tulajdonságaként. Ha a két *elszámol* mondatbeli jelentését vizsgáljuk, szintén ugyanezt állapíthatjuk meg, tehát a formai alapon hozott kompozicionalitási ítélet megjósolja a szemantikai kompozicionalitást is.

1.2 Skaláris argumentumszerkezet-fogalom

A klasszikus értelmezés szerint az argumentumszerkezet azt mutatja meg, hogy hány vonzata van az igének (argumentumszerkezeti variánsnak), illetve, hogy ezek milyen grammatikai tulajdonságokkal rendelkeznek. Azonban a vonzatság nem feltétlenül jelenti azt, hogy kötelezően meg kell jelennie az ige mellett, hiszen vannak olyan esetek, amikor nem jelenik meg a vonzat expliciten. A szabad bővítmenyt pedig nem egy, az ige mellett opcionálisan megjelenő szerkezetnek kell tekinteni, mivel esetenként gyakrabban, máskor ritkábban szerepelhetnek a mondatban. Ezt elfogadva tehát nem a klasszikus vonzat és szabad bővítmeny bináris oppozíciójára kell támaszkodnunk egy ilyen vizsgálat során. Hasonló megoldást javasol Kálmán (2016) is. Szerinte sem lehet kategorikusan eldönteni egy bővítmenyről, hogy az vonzat vagy szabad bővítmeny, hanem egy graduális megkülönböztetést ismertet.

Vizsgálatom során Szécsényi (2019) skaláris argumentumszerkezeti megközelítést követem. Ez alapján egy spektrumon értelmezem ezeket a vonzat és a szabad bővítmeny fogalmakat a korpuszban való megjelenésükből származó gyakorisági értékek alapján. A vizsgált igék összes bővítményéhez rendelt gyakorisági érték alapján jön létre az ige argumentumszerkezetét jellemző skaláris értékekből álló vektor. A skála, melyen a gyakorisági értékeket elhelyezhetjük, egyik végpontja a MINDIG MEGJELENIK, a másik pedig a NEM JELENIK MEG, így azt kell jellemeznünk, hogy valami inkább vonzat vagy inkább szabad bővítmeny. Jelen tanulmánynak nem célja meghatározni, hogy a gyakorisági értékek alapján hol lehetne határt húzni a klasszikus értelemben használt

vonzat és szabad bővítmény fogalmak között, maga az elméleti keret sem ad meg konkrét értéket ezen kategóriák meghatározásához. A későbbiekben ugyan megjelennek ezek a fogalmak a tanulmányban, azonban csupán abból a célból, hogy az olvasónak könnyebb legyen azonosítani, hogy az adott bővítménytípus melyik bináris kategóriába tartozhatna. Az 1. ábrán látható a skaláris argumentumszerkezet-fogalom modellje.

A spektrumon a hagyományosan vonzatnak nevezett bővítmény a MINDIG MEGJELENIK végponthoz közel jelenik meg. Azonban nem lehet a vonzatságot a skála végpontjának tekinteni, mivel maga a klasszikus értelemben vett vonzat is olyan bővítmény, mely például egy alanyi vagy tárgyi névmási törlés esetén elmaradhat a mondatban az ige mellől. Ebben az esetben nem arról van szó, hogy nincs jelen a mondatban vonzat, hiszen a kontextus alapján ki lehet következtetni, hogy milyen bővítménynek kellene ott szerepelnie. A szabad bővítmények a skála egy terjedelmes részén mozoghatnak attól függően, hogy milyen előfordulási gyakoriságot mutatnak az egyes bővítménytípusok a korpuszban. Vannak olyan bővítménytípusok is, amelyek a skála jobb széléhez közel állnak, ezek szinte soha nem jelenhetnek meg (pl. *Mari kitekint *a kocsióért*). Ezeket a skaláris jellemzőket korpusz alapján határozhatjuk meg, hiszen így lehet az egyes argumentumszerkezeteket összesíteni, és olyan kvantitatív kutatást végezni, melynek eredményeképpen az arányokat egy spektrumon tudjuk értelmezni.



1. ábra. Skaláris argumentumszerkezeti modell

A korpuszban szereplő igeik a bővítményei közötti viszony skaláris minősítése alapján különböző argumentumszerkezeti variánsokba sorolhatóak. Az ige mellett megjelenő bővítmények gyakorisági értéke az ige egyedi jellemzője, nevezhetjük az ige ujjlenyomatának is.

Ezek alapján elhagyható a vonzat és a szabad bővítmény fogalom, helyette ezekkel a skaláris értékekkel jellemezhetők az igei bővítményszerkezetek.

Az, hogy egy igeikötő megjelenése hatással van az igei argumentumszerkezetre, már intuitív módon is belátható, ha az [3] és [4] példát összevetjük. [3]-ban látható az igeikötő nélküli, [4]-ben pedig az igeikötős előfordulása a *készül* igeének:

[3] *Peti készül a randira.*

[4] *Peti kikészült a randitól.*

Ezen példák alapján látható, hogy a *ki* igeikötő megjelenése az igeikötő nélkülihez képest bevezethet egy új ablatívuszi bővítményt.

Az, hogy az igeikötő milyen további hatással lehet egy ige bővítményszerkezetére, korpuszvizsgálatot igényel, ugyanis az így kapott előfordulási gyakorisági értékek alapján lehet további állításokat tenni. A vizsgálat során az volt tapasztalható, hogy az igeikötő megjelenése egy teljesen új bővítmény bevezetésén kívül megerősítheti, gyengítheti, esetleg letilthatja bizonyos bővítmények jelenlétét az igeikötő nélküli argumentumszerkezetekhez képest.

1.3 Célkitűzések, hipotézisek

Célkitűzéseim között szerepel az *el* igekötő korpuszalapú vizsgálatának példáján annak bemutatása, hogyan változtatja meg nem szisztematikusan az igekötő megjelenése az igekötő nélküli argumentumszerkezetet. Ezt az igekötők nem megjósolható viselkedésének bemutatása alapján teszem. A cikk következő szakaszában egy olyan módszert mutatok be, amely a Szécsényi (2019) által javasolt skaláris argumentumszerkezeti megközelítésre támaszkodva, pusztán szintaktikai módszerek segítségével tesz kísérletet a nem kompozicionális igekötős igék azonosítására.

Vizsgálatom céljaként azt tűztem ki, hogy egyedi argumentumszerkezet-változtató hatását tanulmányozzam az *el* igekötőnek. Hipotézisem szerint az igekötő megjelenése okozhat szisztematikus változást az igei argumentumszerkezetben, ezekben az esetekben az aszemantikus kompozicionalitásfogalom szerint az igekötő kompozicionálisan kapcsolódik az igéhez. Azonban amennyiben az általa okozott változás nem szisztematikus, hanem esetleges, akkor arra következtethetünk, hogy az az igekötő-ige pár nem kompozicionális összetett kifejezés, melyet egy lexikai egységként kell felvennünk a lexikonba. Ennél a vizsgálatnál arra számíthatunk, hogy inkább nem szisztematikus változások fognak kirajzolódni az adatok alapján, illetve, hogy a szintaktikailag nem szisztematikus változáson átment igekötős ige jelentésében sem kompozicionális.

2 Módszertan

2.1 A korpusz

Az előző szakaszban kitérőt céllok és hipotézisek vizsgálatához hét igét (*halad, jár, lát, készít, készül, felejt, számol*) és az *el* igekötőt választottam ki. Ezeknek az igéknek az összes igekötő nélküli és *el* igekötős előfordulását kértem le a Szeged Dependencia Treebankből (Vincze et al. 2010). Az igekötőnek az elváló és nem elváló előfordulásait is figyelembe véve a korpuszom összesen 2048 igei előfordulást tartalmazott. Nagy segítség volt a kutatás során, hogy olyan korpuszt használtam, ami függőségi nyelvtannal lett előelemézve, hiszen így az igével függőségi viszonyban lévő kifejezések már jelölve voltak: vonzatok, szabad határozók, mondathatározók, módosítószók, mellékmondatok. Így az is rögtön láthatóvá vált, hogy az igékhez kapcsolódó bővítmények milyen morfológiai és szintaktikai tulajdonsággal rendelkeznek (jelen kutatásnál pl.: eset).

2.2 Az annotálási folyamat

Az annotáció célja a Szeged Dependencia Treebankből lekért mondatokba való olyan információk kódolása volt, melyek segítségével az igekötők szisztematikus, illetve nem szisztematikus argumentumszerkezet-változtató hatását nyomon tudtam követni. A korpuszt alkotó 2048 igei előfordulást tartalmazó mondatokon manuális annotációt végeztem az MMAX2 (Müller–Strube 2006) annotáló szoftver segítségével. A Szeged Dependencia Treebank szintaktikai előelemzettsége miatt az annotáló program segítségével ki voltak emelve az igék és a hozzájuk tartozó bővítmények is.

Az annotálás előtt kigyűjtöttem, hogy milyen argumentumszerkezetekkel jelentek meg a fent felsorolt igék a korpuszban mind igekötővel, mind igekötő nélkül. Az annotáció első lépéseként megállapítottam, hogy az egyes mondatokban előforduló igék milyen argumentumszerkezettel szerepeltek. A korpuszban lévő mondatokban voltak olyan esetek, amikor a bővítmény nem volt expliciten kitéve, ilyenkor a kontextus alapján lehetett kikövetkeztetni, hogy milyen bővítménynek kéne állnia az ige mellett. Ezután megnéztem, hogy igekötős-e az ige, végül bejelöltem azokat a bővítményeket a mondatokban, amelyek közvetlenül az igehez kapcsolódnak, tehát nem mondat szintű összetevők (például: *sajnos*, *szerintem*, *valószínűleg* stb.). Az annotálást segítette, hogy a függőségi elemzés a legtöbb bővítményt azonosította, azonban mind a mondat, mind a predikátum szintű módosítókat az igehez kapcsolta, és nekem ezek közül csak a predikátum szintűekre volt szükségem a kutatáshoz.

2.3 Adatfeldolgozás

A manuális annotációt követően az MMAX2-ben kódolt információk egy Excel-táblázatba lettek visszafejtve. A 2. ábrán egy részlet látható a táblázat azon részéből, amely a nyers információkat tartalmazza.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	egy	word	lemm	sub	vm	VN	NC	AC	DA	IN	HK	va	bá	bá	bá	hc	ná	tó	ra	or
2	1	w01_49	felejt	2	el	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	1	w01_30	felejt	2	el	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
4	1	w01_96	felejt	2	el	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	1	w01_34	felejt	2	el	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	1	w01_45	felejt	2	el	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
7	1	w01_52	felejt	2	el	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
8	1	w01_52	felejt	4	el	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
9	1	w01_42	felejt	3	none	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
10	1	w01_42	felejt	4	el	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
11	1	w01_57	felejt	3	none	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
12	1	w01_36	felejt	2	none	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
13	1	w01_52	felejt	4	none	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
14	1	w01_43	felejt	2	el	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
15	1	w01_59	felejt	2	el	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
16	1	w01_26	felejt	2	el	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
17	1	w01_51	felejt	2	el	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
18	1	w01_24	felejt	2	el	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	1	w01_13	felejt	2	el	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0

2. ábra. A visszafejtett nyers információk táblázatának részlete

Ez a táblázat tartalmazza az összes információt az összes mondat tekintetében. A B oszlopban található az igék egyedi azonosítói, a C oszlopban a mondatokban szereplő ige szótöve. A következő oszlopban (D) látható számok azt mutatják, hogy az ige az adott mondatban melyik argumentumszerkezeti variánssal szerepel. A korpuszban szereplő különböző argumentumszerkezeti variánsokat az annotáció előtt intuíción alapján azonosítottam, mely az annotáló szoftverrel való munkát segítette, hiszen néhány előre azonosított kategória már rendelkezésre állt, és ha új argumentumszerkezettel találkoztam, akkor bővítettem a listát. A jövőbeli célkitűzések között szerepel az argumentumszerkezeti variánsok automatikus meghatározása (Szécsényi 2019). A D oszlopban található számok az adott argumentumszerkezeti variáns helyét mutatja a variánsokat

tartalmazó listán. Az E oszlop azt az információt tartalmazza, hogy az ige igekötős-e. Az F oszlopban 0 és 1 jelöli, hogy az adott mondatban szerepelt-e elváló igekötő, vagy sem. Az ezután következő oszlopok az ige mellett megjelenő bővítmények jellemzésére szolgálnak, azt adják meg, hogy hány bővítmény szerepel a mondatban a felsorolt bővítménytípusból, melyeket a főneveknél az esetekkel (NOM, ACC, DAT stb.) jellemeztem.

A nyers információkat tartalmazó táblázatból ezután automatikus módon hozhatók létre igékre lebontott összesítő táblázatok, melyek alapján megfigyelhető az igekötő egyes igékre gyakorolt argumentumszerkezet-változtató hatása. A 3. ábrán látható a *számol* ige táblázatából egy részlet. A felső sorokban, ahol a B oszlop üres, ott az összes argumentumszerkezeti variánusra vonatkozó adatokat láthatjuk.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	tő	variá	megjegyzé	vm	db	pv	VM	NOM	ACC	DAT	INF	HKM	val
2	számol			*	113		6	35	14	2	0	16	82
3							0,05	0,31	0,12	0,02	0	0,14	0,73
4				none	93		1	32	11	1	0	16	66
5							0,01	0,34	0,12	0,01	0	0,17	0,71
6				el	20		5	3	3	1	0	0	16
7							0,25	0,15	0,15	0,05	0	0	0,8
8													
9		1		*	13		1	2	2	0	0	0	0
10							0,08	0,15	0,15	0	0	0	0
11		1		none	12		0	1	2	0	0	0	0
12							0	0,08	0,17	0	0	0	0
13		1		el	1		1	1	0	0	0	0	0
14							1	1	0	0	0	0	0

3. ábra. A *számol* ige táblázatának részlete

Az összesítő táblázatokat a következő módon foglaltam össze: összevontam azokat a sorokat, melyekben az ige ugyanazzal az argumentumszerkezeti variánssal szerepelt, így megkaptam összesítve, hogy mennyi előfordulás tartozik egy variánshoz. Összegeztem azt is, hogy ezeknél a különböző bővítménytípusok hányszor szerepeltek, így azt is ki tudtam számolni (elosztva a bővítmények előfordulási számát az argumentumszerkezeti variánshoz tartozó összes mondat számával), hogy milyen gyakorisággal fordultak elő az egyes argumentumszerkezeti variánsoknál a különböző esetű bővítmények. Az így megkapott gyakorisági értékek a korábban már említett skaláris értékek.

A táblázat felső sorában hasonló fejléc látható, mint a 2. ábrán látható nyers információkat tartalmazó táblázatban. Az egyes igék táblázatát tanulmányozva először az egy igével előforduló összes argumentumszerkezeti variánssal vonatkoztatva kiszámoltam, hogy hányszor és milyen gyakorisággal szerepelnek a különböző bővítménytípusok (2–3. sor), illetve azt is, hogy igekötő nélkül (4–5. sor) és *el* igekötővel (6–7. sor) milyen gyakoriság jellemzi. Ezután variánsokra lebontva is megadtam ugyanezt (az első variáns adatai itt a 9–14. sorban találhatóak). A gyakorisági adatok nagyságát színárnyalatokkal is jelöltem, amely segítette a kiugróan magas arányú előfordulások detektálását. Az igekötők szisztematikus (kompozicionális) és nem szisztematikus (nem

kompozicionális) argumentumszerkezet-változtató hatásának vizsgálatát ezen táblázatokban szereplő adatok alapján végeztem el.

A következőképpen értékeltem ki a táblázatokban összesített adatokat: összehasonlítottam, hogy az igekötő nélküli és az igekötős sorokban milyen bővítménygyakorisági különbségek fedezhetők fel. Először megnéztem azt, hogy az összes argumentumszerkezet igekötős, illetve igekötő nélküli előfordulásai között milyen különbségek mutatkoztak meg (5. és 7. sor). Erre azért volt szükség, mert így egy általános képet kaptam arról, hogy milyen tendenciaszerű változások történtek az igei argumentumszerkezetben az igekötő megjelenésével. Ez egyfajta iránymutató volt arra vonatkozóan, hogy milyen különbségeket kell majd keresnem az argumentumszerkezeti variánsoknál, melyik variánsnál látható különbség okozza az összesített résznél is látható változást.

Ezután megvizsgáltam, hogy argumentumszerkezet-specifikusan található-e valamilyen különbség az igekötő nélküli és az igekötős előfordulások között. Abban az esetben, ha egyedi, nem szisztematikus változás látható az igekötő nélküli és az igekötős előfordulások között egy argumentumszerkezeti variáns esetén, akkor arra lehet következtetni, hogy az az igekötő-ige pár rendelkezik nem kompozicionális olvasattal. Ha nem csak az éppen vizsgált ige egy argumentumszerkezeti variánsánál fedezhető fel egy bizonyos változás az argumentumszerkezetben az igekötő megjelenésével, hanem szisztematikus módon változtatja meg az igekötő az igei bővítményszerkezetet más igeéknél is, akkor kompozicionálisnak tekinthető az adott igekötő-ige pár (Gyulai 2019).

Néhány argumentumszerkezeti variáns esetében egyes igekötős változatok olyan kevésszer szerepeltek a korpuszban, hogy az alapján nem lehetett általánosításokat végezni, ezért az 5 alatti előfordulásokat már nem értékeltem ki.

3 Eredmények

Vizsgálatom során az alanyi bővítmény gyakorisági változását nem vettem figyelembe. Ez a bővítménytípus a kutatás szempontjából nem volt releváns, mivel minden vizsgált ige elvár maga mellé egy alanyt abban az esetben is, ha nem igekötős, és abban az esetben is, ha az. Továbbá nem vettem figyelembe azokat az adatokat sem, ahol öt alatt volt az előfordulások száma. A következő szakaszokban minden vizsgált igenél sorra veszem, hogy hogyan változtatja meg az *el* igekötő megjelenése az igekötő nélküli argumentumszerkezetet. Megjegyzendő, hogy az ebben a fejezetben bemutatott eredmények korpuszadatokat tükröznek. A vizsgálat eredményeként kapott adatok közül azokat emelem ki, amelyek a vizsgálat szempontjából relevánsak. Mindegyik táblázat tartalmazza az összes argumentumszerkezetre vonatkozó adatot, majd azokat az argumentumszerkezeti variánsokat szemléltetem, amelyek jól illusztrálják az *el* igekötő hatását/hatásait az argumentumszerkezetre. Minden szakasz végén összegzem, hogy milyen hatást/hatásokat figyelhettük meg az igekötő megjelenésének. Bizonyos esetekben egy ige kapcsán több, egymástól független hatást is meg lehetett állapítani, ezért szerepel egyes igeéknél akár két formális összegzés is a szakasz végén. Az *el* után szereplő szám az alsó indexben a különböző hatások elkülönítésére és a Következtetések fejezet összegző táblázatában való könnyebb eligazodást szolgálja. A következő szimbólumokat használom az igekötő argumentumszerkezet-változtató hatásainak formális leírásához:

- \uparrow (például $\langle \text{VAL } \uparrow \rangle$ = megnövelte az igekötő megjelenése a *-vAl* ragos bővítmény gyakoriságát, amely már eddig is jelen volt az argumentumszerkezetben);
- \downarrow (például $\langle \text{RÓL } \downarrow \rangle$ = csökkentette az igekötő megjelenése a *-rÓl* ragos bővítmény gyakoriságát, amely már eddig is jelen volt az argumentumszerkezetben);
- $+$ (például $\langle \text{RA } + \rangle$ = behozza az igekötő az eddig az argumentumszerkezetben nem szereplő *-rA* ragos bővítményt az argumentumszerkezetbe);
- 0 (például $\langle \text{BÓL } 0 \rangle$ = letiltja, az eddigi előfordulási gyakorisághoz képest valószínűtlenné válik a *-bÓl* ragos bővítmény az igekötő megjelenése);
- 1 (például $\langle \text{BA } 1 \rangle$ = kötelezővé teszi az igekötő megjelenése a *-bA* ragos bővítményt, mely már szerepelt az argumentumszerkezetben).

3.1 (El)lát

A *lát* ige táblázatában az összes argumentumszerkezeti variánst tekintve látható, hogy igen magas arányú volt mind az igekötő nélküli, mind az igekötős előfordulásoknál az akkuzatívuszi bővítmények aránya. Szembetűnő ezen kívül az igekötős előfordulásoknál a *-vAl* ragos bővítmények kiugró aránya az igekötő nélküliekhez képest, tehát ez a bővítménytípus tiltottból megengedetté válik. Ezeket a változásokat az 1. táblázat szemlélteti. Az összes argumentumszerkezetet tekintve látható, hogy a tárgyi bővítmények előfordulási gyakorisága megnőtt az *el* igekötő megjelenésével.

		db	ACC	-vAl
összes	igekötő nélkül	225	0,72	0,01
	<i>el</i> igekötővel	126	0,89	0,34
<i>valaki (el)lát valamivel</i>	igekötő nélkül	0	--	--
	<i>el</i> igekötővel	41	0,8	0,98

1. táblázat. Az *(el)lát* ige táblázatának részlete

A fentebb említett változás okát a *valaki ellát valakit valamivel* argumentumszerkezetnél találhatjuk. Látható ennél a variánsnál, hogy az *el* igekötős előfordulásoknál igen magas (0,98) arányban fordult elő a *-vAl* ragos bővítménytípus, illetve egyáltalán nincs ennél a variánsnál igekötő nélküli előfordulás ebben a korpuszban. Ebből arra következtethetünk, hogy az igekötő megjelenése szinte kötelezővé teszi ennek a bővítménytípusnak a jelenlétét, és letiltja az igekötő nélküli előfordulásokat.

$el_1: \langle \text{ACC } \uparrow, \text{VAL } 1 \rangle$

3.2 (El)halad

Az *el* igekötő *halad* ige argumentumszerkezetére való hatásait a 2. táblázat tartalmazza:

		db	LOCY
összes	igekötő nélkül	86	0,06
	<i>el</i> igekötővel	16	1
<i>valaki (el)halad valami mellett</i>	<i>el</i> igekötővel	16	1

2. táblázat. Az *(el)halad* ige táblázatának részlete

A *halad* ige esetén az *el* igekező megjelenése nem okozott különösebb változást az argumentumszerkezetben. Az *el* igekezővel csak a *valaki elhalad valahol* argumentumszerkezettel szerepelt a korpuszban és igekező nélkül is megengedett a helyhatározói bővítménytípus (LOCY). Az *el* igekező megjelenése az eddig szabad bővítmény státuszú bővítményt vonzattá tette. A visszafejtett adatok Excel-táblázat (2. ábra) fejlécében szereplő bővítménytípusok között nem szerepel külön a névutó, azonban az annotáció során látható volt, hogy az igekezős ige csak a *valaki elhalad valami mellett* konstrukcióban szerepelt.

*el*₂: <LOCY 1>

3.3 (El)készít, (el)készül

A *készít* ige esetén (3. táblázat) a tárgyi bővítmények magas előfordulási gyakorisága volt jellemző mind az igekező nélküli, mind az igekezős igék esetén.

Nagyon kis mértékben (0,05) csökkent az ACC típusú bővítmények előfordulási gyakorisága az igekező megjelenésével. Ezenkívül az látható, hogy a *-rÓl* típusú bővítmény, mely egy nem közvetlenül az igehez kapcsolódó, de amellel megjelenő önállóan viselkedő bővítmény, az igekező nélküli előfordulásoknál szabad bővítmény státuszú volt, azonban az igekező megjelenése letiltotta ezt a bővítményt. Az adatok alaposabb vizsgálatából kiderült, hogy a *-rÓl* bővítmény nem az ige bővítménye volt, hanem csak bizonyos típusú tárgyak mellett jelent meg, például *valaki képet készít valamiről*, de *valaki tortát készít *valakiről*. Az *el* igekező megjelenése kapcsán megfigyelhetjük annak perfekтивizáló hatását, ez azonban nem magyarázza a *-rÓl* bővítménytípus eltűnését. Az összes igekező nélküli és igekezős előfordulás szinte kivétel nélkül ugyanahhoz az argumentumszerkezeti variánshoz tartozott. A *készül* ige adatait a 4. táblázat tartalmazza.

*el*₃: <RÓL 0>

		db	ACC	-rÓl
összes	igekező nélkül	168	0,96	0,26
	<i>el</i> igekezővel	64	0,91	0

3. táblázat. Az *(el)készít* ige táblázatának részlete

Látható, hogy míg a *készít* ige esetén megengedett volt az igekező nélküli és az igekezős igéknél is a tárgyi bővítmény jelenléte, a *készül* ige esetén pedig az ige tárgyatlan mivolta miatt nem szerepelt ilyen típusú bővítmény. Ha a *-rA* bővítményt az összes argumentumszerkezeti variánst tekintve megvizsgáltuk, látható volt, hogy, míg az igekező nélküli igék esetén szabad bővítmény státuszú volt ez a bővítménytípus, addig az *el* igekező megjelenése lecsökkentette az előfordulás gyakoriságát. A delatívuszi (*-rÓl*) bővítményt tekintve hasonló tendencia mutatkozott meg, mint ami a *-rA* bővítménynél volt látható, azzal a különbséggel, hogy itt az igekező nem csupán lecsökkentette a delatívuszi bővítményt, hanem le is tiltotta, ugyan az igekező nélküli előfordulásoknál sem volt jelen túl magas arányban.

		db	ACC	-rA	-rÓI
összes	igekötő nélkül	273	0,09	0,43	0,1
	<i>el</i> igekötővel	100	0	0,19	0,01
<i>valami (el)készül</i>	igekötő nélkül	107	0,02	0,08	0,24
	<i>el</i> igekötővel	90	0	0,18	0,01

4. táblázat. Az *(el)készül* ige táblázatának részlete

A *valami készül* argumentumszerkezeti variánst érdemes tüzetesebben megvizsgálni, hiszen itt jelent meg majdnem az összes *el* igekötős ige. Argumentumszerkezet-specifikusan a tárgyi bővítményt tekintve nem történt változás, ugyanúgy nem volt jellemző ez a bővítménytípus, mint az összes argumentumszerkezetet tekintve. A szublatívuszi (-rA) bővítményt tekintve azonban felcserélődtek az arányok, hiszen az igekötő megjelenése ennél az argumentumszerkezeti variánsnál növelte az ilyen típusú bővítmény előfordulási gyakoriságát. Ami a delatívuszi (-rÓI) bővítményeket illeti, határozottabban kirajzolódott az argumentumszerkezet-változtató hatása az igekötőnek, mint az összesített értékeknél, azonban itt nagyobb csökkenésről beszélhetünk, mivel argumentumszerkezet-specifikusan nagyobb arányban volt jelen ez a típusú bővítmény az igekötő nélküli előfordulásoknál, mint összesítve.

Megfigyelhető továbbá a *valami készül* variánsnál, hogy amennyiben az igekötő nélküli előfordulásoknál volt delatívuszi (-rÓI) esetű bővítmény az ige mellett, akkor szublatívuszi (-rA) nem lehetett, illetve, ha jelen volt az *el* igekötő, akkor fordított a helyzet: ha volt szublatívuszi (-rA) esetű bővítmény, nem lehetett delatívuszi (-rÓI). Ezek alapján elmondható, hogy ez a két bővítménytípus kiegészítő disztribúcióban áll egymással.

Összevetve a két ige táblázatát azt a mintázatot fedezhetjük fel, hogy mindkét ige esetében hasonló előfordulási gyakoriságról (0,26 és 0,24) tiltotta le az igekötő megjelenése a delatívuszi bővítményt. A két ige közötti kapcsolat miatt tapasztalhatjuk ezt a jelenséget.

$$el_4: \langle RA \uparrow, RÓL \ 0 \rangle$$

3.4 (El)számol

A *számol* ige táblázatát (5. táblázat) tanulmányozva látható, hogy a -vAI ragos bővítmények előfordulási gyakorisága igen magas volt az összes argumentumszerkezeti variánst tekintve mind az igekötő nélküli, mind az igekötős igék esetén.

		db	-vAI
összes	igekötő nélkül	93	0,71
	<i>el</i> igekötővel	20	0,80
<i>valaki (el)számol valamivel</i>	igekötő nélkül	52	1,02
	<i>el</i> igekötővel	15	1,07

5. táblázat. Az *(el)számol* ige táblázatának részlete

Ebből arra a következtetésre juthatunk, hogy az igekötő megjelenése nem okozott formai különbséget az igekötő nélküli argumentumszerkezethez képest (*valaki számol*

valamivel vs. *valaki elszámol valakivel*), egy kicsivel megnövelte az ilyen esetű bővítmenyek jelenlétét.

Ha megfigyeljük argumentumszerkezet-specifikusan ezt a bővítmenytípust, látható, hogy az összes igekötő nélküli és igekötős előfordulásnál jelen volt a mondatban, tehát kötelező vonzat státuszban volt. Ennél az argumentumszerkezetnél formailag nem mutatkozott meg az igekötő argumentumszerkezet-változtató hatása, azonban intuitív módon tudjuk, hogy valamiféle jelentésváltozáson mégis átment az igekötős ige.

3.5 (El)jár

A *jár* igénél az *el* igekötő megjelenése (lásd 6. táblázat) megerősítette az inesszívuszi (-*bAn*) ragos bővítmeny jelenlétét az igei argumentumszerkezetben, illetve letiltotta a -*vAl* ragosakat.

A -*bA* ragos bővítmenyek aránya lecsökkent. A -*rA* ragos és a TO típusú bővítmenyeknél az összesített adatok alapján nem látható lényeges változás az *el* igekötő megjelenésével, azonban az egyes argumentumszerkezeti variánsok vizsgálatánál mutatkoztak különbségek.

A *jár* ige esetén azt a megfigyelést tehetjük, hogy jellemzően a *valaki jár valahol* és *valami jár valamivel* argumentumszerkezettel jelent meg a korpuszban, amely azonban a vizsgálat szempontjából nem releváns, mert nem jelent meg az *el* igekötő. Két argumentumszerkezeti variánsnál volt jelen az *el* igekötő. A *valaki jár* argumentumszerkezetnél az látható, hogy az *el* igekötős igénél megnövekedett a -*bAn* és a -*vAl* ragos bővítmenyek száma. Az annotáció folyamán látható volt, hogy ez a bővítmenytípus leginkább a korpusz jogi szövegeket tartalmazó részében fordult elő (például *valaki eljár valamilyen ügyben*).

A másik argumentumszerkezeti variáns, ahol az *el* igekötő megjelent, a *valaki eljár valahova*. A -*bAn* ragos bővítmeny tiltott státuszról megengedetté vált. A -*bA* bővítmenytípus előfordulási gyakorisága visszaesett, miközben a -*rA* és a TO bővítmenyeké megnőtt. Az *el* igekötő argumentumszerkezet-változtató hatásai közül ez tekinthető a legtipikusabbnak előzetes várakozásaink szerint. Az, hogy az igekötő megjelenésének hatására lecsökkent a -*bA* ragos bővítmenyek gyakorisága, azzal magyarázható, hogy a korpuszban nagy számban fordultak elő iskolai fogalmazások, melyekben az iskolába jár konstrukciók felülreprezentáltak.

		db	- <i>bAn</i>	- <i>vAl</i>	- <i>bA</i>	- <i>rA</i>	TO
összes	igekötő nélkül	229	0,16	0,31	0,14	0,05	0,03
	<i>el</i> igekötővel	73	0,34	0,07	0,03	0,07	0,04
<i>valaki (el)jár</i>	igekötő nélkül	46	0,07	0,02	0	0,04	0,02
	<i>el</i> igekötővel	60	0,38	0,07	0	0,02	0
<i>valaki (el)jár valahova</i>	igekötő nélkül	46	0	0,07	0,67	0,20	0,11
	<i>el</i> igekötővel	9	0,11	0,11	0,22	0,44	0,22

6. táblázat. Az (el)jár ige táblázatának részlete

$$\begin{array}{c}
 \text{els: } \langle \text{BAN } \uparrow, \text{ VAL } \uparrow \rangle \\
 \text{el}_6: \langle \text{BAN } +, \text{ BA } \downarrow, \text{ RA } \uparrow, \text{ TO } \uparrow \rangle
 \end{array}$$

3.6 (El)felejt

A 7. táblázatban láthatjuk, hogy milyen változásokat okozott az *el* igekötő a *felejt* ige argumentumszerkezetében.

A *felejt* ige igekötő nélküli előfordulására a tárgyi (0,61) bővítmények jelenléte volt a jellemző. Szintén jelen volt a tárgyi bővítmény, azonban az igekötő megjelenése kis mértékben lecsökkentette az ilyen típusú bővítményeket az igekötő nélküli előfordulásokhoz képest. Ezenkívül megfigyelhető, hogy míg az igekötő nélküli előfordulásoknál szinte tiltva voltak az infinitívuszi bővítmények, addig az igekötős előfordulásoknál megengedetté vált a jelenlétük. A *hogy* kötőszavas mellékmondatokat tekintve látható, hogy míg az igekötő nélküli előfordulásoknál nem volt megengedett a jelenléte, addig az *el* igekötős előfordulásoknál már szabad bővítmény státuszú volt. Ami a helyhatározói bővítményt (LOCY) illeti, megfigyelhető, hogy az igekötő nélküli előfordulások esetén engedélyezett volt a jelenléte, azonban az *el* igekötő megjelenése letiltotta azt. Az annotáció során látható volt, hogy az összes helyhatározói bővítmény az *otthon* igekötőszerű viselkedést mutató kifejezés volt.

		db	ACC	INF	HKM	LOCY
összes	igekötő nélkül	23	0,61	0,04	0	0,43
	<i>el</i> igekötővel	139	0,55	0,17	0,16	0,01
<i>valaki (el)felejt</i> + INF	igekötő nélkül	1	1	0	0	0
	<i>el</i> igekötővel	21	0,19	0,95	0	0
<i>valaki (el)felejt</i> + HKM	igekötő nélkül	0	--	--	--	--
	<i>el</i> igekötővel	24	0,17	0,08	0,92	0

7. táblázat. Az *(el)felejt* ige táblázatának részlete

A *valaki felejt* + INF argumentumszerkezeti variánsnál mutatkozott meg az infinitívuszi bővítmények igen magas aránya az *el* igekötő megjelenésével, melynek következtében az akkuzatívuszi bővítmények előfordulási gyakorisága csökkent. Ennél a variánsnál egyáltalán nem volt jelen *hogy* kötőszavas mellékmondat és a helyhatározói bővítmény is tiltott volt. Igekötő nélküli előfordulás ennél a variánsnál csupán egy volt, tehát az adatok ebben az esetben nem értékelhetők ki.

A *valaki felejt* + HKM argumentumszerkezeti variánsnál csak az *el* igekötő előfordulás szerepelt. Itt is a tárgyi bővítmények aránya körülbelül ugyanannyira csökkent, mint az előző variáns esetében. A *hogy* kötőszavas mellékmondatok aránya ezzel szemben igen kimagasló volt, míg az infinitívuszi bővítmény tiltott státuszba került, ugyanígy helyhatározói bővítmény sem szerepelt ennél a variánsnál.

*el*₇: <INF 1, LOCY 0>

*el*₈: <HKM 1, LOCY 0>

4 Következtetések

A következőkben összegzem és a 8. táblázatban szemléltetem, hogy milyen változások mentek végbe az igekötő megjelenésének hatására.

Nem szisztematikus változásként említhető, hogy új bővítménytípusként bevezette az *el* igekötő az argumentumszerkezetbe a *-bAn* ragos, az infinitívuszi bővítményt és

a *hogy* kötőszavas mellékmondatot. Megnövekedett az akkuzatívuszi, *-rA*, *-vAl*, *-bAn* ragos és a TO bővítménytípusok aránya az igekötő nélküli előfordulási gyakorisághoz képest. Csökkent az előfordulási gyakorisága a *-bA* ragos bővítménytípusnak, illetve teljesen letiltotta az igekötő megjelenése a *-rÓl* ragos bővítményt. Végezetül kötelezővé tette az igekötő megléte az újonnan megjelent infinitívuszi bővítményt és a *hogy* kötőszavas mellékmondatot, illetve a *-vAl* ragos és a helyhatározói bővítményeket.

Az 1.3 szakaszban megfogalmazott hipotézisnek megfelelően látható, hogy az igekötő megjelenése leginkább nem szisztematikus, egyedi változásokat okozott a kiválasztott igék argumentumszerkezetében, mivel nem jelent meg több igénél ugyanaz az argumentumszerkezet-változtató hatás. A talált adatok között azért fordulhattak elő sűrűbben nem kompozicionális igekötő-ige párok, mert ezek kötött frázisokká vált konstrukciók, így gyakrabban használjuk őket. A kompozicionális összetételek hiányát az igék kiválasztása is befolyásolta, mivel olyan igéket vizsgáltam, amelyek esetében intuíció alapján számítani lehetett nem kompozicionális argumentumszerkezeti variánsok megjelenésére. Mivel a jelen vizsgálat célja csak a módszer használhatóságának vizsgálata volt, ezért a kétszer vagy háromszor megjelenő argumentumszerkezet-változtató hatást már szisztematikusnak lehet tekinteni, azonban egy több igét lefedő vizsgálatnál pontosabb definícióra lenne szükség a szisztematikusság tekintetében. Ahhoz, hogy a szisztematikus változásokat is tanulmányozzuk, más igekötő-ige párokat kell választani, melyek várhatóan gyakrabban fordulnak elő a korpuszban kompozicionálisan (erre vonatkozóan lásd Gyulai 2018, ahol a *ki* és az *át* igekötős szerkezeteket vizsgáltam). Az adatok alapján nem a perfektívizálás az igekötő legjellemzőbb hatása, azonban ez kompozicionális hatásként említhető. A szisztematikus változások szemléltetéséhez beemelek egy táblázatot (9. táblázat), amely a *ki* igekötő hasonló vizsgálatának összegző táblázata (Gyulai 2018).

	A C C ↑	R Ó L 0	B A ↓	V A L 1	V A L ↑	L O C Y 0	L O C Y 1	R A ↑	T O ↑	B A N +	B A N +	I N F 1	H K M 1
<i>el₁(lát)</i>	•			•									
<i>el₂(halad)</i>							•						
<i>el₃(készít)</i>		•											
<i>el₄(készül)</i>		•						•					
<i>el₅(jár)</i>					•					•			
<i>el₆(jár)</i>			•					•	•		•		
<i>el₇(felejt)</i>						•						•	
<i>el₈(felejt)</i>						•							•

8. táblázat. Az *el* igekötőnek az igei argumentumszerkezetre gyakorolt különböző hatásainak összesítő táblázata

A 9. táblázat magyarázataként megjegyzendő, hogy a *ki₂*, *ki₃* és *ki₇* igekötős argumentumszerkezeti variánsoknál szisztematikus argumentumszerkezeti változás figyelhető meg, tehát kompozicionális igekötős igékről van szó: az igekötő letiltja a tárgyat, illetve bevezeti a *-bÓl* ragos bővítményt. A *ki₄* argumentumszerkezeti variáns esetén pedig nem kompozicionális összetételről beszélhetünk, hiszen az igekötő egyedi hatását figyelhetjük meg.

	csökkent			növel			
	ACC	BA	RA	BÓL	BAN	ON	MODE
<i>ki₁(lép)</i>		0				1	
<i>ki₂(lép)</i>		0		+			
<i>ki₃(tör)</i>	0			1			
<i>ki₄(tör)</i>	0				↑		
<i>ki₅(néz)</i>	0						1
<i>ki₆(néz)</i>	0					1	0
<i>ki₇(ugrik)</i>	0		0	↑			

9. táblázat. A *ki* igekötőnek az igei argumentumszerkezetre gyakorolt különböző hatásainak összesítő táblázata

Azok a változások, melyeket az *el* igekötő megjelenése kapcsán megfigyelhettünk, nem tekinthetők kompozicionálisnak. Ezek az igekötős igék valóban jelentésükben sem kompozicionálisak, tehát az alkalmazott módszer alkalmasnak látszik arra, hogy csupán szintaktikai tulajdonságokat figyelembe véve kiszűrjük a nem kompozicionális olvasattal rendelkező igekötős igéket.

A vizsgálat hiányossága a korpusz kis mérete, hiszen gyakran nem volt elég elemszámú előfordulás ahhoz, hogy értékelhető eredményeket lehessen kapni. Jövőbeli célkitűzéseim között szerepel a korpusz bővítése mind a vizsgált igék, mind a vizsgált igekötők szempontjából, illetve a munkafolyamatok automatizálása a korpusz tényleges bővítésére is lehetőséget adna.

Köszönetnyilvánítás

A kutatást az EFOP-3.6.1-16-2016-00008 azonosítójú, EU társfinanszírozású projekt támogatta.

Irodalom

- Gyulai L. 2018 *Az igekötők argumentumszerkezet-változtató képességének korpuszalapú vizsgálata*. TDK-dolgozat. Szegedi Tudományegyetem: Szeged.
- Kálmán L. 2016. Bővítésminták mint konstrukciók. In: Kas B. (szerk.) „Szavad ne feledd!” *Tanulmányok Bánréti Zoltán tiszteletére*. Budapest: MTA Nyelvtudományi Intézet. 61–72.
- Komlósy A. 1992. Régenek és vonzatok. In: Kiefer Ferenc (szerk.) *Strukturális magyar nyelvtan I. Mondattan*. Budapest: Akadémiai Kiadó. 305–525.
- Müller, Ch. – Strube, M. 2006. Multi-Level Annotation of Linguistic Data with MMAX2. In: Braun, S. – Kohn, K. – Mukherjee, J. (szerk.) *Corpus Technology and Language Pedagogy. New Resources, New Tools, New Methods*. Frankfurt: Peter Lang. 197–214.
- Partee, B. H. 2004. Compositionality. In: Partee, B. H. *Compositionality in Formal Semantics. Selected Papers by Barbara H. Partee*. Blackwell: London. 153–181.
- Sass B. 2009. „Mazsola” – eszköz a magyar igék bővítémszerkezetének vizsgálatára. In: Váradí T. (szerk.) *Válogatás az I. Alkalmazott Nyelvészeti Doktorandusz Konferencia előadásaiból*. Budapest: MTA Nyelvtudományi Intézet. 117–129.
- Szécsényi T. 2019. Argumentumszerkezet-variánsok korpusz alapú meghatározása. In: Berend G. – Gosztolya G. – Vincze V. (szerk.) *XV. Magyar Számítógépes Nyelvészeti Konferencia* (Szeged, 2019. Január 24–25.) Szeged: SZTE TTIK Informatikai Intézet. 315–329.

- Szécsényi T. – Gyulai L. 2019. *Aszemantikus kompozicionalitás és az igekötő-ige szerkezetek kompozicionalitása. Újdonságok a szemantikai és pragmatikai kutatásokban*. Szeged, 2019. április 12.
- Vincze V. – Szauter D. – Almási A. – Móra Gy. – Alexin Z. – Csirik J. 2010. Hungarian Dependency Treebank. In: *Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC'10)*. Valletta, Málta.

A karakterábrázolás alapfrekvencia-jellemzői

Huszár Anna

ELTE BTK Nyelvtudományi Doktori Iskola, MTA Nyelvtudományi Intézet
huszar.anna@nytud.hu

Kivonat: A meseolvasás egy speciális területe a beszédnek. A mesét olvasók beszédük változatosságával ragadják meg, illetve tartják fenn a hallgatóság figyelmét, az egyes karakterek jellegzetességeit a hangjukkal jelenítik meg. Ebből adódóan a meseolvasás szupraszegmentális szerkezete eltérhet más beszédműfajokétól (Fackrell et al. 2000; Theune et al. 2006). A mesén belül az egyes karakterek, illetve bizonyos karaktertípusok megjelenítésének módja is különbözik (Doukhan et al. 2011). A jelen kutatásban sztereotip tulajdonságokkal rendelkező karakterek ábrázolását vizsgáltam. Elemeztem az átlagos alapfrekvenciát, a hangterjedelmet és a zöngeminőséget. Az eredmények szerint szignifikánsan eltért az átlagos alapfrekvencia és hangterjedelem a karakterek mentén. A irreguláris részek aránya is tendenciaszerűen eltérést mutatott a karakterek mentén.

1 Bevezetés

A beszédprodukción számos tényező befolyásolhatja, többek között a beszélő személye, a hallgatóság, a beszédtema, a beszédhelyzet és a közlés célja (Gósy 2004). A beszélő életkora, neme, szociológiai háttere, személyisége, aktuális érzelmi és fizikai állapota mind hatással lehetnek a beszédprodukción (Balázs–Bóna 2016; Bóna 2016; Gocsál 1998; Gocsál–Huszár 2003; Gósy 2004; Gósy–Krepsz 2015; Gósy–Krepsz 2018; Gyarmathy 2007; Németh 2013). A beszélő személyén kívül a hallgatóság is meghatározó lehet, a résztvevők száma, illetve az, hogy ismerik-e egymást, illetve milyen viszonyban állnak egymással (pl. barátok, anya-gyermek, tanár-diák stb.). A beszéd típus is meghatározó tényező, más a spontán beszéd, és más a felolvasás prozódiai szerkezete (Bata–Grácsi 2009; Bóna 2011; Jacewicz et al. 2010; Markó 2009; Váradi 2010). Prozódiai vagy szupraszegmentális szerkezet alatt értem a beszéd azon nem szegmentális jellemzőit, melyek az idő, a frekvencia és az intenzitás változásaival írhatók le (Hardcastle–Laver 1997).

Az egyes beszédműfajok szintén eltérhetnek egymástól (pl. hírolvasás, politikai beszéd, reklámszöveg, regény és meseolvasás) több paraméter mentén is: beszédtempó, artikulációs tempó, szünetezés, beszédszakaszok (azaz a szünettől szünetig tartó egységek) hossza, alapfrekvencia (átlagos alapfrekvencia és hangterjedelem) és formánszerkezet (Abe–Mizuno 1994; Eskenazi 1993; Hirschberg 2000; Roekhaut et al. 2010).

Huszár Anna: A karakterábrázolás alapfrekvencia-jellemzői. In Váradi Tamás (sorozatszerkesztő), Ludányi Zsófia, Grácsi Tekla Etelka (szerkesztő): *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2019. XIII. Alkalmazott Nyelvészeti Doktoranduszkonferencia*. Budapest: MTA Nyelvtudományi Intézet. 2019. 61–73. DOI: 10.18135/Alknyelvdok.2019.13.5

A meseolvasás egy speciális területe a beszédnek, prozódiai szerkezete eltér más beszédműfajokétól (Fackrell et al. 2000; Theune et al. 2006). A mesét olvasók beszédük változatosságával tartják fenn a hallgatóság figyelmét, az egyes karakterek jellegzetességeit a hangjukkal jelenítik meg.

Többen vizsgálták a mese prozódiai szerkezetét gépi felolvasás létrehozásának céljából. Ahhoz, hogy minél tökéletesebb gépi felolvasóalgoritmust hozzanak létre, szükséges minél pontosabban ismerni a mese szupraszegmentális szerkezetét, több akusztikai elemzés született a témában (Fackrell 2000; Jokisch et al. 2005; Roekhaut et al. 2010; Theune et al. 2006; Verma et al. 2015). A meseolvasás temporális szerkezete jelentősen eltér más beszédműfajokétól, lassabb beszéd- és artikulációs tempó, hosszabb szünetek, magasabb szünetarány és hosszabb beszédszakaszok jellemzik (Fackrell et al. 2000; Jokisch et al. 2005; Roekhaut et al. 2010; Theune et al. 2006; Verma et al. 2015). A meseolvasás alapfrekvencia-jellemzői is különböznek a többi beszédműfajokétól, magasabb átlagos alapfrekvencia és nagyobb hangterjedelem, nagyobb fokú hangsúlyozás és az intenzitás nagyobb fokú variabilitása jellemzi (Jokisch et al. 2005; Roekhaut et al. 2010; Theune et al. 2006; Verma et al. 2015).

Magyar kutatások is vizsgálták a mese prozódiai szerkezetét (Imre 2005; Olaszky 2005). Egy kutatásban Olaszky (2005) a hírolvasással, novella és reklámok felolvasásával vetette össze a mesemondás prozódiját. Az eredmények szerint a mesemondás időszerkezete a magyar nyelvben is eltér más beszédműfajokétól: a mesemondásra lassabb beszéd- és artikulációs tempó jellemző a hírolvasáshoz képest. A mesemondásra általában rövidebb, jól tagolható mondatok, több mondatbeli szünet jellemző, melyek időtartama hosszabb, mint más beszédműfajokban. A mesemondás frekvenciaszerkezete szintén eltér más beszédműfajokétól, általában magasabb és variábilisabb alapfrekvencia és nagyobb hangterjedelem jellemzi.

A mesemondás egyik jellemzője, hogy a mesemondó nemcsak narrátor, hanem sok esetben ő jeleníti meg beszédével az összes karaktert. A mesemondó prozódiai eszközöket használ annak érdekében, hogy az egyes karaktereket elkülönítse egymástól, illetve kifejezően ábrázolja azokat. Adell és munkatársai (2005) többek között a meseolvasás szupraszegmentális szerkezetét vetették össze más beszédműfajokkal (érzelmes beszéd és parlamenti beszéd). Többek között mérték alapfrekvencia-jellemzőket: az alapfrekvencia átlagát, szórását, minimumát, maximumát, a hangterjedelmet és a dallammenetet elemezték. Az intenzitásnak szintén nézték az átlagát, szórását, minimumát és maximumát is. Az artikulációs tempót és a szünetek időtartamát is elemezték. A mese különböző részeinek szupraszegmentális szerkezetét is vizsgálták, és megállapították, hogy a narratív részek, a leírások és az egyes karakterek megszólalásai prozódiailag elkülönülnek egymástól.

Egy másik kutatás (Montaño et al. 2013) szintén a meseolvasás narratív, leíró és dialogikus részeinek prozódiai szerkezetét vizsgálta. A narrátorhoz tartozó beszédegységeket különböző részekre bontva elemezték: semleges, leíró, karakter megszólalása utáni és többféle érzelmi töltetű rész. Az átlagos alapfrekvencia és annak szórásának, az átlagos intenzitás és az artikulációs tempó alakulását vizsgálták. Az említett kategóriák között főleg az alapfrekvenciában (átlag és szórás tekintetében is) és az intenzitás-ban találtak eltéréseket, az artikulációs tempóban a kategóriák többsége között nem volt jelentős különbség. A karakterek megszólalása utáni részekre alacsonyabb alapfrekvencia és intenzitás volt jellemző. A leíró részekre magasabb alapfrekvencia volt jellemző (átlag és szórás tekintetében is), intenzitásában azonban nem különböztek jelentősen a semleges stílusú részeketől. A különböző érzelmi töltettel rendelkező részek prozódiai szerkezetét is elemezték, és lassabb artikulációs tempót mérték a semleges részekhez

viszonyítva. Továbbá magasabb átlagos intenzitást és alapfrekvenciát mértek (illetve az alapfrekvencia nagyobb szórását) a következő érzelmek esetében: düh, öröm, meglepetés és félelem. A szomorúságra alacsonyabb átlagos intenzitás és alapfrekvencia, illetve az alapfrekvencia kisebb szórása volt jellemző.

Doukhan és munkatársai (2011) szintén a meseolvasás szupraszegmentális sajátosságait vizsgálták. Tizenkét mese felolvasásában elemezték az artikulációs tempót, a szünetek százalékos időtartamarányát, az átlagos alapfrekvenciát, az átlagos intenzitást, a szótagok közti átlagos alapfrekvencia- és intenzitáskülönbségeket és a hangterjedelmet. A narrátor szakaszait a következő részekre bontva vizsgálták: cím, bevezető szöveg, a kiváltó esemény bemutatása, jelenetek, refrén, utószó. Eredményeik szerint a cím prozódiai sajátosságai jelentősen elkülönültek a narrátor többi szakaszának sajátosságaitól: a címnél mérték átlagosan a legmagasabb intenzitást, a legnagyobb szótagok közti alapfrekvencia-különbséget, a legnagyobb hangterjedelmet, és a leglassabb artikulációs tempót (a szünetarány is a cím esetében volt a legalacsonyabb, azonban a címek általában lényegesen rövid szövegek). A bevezetésben általában nagy volt a szótagok közötti alapfrekvencia-különbség, ami a hallgatóság figyelmének megragadásával is magyarázható. A refrénekre alacsony szünetidőtartam-arány volt jellemző, illetve alacsony volt a szótagok közti alapfrekvencia-különbség is a többi narrátori részhez képest. Az utószóra szintén alacsony szótagok közti alapfrekvencia-különbséget mértek, alacsony alapfrekvenciát és intenzitást, kis hangterjedelmet és intenzitásváltozást, illetve lassú artikulációs tempót. A narrátor szakaszain kívül a karakterek megjelenítését is vizsgálták. A megjelenített karaktert neme és életkora (gyerekszereplő, felnőtt vagy idős) alapján csoportosították. A meseolvasó beszédének alapfrekvenciája és intenzitása magasabb volt, amikor karaktereket jelenített meg, mint amikor a narrátor részeit (ez alól a nagy medve kivétel, ennél a karakternél mérték a legalacsonyabb alapfrekvenciát). A gyerekszereplők esetében általában magasabb és egyenletesebb alapfrekvenciát figyeltek meg; a fiúknál magasabb, a lányoknál alacsonyabb intenzitást. A karakterek az alapján is elkülöníthetőek voltak prozódiai szempontból, hogy fő- vagy mellékszereplők voltak. A gyerek főszereplőkre általában jellemző volt a magasabb átlagos alapfrekvencia és a kisebb szótagok közti alapfrekvencia-különbség. A felnőtt mellékszereplők megjelenítése a narrátoréhoz volt hasonló prozódiaileg. Az idősebb szereplők esetében nagyobb volt a szótagok közti alapfrekvencia-különbség. Nem elhanyagolható azonban a megjelenítendő karakterek életkora sem, a szupraszegmentális szervezetre a karakterek életkora és a szerep típusa együttesen lehetett hatással.

A fenti kutatásokból látható, hogy a meseolvasás prozódiai szerkezete több tényező mentén is különbözik a többi beszédműfajétól. A mese egyes részei (pl. narrátor és a különféle karakterek megjelenítése) több prozódiai jegyben is eltérhetnek egymástól. A fenti kutatások főleg adott mesék felolvasásának prozódiai elemzését tűzték ki célul. A narrátor szerepén kívül az egyes karakterek megjelenítését is vizsgálták, a karakterek megjelenése egy mesében azonban nem egyenletes, vannak karakterek (pl. főszereplők), akik gyakran szóhoz jutnak a mese során, vannak azonban olyan karakterek is (pl. mellékszereplők), akik csak ritkán kapnak szót, ez nehezíti az összehasonlíthatóságukat. Ezen túl az egyes karakterek szövege eltérő, tehát ez is nehezíti a megjelenítésük vizsgálatát. A prozódia befolyásolhatja a mondatok hossza, szótagszáma, összetettsége, hogy tartalmaznak-e alá- vagy mellérendeléseket, így az egyes karakterek közti különbségek egy része adódhat az eltérő mondatszerkezetből is. Az összehasonlítás érdekében célszerű ugyanazon szövegen vizsgálni az eltéréseket.

A jelen kutatás célja sztereotip tulajdonságokkal rendelkező karakterek ábrázolásának vizsgálata adott szövegen. A karakterek megválasztásánál törekedtem arra, hogy

minél több tulajdonság mentén eltérjenek egymástól, a következő karakterek megjelenítését vizsgáltam: IZGÁGA HÖRCSÖG, LUSTA LAJHÁR, KECSSES ŐZ és RAVASZ RÓKA. Az elemzés során arra a kérdésre keresem a választ, hogy mennyiben térnek el az egyes alapfrekvencia-jellemzők a karakterek mentén, illetve a semleges felolvasáshoz viszonyítva. Hipotéziseim szerint a karakterek mentén eltérés mutatkozik (i) az átlagos alapfrekvenciában, (ii) a hangterjedelemben és (iii) az irregularitás megjelenésében is.

Az IZGÁGA HÖRCSÖGHÖZ társítható az izgatottság és a vidámság. Az izgatottságra az alapfrekvencia nagyobb mértékű variabilitása (Trouvain–Barry 2000), a vidámságra pedig magasabb és variábilisabb alapfrekvencia jellemző (Scherer et al. 1991). Tehát az IZGATOTT HÖRCSÖG esetében magasabb alapfrekvenciát és nagyobb hangterjedelmet várok a kontroll kondícióhoz viszonyítva.

A LUSTA LAJHÁRHOZ társítható a szomorúság, a csalódottság vagy akár az unottság. A szomorúságot jellemezheti alacsonyabb és kevésbé variábilis alapfrekvencia, kisebb hangterjedelem és nagyobb fokú irregularitás (Burkhardt–Sendlmeier 2000; Paeschke 2004; Scherer et al. 1991). Az unottságra jellemző lehet alacsonyabb alapfrekvencia, nagyobb hangterjedelem és nagyobb fokú irregularitás (Burkhardt–Sendlmeier 2000; Paeschke 2004; Scherer et al. 1991). A csalódottságra kisebb hangterjedelem jellemző (Maekawa 1998). Tehát a LUSTA LAJHÁR esetében alacsonyabb alapfrekvenciát és nagyobb fokú irregularitást várok a kontroll kondícióhoz viszonyítva.

A KECSSES ŐZHÖZ társítható a nőiesség. A magasabb alapfrekvenciájú hangot nőiesebbnek ítélik meg (Feinberg et al. 2008). A KECSSES ŐZ esetében tehát magasabb alapfrekvenciát várok a kontrollkondícióhoz képest.

A RAVASZ RÓKÁHOZ társítható az irónia. Az iróniára jellemző a magasabb és variábilisabb alapfrekvencia (Bryant–Fox 2005). A RAVASZ RÓKA esetében tehát magasabb alapfrekvenciát és nagyobb hangterjedelmet várok.

2 Módszertan

Adatközlőim középiskolás dráma szakos diákok voltak (nyolc tizenegyedik osztályos lány). Feladatuk egy adott szöveg felolvasása volt, melyet előre megkaptak. A felolvasandó szöveg a kísérletre készült, öt kisebb egységből áll. Összesen 40 mondatot tartalmaz, melyek között vannak rövidebb (kb. 5-6 szótagos) és hosszabb (kb. 9-11 szótagos) mondatok is. A szöveg felolvasása körülbelül egy percet vett igénybe. A szöveget először semleges stílusban kellett felolvasniuk (kontroll kondíció), majd a négy különböző sztereotíp tulajdonsággal rendelkező állatfigurákat (RAVASZ RÓKA, LUSTA LAJHÁR, KECSSES ŐZ és IZGÁGA HÖRCSÖG) megjelenítve. Az adatközlők rajz-illusztrációkat (1. ábra) kaptak segítségül a karakterek megformálásához, arra azonban nem kaptak semmilyen instrukciót, hogy az adott karaktereket hogyan jelenítsék meg.

A felvételeket csendesített szobában, fejmikrofonnal rögzítettem Audacity (Audacity Team 2018) programmal. A hanganyagot Praat szoftverrel (Boersma–Weenink 2018) szakasz- és mondatszinten annotáltam. Az esetleges megakadásokat az elemzés során nem vettem figyelembe. Szegmentáltam az irregularis részeket a rezgéskép, a hangszínek és az auditív információk alapján (Dilley et al. 1996; Markó 2013). Az alapfrekvenciát a mondatok nem irregularis részein 5 ms-onként, 75–600 Hz között, egy Praat szkript segítségével mértük. Az átlagos alapfrekvenciát mondatokra számoltam, Hz-ben kifejezve. A hangterjedelmet félhangban kifejezve adtam meg mondatokra az alábbi képletet használva:

$$(1) 12 * \log_2 \frac{f_{0max}}{f_{0min}}$$



1. ábra. A megjelenítendő karakterek rajz-illusztrációi (RAVASZ RÓKA, LUSTA LAJHÁR, KECSES ŐZ ÉS IZGÁGA HÖRCSÖG)

Ahol f_{0max} az adott mondatban mért maximális alapfrekvencia, f_{0min} pedig az adott mondatban mért minimális alapfrekvencia.

Az irreguláris részek arányát százalékosan adtam meg. Egyrészt kiszámoltam az irreguláris részek időtartamának arányát, azaz megadtam, hogy a mondatok teljes idejének hány százaléka az irreguláris szakaszok összeitartama (az időtartamértékeket szintén Praat szkripttel nyertem ki). Másrészt az irreguláris szótagok arányát is megadtam, tehát azt, hogy a teljes szöveg szótagszámának hány százalékát teszik ki az irreguláris szótagok.

Az adatokon statisztikai elemzéseket végeztem, kevert modelleket állítottam fel (Linear Mixed Model, LMM), az SPSS 20.0 szoftverben számoltam. A kondíció mint független változó hatását vizsgáltam a következő függő változókra: átlagos alapfrekvencia és hangterjedelem. A beszélőt és a mondatot mindkét esetben random faktorként kezeltem. A csoportok közötti összehasonlításhoz a Tukey-féle post hoc tesztet használtam.

A kondíciónak az irregularitásra gyakorolt hatását a Friedman-féle összetartozó min-tás varianciaanalízissel vizsgáltam, melyben a függő változó az irreguláris részek időtartam-, illetve szótagaránya volt, a független változó pedig a kondíció. A csoportok közötti összehasonlítást a Wilcoxon Signed Rank tesztel végeztem.

3 Eredmények

3.1 Átlagos alapfrekvencia

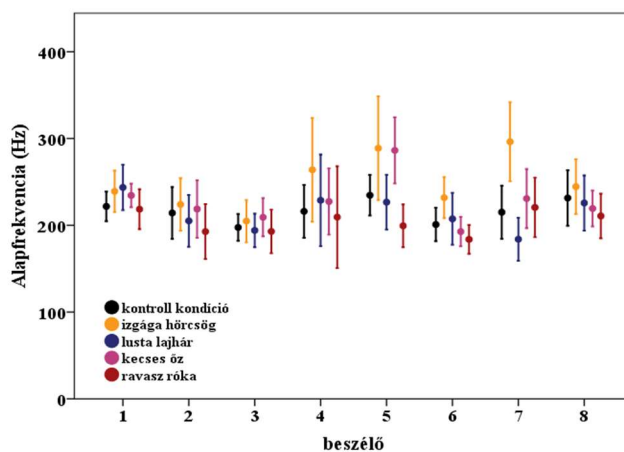
Az alapfrekvencia értékeit a 2. ábra mutatja Hz-ben kifejezve, az ábrán az átlag- és szórásértékek láthatók. A grafikon az egyes kondíciók (kontroll kondíció és a négy karakter) közti eltéréseket szemlélteti, beszélőnkénti bontásban.

Az átlagos alapfrekvencia minden beszélő esetén magasabb az IZGÁGA HÖRCSÖG megvalósításakor, mint a kontroll kondícióban. Egy adatközlő kivételével mindenki átlagosan alacsonyabb alapfrekvencián valósította meg a RAVASZRÓKA karakterét a kontrollhoz viszonyítva. A KECSES ŐZ esetében a beszélők többségénél (nyolcból hat beszélő) átlagosan magasabb alapfrekvencia volt jellemző a kontroll kondícióhoz

viszonyítva. A LUSTA LAJHÁR ábrázolása volt a legváltozatosabb, az adatközlők közül öten alacsonyabb, hárman magasabb átlagos alapfrekvencián valósították meg ezt a karaktert.

A grafiknról továbbá leolvasható, hogy az adatközlők eltérő mértékben változtattak az alapfrekvencián. Voltak beszélők, akiknél az átlagos alapfrekvencia karakterenként csak kis mértékben tért el (pl. első és harmadik beszélő), azonban náluk a szórás is kisebb volt, tehát a karakterek közti változtatások lehetnek következetesek. Néhány adatközlő azonban jelentősen nagyobb mértékben változtatott az alapfrekvencián az egyes karakterek esetében (pl. negyedik, ötödik és hetedik beszélő), az ő esetükben a szórás is nagyobb volt.

Az 1. táblázat az átlagos alapfrekvenciát és annak szórását mutatja kondíciókra bontva.



2. ábra. Az alapfrekvencia alakulása a kondíció függvényében, beszélőnkénti bontásban (átlag- és szórásértékek)

kondíció	f_0 -átlag (Hz)	szórás (Hz)
kontroll	216,4	28,1
izgága hörcsög	249,6	49,4
lusta lajhár	213,5	37,0
kecses őz	227,5	38,3
ravasz róka	203,3	34,3

1. táblázat. Az átlagos alapfrekvencia és szórása kondícióként

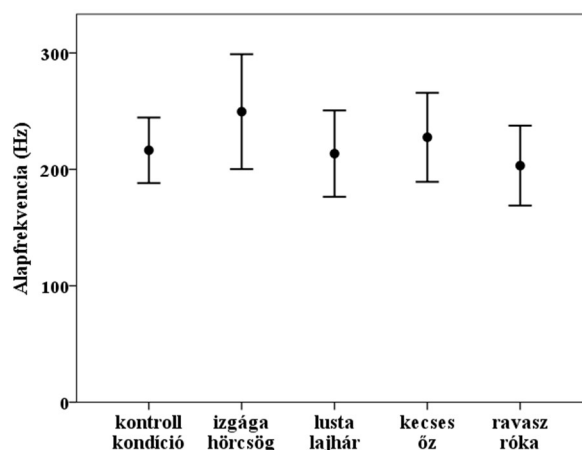
A táblázatról leolvasható, hogy az alapfrekvencia szórása a kontroll kondícióban volt a legkisebb és az IZGÁGA HÖRCSÖG esetében a legnagyobb.

A 3. ábra az alapfrekvencia alakulását szemlélteti a kondíció függvényében. (Az ábrán átlag- és szórásértékek láthatóak.)

A 3. ábrán látható, hogy az IZGÁGA HÖRCSÖG és a KECSES ŐZ karakterét általában magasabb, a LUSTA LAJHÁR és a RAVASZ RÓKA karakterét általában alacsonyabb alapfrekvencián valósították meg a beszélők. A kontrolltól való legnagyobb mértékű eltérés az IZGÁGA HÖRCSÖG, míg a legkisebb mértékű eltérés a LUSTA LAJHÁR esetében volt

tapasztalható. Általában megállapítható, hogy az adatközlők nagyobb mértékben változtatták az alapfrekvenciájukat felfelé, mint lefelé a kontrollhoz képest.

A statisztikai modell (LMM) alátámasztotta, hogy a kondíció hatással van az alapfrekvenciára ($\chi^2(1) = 324,32; p < 0,05$). A Tukey-féle post hoc teszt alapján elmondható, hogy az alapfrekvencia a LUSTA LAJHÁR kivételével mindegyik karakter esetében szignifikánsan eltért a kontrolltól (IZGÁGA HÖRCSÖG: $Z = 12,259; p < 0,001$; LUSTA LAJHÁR: $Z = -0,894; p = 0,815$; KECSSES ÖZ: $Z = 3,291; p = 0,004$ és RAVASZ RÓKA: $Z = -3,298; p = 0,004$).



3. ábra. Az alapfrekvencia alakulása a kondíció függvényében (átlag- és szórásértékek)

3.2 Hangterjedelem

A hangterjedelem beszélőnkénti alakulását a karakterek mentén a 4. ábra mutatja félhangban kifejezve. Az ábrán az átlag- és szórásértékek láthatók.

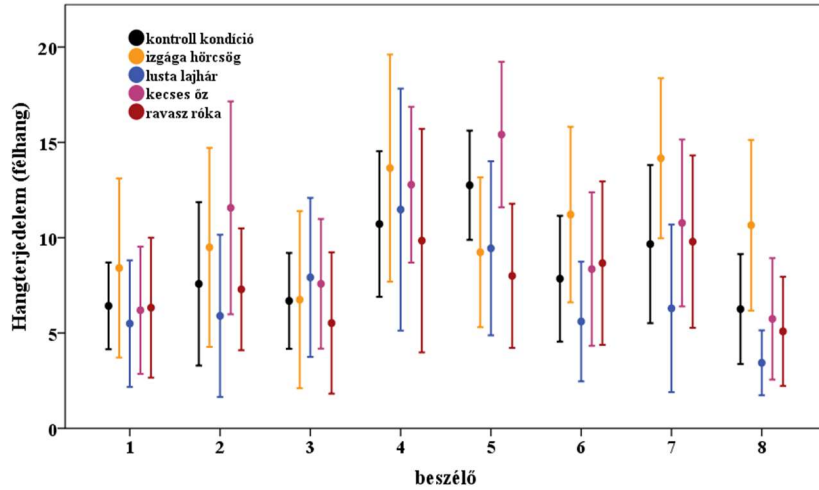
Egy adatközlő kivételével mindenki átlagosan nagyobb hangterjedelemben valósította meg az IZGÁGA HÖRCSÖGöt a kontroll kondícióhoz képest. Két beszélő kivételével a LUSTA LAJHÁRT és a RAVASZ RÓKÁT átlagosan kisebb, a KECSSES ÖZet nagyobb hangterjedelemben valósították meg a kontrollhoz képest.

Beszélőnként eltérő mértékű változtatások figyelhetők meg, vannak azonban hasonló stratégiák: az első, a hatodik, a hetedik és a nyolcadik beszélőnél egyaránt megállapítható, hogy IZGÁGA HÖRCSÖG hangterjedelme jelentősen nagyobb volt a többi karakterhez viszonyítva, a LUSTA LAJHÁRÉ a legalacsonyabb, míg a KECSSES ÖZ és RAVASZ RÓKA nem sokban tért el a kontroll kondíciótól.

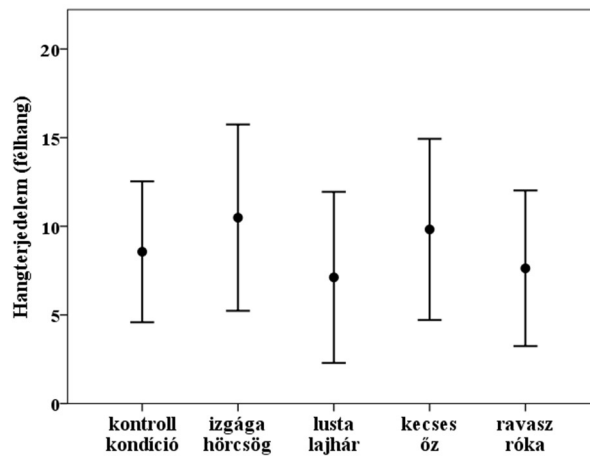
Az 5. ábra a hangterjedelmet a kondíció függvényében mutatja, beszélőnkénti bontás nélkül.

Az 5. ábra alapján megfigyelhető, hogy azon karaktereket, melyeket a kontroll kondícióhoz képest átlagosan magasabb alapfrekvencián valósítottak meg, általában nagyobb hangterjedelemmel is, illetve, melyeket átlagosan alacsonyabb alapfrekvencián valósítottak meg, azok esetében a hangterjedelem is kisebb volt.

A statisztikai modell (LMM) megerősítette, hogy a kondíció hatással van a hangterjedelemre ($\chi^2(4) = 157,66; p < 0,05$). A Tukey-féle post hoc teszt alapján elmondható, hogy a hangterjedelem mindegyik karakter esetében szignifikánsan eltért a kontrolltól (IZGÁGA HÖRCSÖG: $Z = 5,976; p < 0,001$; LUSTA LAJHÁR: $Z = -4,407; p < 0,001$; KECSSES ÖZ: $Z = 3,711; p < 0,001$ és RAVASZ RÓKA: $Z = -2,556; p = 0,0459$).



4. ábra. A hangterjedelem alakulása a kondíció függvényében, beszélőnkénti bontásban (átlag- és szórásértékek)



5. ábra. A hangterjedelem alakulása a kondíció függvényében (átlag- és szórásértékek)

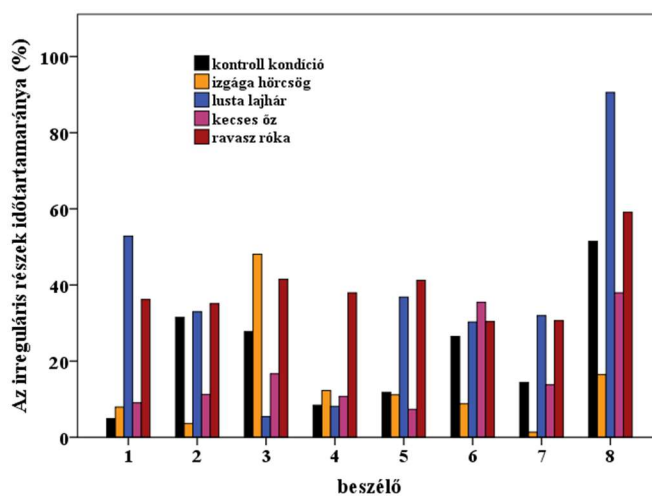
3.3 Zöngeminőség – irregularitás

A 6. ábra az irregularis részek időtartamarányát mutatja a kondíció függvényében, beszélőkre bontva, százalékban kifejezve.

Az irregularis részek időaránya minden beszélő esetében magasabb volt a RAVASZ RÓKA megvalósításakor, mint a kontroll kondícióban. Egy adatközlő kivételével a LUSTA LAJHÁR esetében is magasabb volt az irregularis részek aránya a kontrollhoz képest. A KECSSES ŐZ és az IZGÁGA HÖRCSÖG esetében a beszélők többségénél alacsonyabb volt az irregularis részek időaránya. Megfigyelhető azonban, hogy nagyok voltak az egyéni eltérések, melyek már abból is adódhattak, hogy a kontroll kondícióban is jelentősen eltért az irregularitás aránya az egyes beszélőknél. Az első, a negyedik és az ötödik adatközlőre kevésbé, míg a második, a harmadik, a hatodik és a nyolcadik adatközlőre nagyobb mértékben volt jellemző az irregularitás.

A 7. ábra az irregularis részek időtartamarányát mutatja a kondíció függvényében, beszélőnkénti bontás nélkül, százalékban kifejezve.

A 7. ábra alapján az mondható el, hogy a LUSTA LAJHÁR megvalósítása tért el a legnagyobb mértékben a többi kondíciótól. Az 6. ábrával összevetve (ahol látszik a beszélőnkénti bontás is) azonban látható, hogy a nyolcadik beszélő magas értékei jelentősen befolyásolták az arányokat. Megfigyelhető továbbá, hogy azon karakterek esetében, melyeket a kontrollhoz képest alacsonyabb alapfrekvencián valósítottak meg, általában nagyobb volt az irregularis részek aránya is.



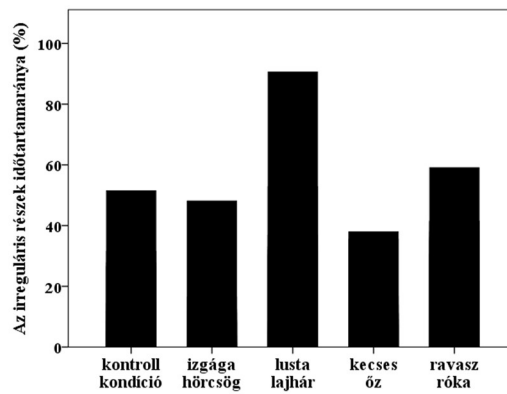
6. ábra. Az irregularis részek időtartamarányának alakulása a kondíció függvényében, beszélőnkénti bontásban

A statisztikai modell alátámasztotta, hogy a kondíció hatással van az irregularis részek időtartamarányára ($\chi^2(4) = 10,9; p = 0,028$). A Wilcoxon Signed Rank teszt alapján elmondható, hogy csak egy karakter, a RAVASZ RÓKA esetében tért el szignifikánsan az irregularis részek időaránya a kontrolltól (IZGÁGA HÖRCSÖG: $Z = -0,98; p = 0,327$; LUSTA LAJHÁR: $Z = -1,68; p = 0,093$; KECSSES ŐZ: $Z = -1,12; p = 0,263$ és RAVASZ RÓKA: $Z = -2,521; p = 0,012$).

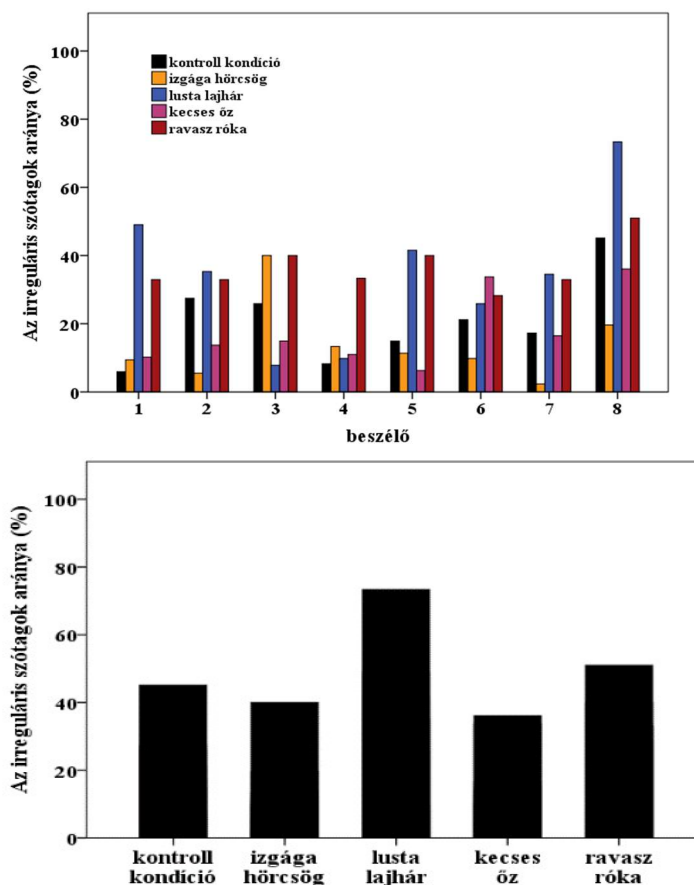
Az irreguláris szótagok arányát is elemeztem, melyek alakulását a 8. ábra szemlélteti. Felül az irreguláris szótagok aránya a kondíció függvényében, alul ugyanez beszélőkre bontva látható. Mindkét grafikon az irreguláris szótagok százalékban kifejezett értékét mutatja az összes szótagszámhoz viszonyítva.

A 8. ábra grafikonjait összevetve a 6. és a 7. ábrával megállapítható, hogy nem volt jelentős különbség az időtartam- és a szótagok aránya között. A tendenciák hasonlóak voltak a két különböző ábrázolási módban.

A statisztikai modell alátámasztotta, hogy az irreguláris szótagok arányára is hatással van a kondíció ($\chi^2(4) = 11,899$; $p = 0,018$). A Wilcoxon Signed Rank teszt alapján pedig elmondható, hogy csak egy karakter, a RAVASZ RÓKA esetében tért el szignifikánsan az irreguláris szótagok aránya a kontrolltól (IZGÁGA HÖRCSÖG: $Z = -1,12$; $p = 0,263$; LUSTA LAJHÁR: $Z = -1,82$; $p = 0,069$; KECSÉS ÖZ: $Z = -0,84$; $p = 0,401$ és RAVASZ RÓKA: $Z = -2,521$; $p = 0,012$).



7. ábra. Az irreguláris részek időtartamaránya a kondíció függvényében



8. ábra. Az irreguláris szótágok arányának alakulása a kondíció és a beszélő függvényében

4 Következtetések

A jelen kutatásban a karakterábrázolás alapfrekvencia-jellemzőinek alakulását vizsgáltam. Arra a kérdésre kerestem a választ, hogy elkülönülnek-e a karakterek mentén a következő paraméterek: átlagos alapfrekvencia, hangterjedelem és irregularitás.

Első hipotézisem az volt, hogy az átlagos alapfrekvencia eltérést fog mutatni a karakterek mentén. Ezt a hipotézisemet alátámasztották az adatok, egy karakter kivételével (LUSTA LAJHÁR) valóban szignifikánsan eltértek az értékek a kontroll kondícióban mérttől. Az IZGÁGA HŐRCSÖG és a KECSÉS ŐZ esetében az elvárásokkal egyezően a kontrollnál magasabb alapfrekvenciát mértem. A RAVASZ RÓKA esetében a kontrollnál alacsonyabb átlagos alapfrekvencia volt megfigyelhető a beszélők többségénél, ez ellentétben van a megfogalmazott elvárással.

A második hipotézisem a hangterjedeleme vonatkozott. Ez a hipotézisem is megerősítést nyert, a kondíció valóban hatással volt a hangterjedeleme, minden karakter esetén szignifikánsan eltértek az értékek a kontroll kondícióban mérttől. Az IZGÁGA

HÖRCSÖG esetében a várttal megegyezően nagyobb hangterjedelmet mértem a kontrollhoz képest, míg a RAVASZ RÓKA esetében a várttal ellenkezően kisebbet. A KECSSES ŐZ esetében nagyobb, a LUSTA LAJHÁR esetében pedig kisebb hangterjedelmet mértem a kontrollhoz képest.

A harmadik hipotézisem az volt, hogy a kondíció hatással lesz az irregularitás mértékére, tehát az irregularis részek aránya eltérést fog mutatni a karakterek mentén. A hipotézis csak részben nyert megerősítést az eredmények által: csupán egy karakter esetében (RAVASZ RÓKA) tértek el szignifikánsan az értékek a kontrolltól. Tendenciaszerűen azonban a többi karakter is különbözött egymástól az irregularitás tekintetében: a RAVASZ RÓKA és a LUSTA LAJHÁR esetében (a várttal egyezően) a kontrollhoz képest gyakoribb, az IZGÁGA HÖRCSÖG és a KECSSES ŐZ esetében ritkább volt az irregularis szakaszok előfordulása. A kutatás eredményei hozzájárulnak a karakterábrázolás alapfrekvencia-jellemzőinek részletesebb megismeréséhez.

Köszönetnyilvánítás

A kutatást a Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal NKFIH - FK-128814 számú pályázata támogatta.

Irodalom

- Abe, M. – Mizuno, H. 1994. Speaking styles conversion by changing prosodic parameters and formant frequencies. *ICSLP 94*. Yokohama, Japan: 1455–1458.
- Adell, J. – Bonafonte, A. – Escudero, D. 2005. Analysis of prosodic features towards modelling of emotional and pragmatic attributes of speech. *Procesamiento del lenguaje natural*, no. 35: 277–283.
- Audacity Team 2018. Elérhető: https://www.audacityteam.org/?fbclid=IwAR1X0rCeE0RuaW3T1ONRRhc2ic3F2TyNb2t5hGy7PwO2jtGy18nBUbBz_g
- Balázs B. – Bóna J. 2016. Életkori sajátosságok a beszédképzésben és a beszédfeldolgozásban. In: Bóna J. (szerk.) *Fonetikai olvasókönyv*. Budapest: ELTE BTK Fonetikai Tanszék. 7–19.
- Bata S. – Grácz T. E. 2009. Hatással van-e a beszédpartner életkora a beszélő beszédének szupraszegmentális jellegzetességeire. In: Keszler B. – Tátrai Sz. (szerk.) *Diskurzus a grammatikában, grammatika a diskurzusban*. Budapest: Tinta kiadó. 74–83.
- Boersma, P. – Weenink, D. 2018. *Praat: doing phonetics by computer*. Elérhető: http://www.fon.hum.uva.nl/praat/download_win.html (letöltés ideje: 2018. 12. 01.)
- Bóna J. 2011. A különböző beszédstílusok az akusztikai-fonetikai és percepció vizsgálatok tükrében. *Alkalmazott Nyelvtudomány* 11: 39–48.
- Bóna J. 2016. *Női beszéd - férfi beszéd a fonetikai és a pszicholingvisztikai vizsgálatok tükrében*. Nyelvtudományi értekezések 166. Budapest: Akadémiai Kiadó.
- Bryant, G. A. – Fox, J. E. 2005. Tree Is there an ironic tone of voice? *Language and Speech* 48(3): 257–277.
- Burkhardt, F. – Sendlmeier, W. F. 2000. Verification of acoustical correlates of emotional speech using formant-synthesis. In: *Proceedings ISCA Workshop on Speech and Emotion*, 5–7. September 2000, Newcastle, Northern Ireland: 151–156.
- Dilley, L. – Shattuck-Hufnagel, S. – Ostendorf, M. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24(4): 423–444.
- Doukhan, D. – Rilliard, A. – Rosset, S. – Adda-Decker, M. – d’Alessandro, C. 2011. Prosodic analysis of a corpus of tales. *InterSpeech 2011*: 3129–3132.
- Eskenazi, M. 1993. Trends in Speaking Styles Research. *Proceedings Eurospeech*. Berlin: 501–509.

- Fackrell, J. – Vereecken, H. – Buhmann, J. – Martens, J-P. – Van Coile, B. 2000. Prosodic variation with text type. *Proceedings of 6th International Conference on Spoken Language Processing (ICSLP 2000)*, Vol.3: 231–234.
- Feinberg, D. R. – DeBruine, L. M. – Jones, B. C. – Perrett, D. I. 2008. The role of femininity and averageness of voice pitch in aesthetic judgments of women's voices. *Perception* 37: 615–623.
- Gocsál Á. 1998. Életkorbecslés a beszélő hangja alapján. *Beszédkutatás* 1998: 122–135.
- Gocsál Á. – Huszár Á. 2003. Csábító hangok. *Beszédkutatás* 2003: 9–18.
- Gósy M. 2004. *Fonetika, a beszéd tudománya*. Budapest: Osiris Kiadó.
- Gósy M. – Krepesz V. 2015. Magánhangzók temporális jellemzői az idő múlásának függvényében. *Beszédkutatás* 2015: 53–65.
- Gósy, M. – Krepesz, V. 2018. Phrase-final Lengthening of Phonemically Short and Long Vowels in Hungarian Spontaneous Speech across Ages. In: Gósy M. – Grácsi, T. E. (szerk.) *Challenges in analysis and processing of spontaneous speech*. Budapest: Research Institute for Linguistics, Hungarian Academy of Sciences (RIL HAS). 99–126.
- Gyarmathy D. 2007. Az alkohol hatása a spontán beszédprodukcóra. *Beszédkutatás* 2007: 108–121.
- Hardcastle, W. J. – Laver, J. (eds.) 1997. *The handbook of phonetic sciences*. Blackwell Publishers, Oxford.
- Hirschberg, J. 2000. A Corpus-based Approach to the Study of Speaking Style. In: Horne, M. (szerk.) *Prosody: Theory and Experiment*. The Netherlands, Kluwer. 335–350.
- Imre A. 2005. Különböző műfajú szövegek szuprasegmentális jellemzői. *Magyar Nyelvőr* 129(4): 510–520.
- Jacewicz, E. – Fox, R. A. – Wei, L. 2010. Between-speaker and within-speaker variation in speech tempo of American English. *The Journal of the Acoustical Society of America* 128: 839.
- Jokisch, O. – Kruschke, H. – Hoffmann, R. 2005. Prosodic reading style simulation for Text-to-Speech synthesis. In: Tao, J. – Tan, T. – Picard, R. W. (szerk.) *Affective Computing and Intelligent Interaction, ser. Lecture Notes in Computer Science*. Springer Berlin Heidelberg, vol. 3784: 426–432.
- Maekawa, K. 1998. Phonetic and phonological characteristics of paralinguistic information in spoken Japanese. *Proceedings International Conference Spoken Language Processing*, vol. 2: 635–638.
- Markó A. 2009. Stigmatizált hanglejtésforma a spontán beszédben. *Beszédkutatás* 2009: 88–106.
- Markó A. 2013. *Az irreguláris zöngé funkciói a magyar beszédben*. Budapest: ELTE Eötvös Kiadó.
- Montaño, R. – Alas, F. – Ferrer, J. 2013. Prosodic analysis of storytelling discourse modes and narrative situations oriented to text-to-speech synthesis. *8th ISCA Workshop on Speech Synthesis*, Barcelona, Spain, August 31: 191–196.
- Németh M. 2013. A tanári beszéd és személyiség. *Anyanyelv-pedagógia* 2013/1. Elérhető: <http://anyanyelv-pedagogia.hu/cikkek.php?id=434> (letöltés ideje: 2019. január 3.)
- Olaszy G. 2005. Prozódiái szerkezetek jellemzése a hírfelolvasásban, a mesemondásban, a novella és a reklámok felolvasásában. *Beszédkutatás* 2005: 21–50.
- Paeschke, A. 2004. Global trend of fundamental frequency in emotional speech. *Proceedings of Speech Prosody* 2004, Nara: 671–674.
- Roekhaut, S. – Goldman, J. – Simon, A. 2010. A Model for Varying Speaking Style in TTS systems. In: *5th International Conference on Speech Prosody*, Chicago: IL.
- Scherer, K. R. – Banse, R. – Wallbott, H. G. – Goldbeck, T. 1991. Vocal cues in emotion encoding and decoding. *Motivation and Emotion* 15: 123–148.
- Theune, M. – Meijs, K. – Heylen, D. – Ordelman, R. 2006. Generating expressive speech for storytelling applications. *Speech, and Language Processing* 14(4): 1137–1144.
- Trouvain, J. – Barry, W. J. 2000. The prosody of excitement in horse race commentaries. In: *Proceedings ISCA Workshop on Speech and Emotion*. 5–7. September, Newcastle, Northern Ireland: 86–91.
- Váradi V. 2010. A felolvasás és a spontán beszéd temporális sajátosságainak összehasonlítása. *Beszédkutatás* 2010: 100–109.
- Verma, R. – Sarkar, P. – Rao, K. 2015. Conversion of neutral speech to storytelling style speech. *Advances in Pattern Recognition (ICAPR)*: 1–6.

Kérdő funkciójú megnyilatkozások enyhe és középsúlyos értelmi fogyatékosok felolvasásában

Jankovics Julianna

ELTE BTK Nyelvtudományi Doktori Iskola
jankovicsjuli@gmail.com

Kivonat: Mind a nemzetközi, mind a magyar irodalomban több alkalommal vizsgálták már a kérdő funkciójú megnyilatkozásokat mentálisan ép személyek beszédében, ám az értelmi fogyatékosokkal élők esetében ez még feltérképezetlen terület. A jelen tanulmányban felnőtt enyhe és középsúlyos értelmi fogyatékosokkal élő személyek többször ismételt felolvasásában vizsgálom az eldöntendő és a kiegészítendő kérdéseket, és vetem össze a nemben és korban illesztett mentálisan ép személyek kérdő funkciójú megnyilatkozásaival. A kutatás fő célja a két adatközlői csoport közötti egyezések és különbségek feltérképezése a felolvasásokban lévő kérdések alapján. A tanulmányban vizsgálom az adatközlők átlagos alapfrekvenciáját, a kérdések dallammenetét, valamint a zöngéminőségét. Az eredmények szerint az értelmi fogyatékosokkal élőknél átlagos alapfrekvenciája tendenciaszerűen magasabb, mint a kontrollszemélyeknél mért értékek; a kérdésekben megállapított hangközértékek jóval nagyobbak a szakirodalomban leírtakhoz képest. A zöngéminőségét tekintve elmondható, hogy mind az értelmi fogyatékos, mind a kontrollcsoporton belüli nőknél nagyobb arányban fordulnak elő részben irreguláris zöngével megvalósuló szakaszok.

1 Bevezetés

A kérdő funkciójú megnyilatkozások vizsgálata régóta jelen van a különböző tárgyú nyelvészeti kutatásokban, tekintve, hogy a kérdések a társas interakciók egyik legfőbb megnyilatkozásformái. A beszélt nyelvben a kijelentés után a kérdés a leggyakrabban használt kifejezésforma (Olaszy 1995). A nyelvészek egyetértenek abban, hogy „a kérdő megnyilatkozások azonosítható szintaktikai szerkezettel, sajátos prozódiai és szemantikai vagy propozicionális tartalommal rendelkeznek” (Markó 2015: 80) – betöltött funkciójuktól függetlenül. A kérdések osztályozása és elemzése szempontjából azonban nincs egyetértés a kutatók között (Markó 2015).

Logikai szemantikai szempontból (amit a tanulmányban vizsgált kérdések csoportosítására alkalmazok) három fő kérdéstípust különböztet meg a szakirodalom. Ez a kiegészítendő, az eldöntendő és a választó kérdés, amelyek mindegyike sajátos intonációs

Jankovics Julianna: Kérdő funkciójú megnyilatkozások enyhe és középsúlyos értelmi fogyatékosok felolvasásában. In Várad Tamás (sorozatszerkesztő), Ludányi Zsófia, Grácz Tekla Etelka (szerkesztő): *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2019. XIII. Alkalmazott Nyelvészeti Doktoranduszkonferencia*. Budapest: MTA Nyelvtudományi Intézet. 2019. 74–91. DOI: 10.18135/Alknyelvdok.2019.13.6

sémával rendelkezik. A kérdő mondatok intonációját az határozza meg, hogy a kérdő funkciót kifejezi-e valamilyen lexikai jelölő (Markó 2017: 139).

A kiegészítendő kérdés (pl. *Ki jön át?*) kérdő névmást tartalmaz, és ennek a kérdőszói mondatrésznek a kategóriája megegyezik az elvárt válaszával, azaz például a *Ki?* kérdésre személyreferencia a válasz (Kugler 2000). Ehhez a kérdéstípushoz szórendi megkötés társul. A kiegészítendő kérdésnek globálisan ereszkedő kontúrja van, vagyis az f_0 -maximum a kérdőszón realizálódik akkor is, ha a kérdőszó nem a mondat elején található (Markó 2017: 139). A dallamesúcsot egy kvartnyi lelépés követi, így dallamkontúrja eső/ereszkedő (Fónagy–Magdics 1967). Több szerző (pl. Olasz 2002) e kérdéstípus egy variánsának tekinti az utolsó szótagban „felkapott” dallamot. Az eldöntendő kérdés (pl. *Adhatok még valamit?*) kérdő névmást nem tartalmaz, ilyenkor a kérdező a kérdés tartalmával vár döntést a partnerétől. A magyarban ezt vagy kérdő intonáció (szökő-eső zárlat) vagy kérdő partikula (*-e, ugye*) jelzi (Kugler 2000). Fónagy és Magdics (1967) szerint ennél a kérdéstípusnál a fellépés átlagos hangközértéke tercnyi, a lelépése pedig kvartnyi. Az eldöntendő kérdő megnyilatkozás intonációját a szótagszám befolyásolja, dallamkontúrája a szökő-eső zárlat jellemző. Ha a kérdés három vagy annál több szótagból áll, akkor az utolsó előtti szótagban lép fel az f_0 , majd az utolsó szótagban lép le. Az egy és két szótagból álló eldöntendő kérdő mondatoknál mindkét folyamat az utolsó szótagban zajlik le, ám ha az utolsó szótag zöngéje túl rövid, a lelépés nem feltétlenül megy végbe (Markó 2017: 139). A választó kérdést (pl. *Tanulunk vagy zenét hallgatunk?*) némely szerző különálló kategóriaként kezeli (pl. Stivers 2010), míg mások az eldöntendő kérdés altípusának tekintik (pl. Kiefer 2000, Kugler 2000). Ez a kérdéstípus a magyarban a szintaktikai szerkezetét tekintve szinte azonos az eldöntendő kérdéssel (a *vagy* kötőszó hozzáadásával) (Kugler 2000). Dallamformáját sajátos intonációs kontúr jellemzi, amely a kötőszóig emelkedő, utána ereszkedő (Markó 2017: 141).

Olasz (2010) ezeken kívül még megkülönbözteti az ellenőrző és a befejezetlen kérdéseket. Az ellenőrző kérdéseket az eldöntendő kérdésekre jellemző dallamszerkezettel valósítjuk meg, míg a befejezetlen kérdések dallamenete – mivel magukban hordozzák a gondolat továbbvitelét – mindig magas alaphangfrekvencia-értéken fejeződik be (2010: 188–189).

A kérdő funkciójú megnyilatkozásokat a magyarban már több évtizede vizsgálják (vö. Bartók 1978; Fónagy–Magdics 1967; Gósy 1993; Gósy–Teren 1994; Varga 1994). Olasz (1995) a kérdés prozódiaját elemezte vizsgálatában, más mondatfajtákkal együtt. A felolvasáson alapuló kísérlet eredményei szerint sem az eldöntendő, sem a kiegészítendő kérdés esetében nem adatolható első szótagi hangsúly, valamint mind a két kérdéstípus jellemző tempóértéke 14 hang/s. Olasz (2002) a magyar kérdés dallamformáinak és intenzitás szerkezetének fonetikai vizsgálatát végezte el összesen 124 izolált és dialógusban szereplő mondaton. A szerző vizsgálatában megállapította, hogy a kérdést minden esetben egy magas-mély f_0 -kontraszt fejezi ki, valamint az f_0 -csúcs megvalósítása egy és két szótagnyi területen valósul meg. Markó (2007) egy kétórás, négyfős társalgásban elemezte a kérdő megnyilatkozásokat közel 200 kérdésen. Vizsgálatában megállapította, hogy a spontán beszédben megjelenő kérdések sokszínűbbek, mint a felolvasásban megjelenők. A változatosság az f_0 -modulációban és a hangköz értékeiben mutatkozott meg. Későbbi kutatásában Markó (2015) közel 5 órányi, 30 beszélőtől származó spontánbeszéd-anyagban (társalgásban) vizsgálta a kérdések (összesen 134 darab) megvalósulását. Hipotézise szerint a kérdések logikai szemantikai típusa, pragmatikai funkciója és prozódiai megvalósulása között összefüggés van, tehát

a valódi (információkérő kérdés, amelyre választ vár a kérdező) kérdések esetében jellemzőbb a prototipikus megvalósulás. Hipotézisét az eldöntendő kérdések esetében támasztották alá az eredmények. A kérdő megnyilatkozások 20-30%-ban nem a szakirodalomban leírt intonációs formában realizálódtak. A nem prototipikus megvalósulások a férfiaknál voltak gyakoribbak, és egyes kérdések pedig irreguláris zöngével valósultak meg. Az irreguláris fonáció az a jelenség, „amikor az alapfrekvencia hirtelen jelentősen, a beszélő jellemző hangterjedelme alá csökken, és ezáltal érzékelhetően megváltozik a hangszínezet” (Bóhm–Ujváry 2008: 108). Funkcióját tekintve sokféle szerepet tölthet be: néhány nyelvben (pl. az angol több nyelvjárásában, a mexikói mazatékban) fonológiai kontrasztot jelöl, kifejezhet különböző érzelmeket és attitűdöket, szegmentumok határát jelezheti stb. (Markó 2013). Tóth (2017) 7 és 18 éves kor közötti (összesen 60 fő) személlyel végzett kutatásában, spontán beszédben és felolvasásban vizsgálta a kérdések megvalósulását. Hipotézisét, amely szerint a spontán beszédben gyakoribbak a nem a szakirodalom szerinti dallammal realizálódott kérdések, alátámasztották az eredmények. Az életkor szerepet játszott a kérdések megvalósulásában, mivel az életkor előrehaladtával csökkent azoknak a kérdéseknek a száma, amelyek a szakirodalom szerinti intonációs sémával valósultak meg. További eredménye volt még a kutatásnak, hogy gyakoriak voltak az irreguláris zöngével megvalósuló kérdések, és ezekből az eredményekből a szerző arra következtetett, hogy az irreguláris zöngé megjelenése leggyakrabban a szegmentumok határát jelzi (Tóth 2017).

A magyar beszédben megjelenő kérdő funkciójú megnyilatkozásokra irányuló kutatások mentálisan ép gyermekek és felnőttek beszédét vizsgálták, ám az értelmi fogyatékosokkal élő beszédében sem a nemzetközi, sem pedig a magyar irodalomban nem találhatók olyan kutatások, amelyek a kérdésekre koncentrálnak. A jelen tanulmányban az *értelmi fogyatékoság, értelmi fogyatékosággal élő* kifejezéseket használom. Mind a nemzetközi, mind a magyar irodalomban látható egyfajta különbség a fogalomhasználat tekintetében a tudományterületek között. A jelenlegi terminus technicus használata a hatályban lévő köznevelési törvény fogalomhasználatát követi (2011. évi CXCV. törvény 2. bekezdésének 3. §-a). A gyógypedagógiában megkülönböztetett tanulásban akadályozottság és értelmileg akadályozottság kategóriáinak legfőbb különbsége az, hogy míg az előbbi kategóriába az enyhe fokú értelmi fogyatékosággal élő személyek sorolhatók, addig a mentálisan súlyosabb fokban sérült emberek alkotják az értelmileg akadályozott személyek csoportját (vö. például Macher 2007, Csákvári 2013). A közép- vagy annál súlyosabb fokban sérült személyek állapotára az intellektuális képességzavar kifejezést is alkalmazzák (vö. Lányiné 2012).

Az értelmi fogyatékoság alatt azt az állapotot értem, amely az alábbi három tulajdonsággal írható le. Az első a normál övezet alatti intelligencia, a második az adaptív működésben kimutatható deficit, a harmadik pedig a korai kezdet. Ez utóbbi azt jelenti, hogy az értelmi fogyatékoság állapota 18 éves kor előtt alakul ki, így különíthető el ez az állapot a felnőttkorban jelentkező degeneratív betegségektől (Hodapp–Dickens 2003, idézi Csákvári–Mészáros 2012: 5). A nyelvi elmaradás szinte minden értelmi fogyatékosággal élő személyre jellemző, hiszen az általános információfeldolgozó funkciók jelentős mértékben sérülnek ebben az állapotban (Lukács–Kas 2014). A jelen kutatásban enyhe és közepsúlyos értelmi fogyatékosággal élő személyek vettek részt. A súlyossági kategóriák megkülönböztetése azért nélkülözhetetlen, mert az értelmi fogyatékoság állapota nem homogén. Míg az enyhe értelmi fogyatékosággal élő személyek IQ-értéke 69–50 pont között, addig a közepsúlyos értelmi fogyatékosággal élőknél ez az érték 49–35 IQ-pont között szóródik (BNO-10 1995; ICD-11 2018).

Az értelmi fogyatékosokkal élők beszédének vizsgálatakor elsősorban a Down- és a Williams-szindrómás személyek beszédére összpontosítanak (vö. például Lányiné 2012; Lukács–Kas 2014). A Down-szindróma olyan zavar, amely anatómiai eltérésekkel és értelmi fogyatékossgal jár együtt. Az ezzel a szindrómával élők között nagy egyéni különbségek fedezhetők fel az értelmi fogyatékossgal súlyosságában és a specifikus képességek sérülésének mértékében is (Lukács–Kas 2014). Náluk gyakori a nagy nyelv, a lapos orrnyél, a kicsi orr, a szűkös szájtér, a fogászati rendellenességek, s nem ritka a halláskárosodás (Stoel-Gammon 2001; Lányiné 2012). További jellemzőként megemlíthető a nazális hangképzés, a rekedtes hangszín, és ők szinte valamennyi nyelvi szinten alacsonyabb teljesítményt érnek el más értelmi fogyatékossgal élő társaikkal viszonyítva (Lezcano–Troncaso 1999; Anilkumar et al. 2017). Beszédük nehezen érthető, a hang minőségét vizsgáló kutatásokban több szerző feltételezése szerint ehhez a szindrómához jellegzetes diszfónia társul (Kent–Vorperian 2013). A Williams-szindróma egy ritka genetikai rendellenesség, az ilyen szindrómával élő személyek főbb jellemzői a manószerű arc, a túlzottan szociális személyiség és az értelmi fogyatékos állapot. A Williams-szindrómával élők nyelvi teljesítménye az értelmi sérüléshez képest kiemelten jó, szókincsük gazdag, mondataik többsége grammatikailag megfelelő (Lányiné 2012; Lukács–Kas 2014).

Az értelmi fogyatékossgal élőknel szinte minden esetben előfordul a megkészt beszédfejlődés, a beszédük főbb jellemzője a gyakori beszédhiba, a dadogásra és hadarásra emlékeztető megakadás, valamint a hangképzés elmosódott, dűnyögő, orrhangzós jellege és a modulálatlan hanghordozás, ami krónikus vagy funkcionális rekedtséggel párosul (Merei 1980; Frint 1982).

A magyar szakirodalomban az értelmi fogyatékossgal élők kérdő funkciójú megnyilatkozásairól csak az a megállapítás olvasható, hogy az értelmi fogyatékosok spontán módon ritkán fogalmazznak meg kérdéseket (Hatos 2000). Egy korábbi kutatásban már vizsgáltam a kérdő funkciójú megnyilatkozásokat: egy enyhe (20 éves) és egy középsúlyos (25 éves) értelmi fogyatékossgal élő nőnél spontán beszédben és felolvasásban elemeztem a kérdések megvalósulását. Eredményeim szerint az enyhe értelmi fogyatékossgal élő női beszélő mindkét beszédtypusban a kontrollhoz hasonlóan valószínűsítette meg a kérdéseket (néhol pragmatikai eltérésekkel), míg a középsúlyos értelmi fogyatékossgal élő adatközlőnél a kérdő intonáció csak a spontán beszédben valósult meg, a kiegészítendő kérdés viszont több alkalommal is kérdő intonációval realizálódott (Jankovics 2017).

A nemzetközi irodalomban nincs tudomásom olyan vizsgálatról, amely a kérdések vizsgálatára irányult volna értelmi fogyatékossgal élőknel. Ám a jelen kutatás szempontjából fontos megemlíteni olyan tanulmányokat is, amelyek a beszéd prozódiai jellemzőit kutatták ebben a populációban. Lee és munkatársai (2009) felnőtt személyeknél vizsgálták az alaphangjellemzőket. Kutatásukban 17–29 év (átlagéletkor: 24,7 év) közötti Down-szindrómás felnőttek (9 fő; 5 nő és 4 férfi) beszédét vetették össze kilenc, nemben és korban illesztett, nem értelmi fogyatékos személy beszédével (átlagéletkor: 23,5 év). Négy feladat – kitartott [ah] hangkapcsolat, emelkedő és ereszkedő glisszandó, felolvasás és egyperces spontán beszéd – alapján megállapították, hogy mindkét nemben a Down-szindrómás személyeknél mérhető magasabb átlagos alaphangjellemző-érték. Albertini és munkatársai (2010) szintén Down-szindrómás személyek beszédében elemeztek különböző akusztikai paramétereket. Vizsgálatukban 30 Down-szindrómás felnőtt (17 férfi, átlagéletkor: 28,7±7,7 év; 13 nő, átlagéletkor 23,2±4,7 év) és 60 kontrollszemély (30 férfi, átlagéletkor: 48,1±8,1 év; 30 nő, átlagéletkor:

44,7±10,3 év) vett részt. A vizsgálat nyelvi anyagát egy, a neuropszichológiában használt afáziaeszt (ENPA, Neuropsychological Examination for Aphasia) adta. A kísérleti személyeknek a hallott szavakat kellett megismételniük két alkalommal, de az elemzés csak az első sorozatra terjedt ki. Az f_0 -értékeken kívül vizsgálták még az intenzitást, a jitter- („a hangszalagrezgések frekvenciaingadozásának mértéke”, Gósy 2004: 31) és a shimmerértékeket („a hangszalagrezgések amplitúdóingadozásának mértéke”, Gósy 2004: 31), valamint az időtartamot. A kutatás során megállapították, hogy a Down-szindrómás férfiaknál és nőknél (férfi: 162,4±30,0 Hz; nő: 213,6±23,2 Hz) szignifikánsan magasabb volt az átlagos alapfrekvencia a kontrollszemélyekhez viszonyítva (férfi: 119,1±11,0 Hz; nő: 193,1±18,8 Hz), továbbá ugyanebben a populációban mindkét nem esetében szignifikánsan alacsonyabb intenzitásértékek születtek a felnőtt kontrollszemélyekhez viszonyítva. Ugyancsak a Down-szindrómás férfiaknál rövidebb időtartamot és alacsonyabb shimmerértékeket mértek a kontrollszemélyekhez képest. A Down-szindrómásoknál mért alacsonyabb shimmerértékek – habár ez nem volt része a vizsgálatnak – utalhatnak arra, hogy a beszédük kisebb arányban valósult meg irreguláris zöngével. Egy másik vizsgálatban, szintén elsősorban az átlagos f_0 -jellemzőkre koncentrálna, huszonnégy 20 és 28 év közötti (átlagéletkor: 25,0 év; 8 nő és 14 férfi), legalább 50 és legfeljebb 65 IQ-ponttal rendelkező Down-szindrómás beszélő értékeit vették össze huszonnégy nemből és korban illesztett ép beszélő értékeivel (Seifpanahi et al. 2011). Az akusztikai elemzést a kitartott [a] hang (öt alkalommal rögzített) mintáin végezték el, mindössze három másodperces anyagot elemezve. A szerzők megállapították, hogy az átlagos alapfrekvencia mindkét nemből szignifikánsan magasabb volt a Down-szindrómás személyeknél a kontrollszemélyek értékeihez viszonyítva. Feltételezhető, hogy a Down-szindrómásoknál mért magasabb f_0 oka az alkati különbségekben és a felső légút anatómiai eltérésében keresendő. Egy korábbi vizsgálatomban öt értelmi fogyatékossgal élő (átlagéletkoruk: 32,4 év) és öt nemből és korban illesztett mentálisan ép (átlagéletkor: 32,0 év) nő vett részt. A kutatásban négy beszéd típusban (felolvasás, interjú, képleírás, tartalomösszegzés) elemeztem az alaphangjellemzőket kijelentő megnyilatkozásokban. A beszéd típus és az alaphangjellemzők összefüggéseire vonatkozó hipotézist csak részben támasztották alá az eredmények. Mindkét csoportban a felolvasásban adatoltam a legmagasabb átlagos f_0 -t, de a spontán beszédben egyértelmű tendenciák nem rajzolódottak ki, inkább a beszélők közötti variabilitás volt jellemző. Egy értelmi fogyatékos személy értékeit kivéve az enyhe és a középsúlyos értelmi fogyatékossgal élőknek (a nemzetközi eredményekhez hasonlóan) magasabb volt az átlagos f_0 -értékük minden beszéd típusban (Jankovics 2019).

A jelen kutatásban enyhe és középsúlyos értelmi fogyatékossgal élő személyek háromszori felolvasásában vizsgálok a kérdő funkciójú megnyilatkozásokat, és ezeket az eredményeket hasonlítom össze a kontrollszemélyek eredményeivel. A kutatásban fő célom a beszélők közötti és a beszélőn belüli variabilitás feltérképezése a felolvasásokban lévő kérdések alapján. A hipotéziseket – a heterogén adatközlői csoport miatt – a kutatásban részt vevő személyekre korlátozva fogalmazom meg. Hipotéziseim szerint: 1. az értelmi fogyatékossgal élők átlagos alapfrekvencia-értéke tendenciaszerűen magasabb, mint a kontrollszemélyek értékei; 2. az egyes felolvasásokban az értelmi fogyatékossgal élők kérdései változatosabb dallammenettel realizálódnak, mint a kontrollszemélyek kérdései; 3. az értelmi fogyatékossgal élő személyek kérdéseiben kevesebb az irreguláris zöngével megvalósuló szakasz.

2 Kísérleti személyek, anyag, módszer

A jelen vizsgálatban összesen hét értelmi fogyatékossgal élő (átlagéletkor: 41,0 év; 3 nő és 4 férfi) és hét nemből és korban (átlagéletkor: 41,4 év; 3 nő és 4 férfi) illesztett (maximum ± 3 év korkülönbség az értelmi fogyatékos és a hozzá illesztett kontrollszemély között) mentálisan ép adatközlő vett részt. A vizsgálatban részt vevő adatközlőket háromkarakteres jelzéssel azonosítottam. Az első karakter a mentális állapotot (E = értelmi fogyatékossgal élő, K = kontrollszemély) jelzi, a második karakter a beszélő nemére utal (N = nő, F = férfi), a harmadik karakter pedig az adatközlő életkorát jelzi. Az 1. táblázatban az értelmi fogyatékossgal élők azonosítóját, mentális állapotát, valamint a hozzájuk illesztett kontrollszemélyek azonosítóját tüntettem fel. Az értelmi fogyatékossgal élők (EF37 kivételével) mindannyian színészek, akik aktívan játszanak különböző előadásokban, a Baltázár és a Mázínház társulatában. Az adatközlők magyar anyanyelvűek.

A vizsgálat előtt mindegyik adatközlővel beleegyező nyilatkozatot írtam alá (abban az esetben, ha gondnokság alatt állnak, akkor a szülő vagy a gondnok írta alá a dokumentumot). A nyilatkozatban önkéntes alapon szerepelt az a kitétel, hogy a felvételvezető betekinthez az adott személy orvosi dokumentációjába. Néhányan ehhez nem járultak hozzá, náluk a gondnok vagy a szülő adott információt a sérülés súlyosságáról. A pontos IQ-értékeket azért nem tüntettem fel, mert egy-egy személy több mint tízéve készült orvosi dokumentumában szereplő pontos IQ-pontszámot nem tartottam releváns információnak (csak a BNO-kódot, ami az enyhe értelmi fogyatékossgal élők esetében F.70, a közép súlyos értelmi fogyatékosok esetében pedig F.71).

Azonosító	Mentális állapot (értelmi fogyatékossgal)	Illesztett kontrollszemély
EN25	Down-szindróma, közép súlyos é. f.	KN24
EN30	Williams-szindróma, enyhe é. f.	KN29
EF37	Williams-szindróma, enyhe é. f.	KF37
EF41	nem adott információt	KF42
EF43	Down-szindróma, közép súlyos é. f.	KF44
EF49	nem adott információt	KF49
EN62	enyhe é. f.	KN65

1. táblázat. A kutatásban részt vevő értelmi fogyatékos adatközlők és a hozzájuk illesztett kontrollszemélyek azonosítói

A hangfelvételeket az ELTE Fonetikai Tanszékének stúdiójában rögzítettem 44,1 kHz-es mintavételezési frekvencián, legalább 16 biten az Audacity (Audacity Team 2018) program segítségével. A felvételnézéshez fejmikrofont használtam. Az adatközlőkkel felolvasást rögzítettem. A felolvasás során egy párbeszédet kellett megfogósítani az adatközlőknek. A szöveg részben megegyezik a GABI (Bóna et al. 2014) adatbázisban található és a Tóth (2017) által használt szöveggel. A szöveget háromszori felolvasásban rögzítettem. A felolvasások között más beszéd típusokat vettem fel, ám ezek nem képezik a jelenlegi kutatás anyagát. Az első felolvasás előtt az adatközlők elolvashatták a szöveget, ennek nem volt időtartambeli megkötése.

A jelen vizsgálatban szereplő adatközlők között jelentősek az egyéni különbségek, mind az életkor, mind az értelmi fogyatékossgal súlyossága, valamint a sérülés fajtája tekintetében. Ebből kifolyólag az adatközlőket vizsgálva hét esettanulmányról van szó,

viszont az összehasonlíthatóság érdekében most egy csoportként kezelem az értelmi fogyatékos személyeket, és ebből kifolyólag alkalmazom a *csoport* kifejezést.

A kutatásban a felolvasásban előforduló kérdő funkciójú megnyilatkozásokat elemeztem. A felolvasási feladatban összesen három kérdés fordult elő, ebből 2 eldöntendő (*Nincs kedved velem jönni?*, *És van pénzed fagyira?*), 1 pedig kiegészítendő (*Miért, hová megyünk?*) kérdés volt. Az utóbbit, amely egy összetett kérdés, két kiegészítendő kérdésként kezeltem, mivel mindkét tagja saját dallammenettel rendelkezett. A kutatásban egy beszélőtől egy felolvasásban 2 eldöntendő és 2 kiegészítendő kérdést vizsgáltam, vagyis a háromszori felolvasásban összesen 12 darab kérdés fordult elő személyenként.

A hanganyagban a Praat 6.0.43 szoftver (Boersma–Weenink 2018) segítségével manuálisan annotáltam a kérdő funkciójú megnyilatkozásokat. A megnyilatkozásokban manuálisan jelöltem a frekvenciacsúcs helyét (tehát azt a szótagot, amelyben az f_0 -max megvalósult), valamint a megelőző és a követő szótagok f_0 -minimumait. A Praat beállított tartománya 70 és 400 Hz között helyezkedett el. Ez alól két személy volt kivétel, EF43 és EF49, akiknél perceptuális döntés alapján 50 és 350 Hz között állítottam be a vizsgálni kívánt frekvenciatartományt. Ezekből az adatokból kiszámítottam a fel- és lelépés értékét, amelyet a tanulmányban félhangokban adok meg. Rögzítettem a zöngé minőségét is (modális vagy irreguláris zöngé). Az irreguláris zöngét manuálisan annotáltam. Ennek időtartamát, valamint az egyes kérdések átlagos alaphangfrekvenciáját automatizálva mértem. Az alaphangfrekvencia számításánál az automatikusan mért tizedmásodpercenkénti értékekből határoztam meg az átlagértéket. A minta méretéből következően statisztikai próbákat nem végeztem el a felvett anyagon.

3 Eredmények

3.1 Alaphangfrekvencia-érték

Az átlagos alaphangfrekvencia értékét elsőként az egyes beszélőknél – a háromszori felolvasást együtt elemezve – számítottam ki automatizálva (1. és 2. ábra). Az értelmi fogyatékossgal élő nők esetében a legmagasabb értéket EN30-nál adatoltam, nála ez az érték 263 Hz (szórás 31 Hz) volt, míg a legalacsonyabb átlagos f_0 -t, 206 Hz-et (szórás 41 Hz) EN62-nél mértem. A kontrollszemélyeknél az életkor tekintetében ugyanilyen sorrendben alakultak az eredmények. A legmagasabb f_0 -t KN29-nél (220 Hz, szórás 46 Hz), a legalacsonyabbat pedig KN65-nél (196 Hz, szórás 56 Hz) mértem (1. ábra). A férfiak esetében az értelmi fogyatékos személyeknél a legmagasabb átlagos alaphangfrekvencia értéke EF41-nek (192 Hz, szórás 42 Hz), a legalacsonyabb f_0 -értéke pedig EF49 (129 Hz, szórás 29 Hz) személynek volt. A nőkkel ellentétben itt nem érvényesült az értelmi fogyatékos és a hozzá illesztett kontrollszemély értéke közötti párhuzam, mert a legmagasabb átlagos f_0 -értéket (a felolvasásokat együtt vizsgálva) KF37-nél (155 Hz, szórás 39 Hz) mértem, míg a legalacsonyabb értéke KF44 személynek (114 Hz, szórás 20 Hz) volt (2. ábra).

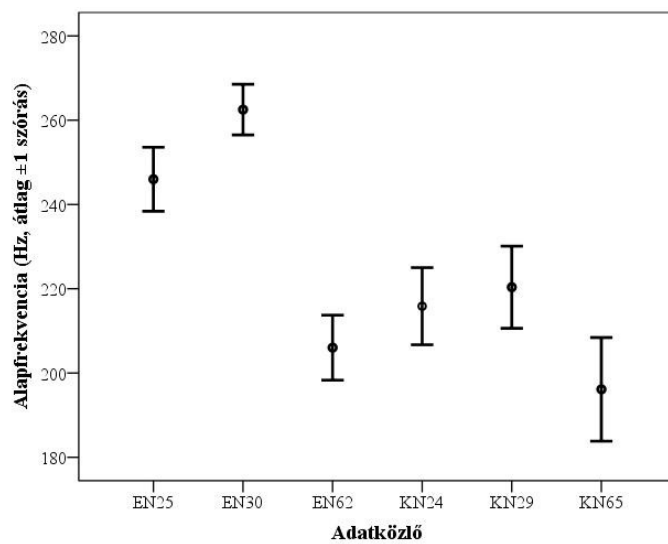
A női résztvevőket vizsgálva megállapítható, hogy az értelmi fogyatékos kísérleti személyeknek a hozzájuk nemben és korban illesztett nőkhöz viszonyítva minden esetben magasabb volt az átlagos f_0 -értékük. Látható, hogy EN62-nek és KN65-nek volt a saját csoportjában a legmélyebb átlagos alaphangfrekvencia-értéke. Azt, hogy idősebb korban – a különböző fiziológiai változások hatására – megváltozik a nők (és a férfiak)

alapfrekvenciája, már több kutatás is feltárta (lásd pl. magyarra Bóna 2012; 2013). Ha a csoporton belüli mindhárom adatközlő értékét együtt számítjuk, akkor az értelmi fogyatékos nők átlagos f_0 -értéke 238 Hz (szórás 45 Hz), míg a kontrollszemélyeké 211 Hz (szórás 49 Hz) volt.

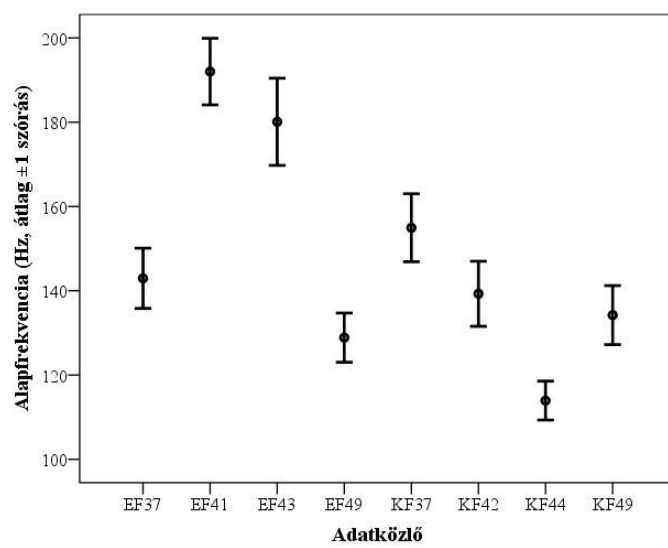
A férfiaknál ilyen tendencia nem állapítható meg. Igaz ugyan, hogy az értelmi fogyatékos csoporton belül a legidősebb adatközlőnél, EF49-nél adatoltam a legmélyebb f_0 -t, a kontrollszemélyeknél viszont ez nem így történt. A csoportok átlageredményeit tekintve elmondható, hogy – a nőkhöz hasonlóan – a férfiaknál is magasabb volt az értelmi fogyatékos csoportban számított átlagos f_0 -érték (163 Hz, szórás 47 Hz) a kontrollszemélyekhez viszonyítva (137 Hz, szórás 35 Hz). A jelen kutatásban is megvalósult a nemzetközi szakirodalomban leírt tendencia, amely szerint az értelmi fogyatékos adatközlőknek magasabb az alapfrekvencia-értékük a nemből és korból illesztett, kontrollszemélyekhez viszonyítva.

A kérdés típusát külön vizsgálva megállapítható, hogy mindkét nem esetében magasabb átlagos alapfrekvenciával valósultak meg az eldöntendő kérdések, mint a kiegészítendő kérdések (3., 4. ábra). Az értelmi fogyatékos nők átlagos f_0 -értéke az eldöntendő kérdéseknél 243 Hz (szórás 40 Hz), míg a kiegészítendő kérdéseknél 227 Hz (szórás 52 Hz) volt. A mentálisan ép nőknek az átlagos f_0 -értéke az eldöntendő kérdésekben 218 Hz (szórás 47 Hz), a kiegészítendő kérdésekben pedig 197 Hz (szórás 51 Hz) volt (3. ábra). Az értelmi fogyatékos férfiaknál az eldöntendő kérdésekben 169 Hz (szórás 43 Hz), a kiegészítendőkében 148 Hz (szórás 53 Hz) volt az átlagos alapfrekvencia-érték. A kontrollcsoportban az eldöntendő kérdések esetében 143 Hz-et (szórás 32 Hz), a kiegészítendőké esetében 124 Hz-et (szórás 39 Hz) adatoltam (4. ábra).

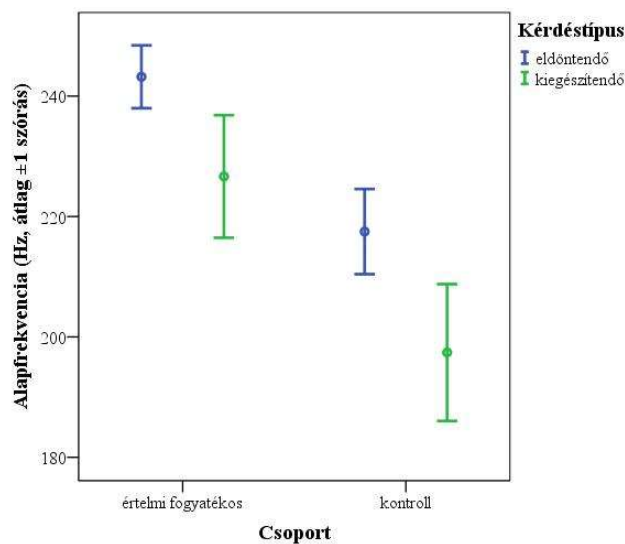
A kérdéstípusokban mért alapfrekvencia-értékek alakulásában vélhetően a szakirodalom által leírt sémák játszanak szerepet (vö. Bevezetés). Míg az eldöntendő kérdéseknél az utolsó előtti szótagon realizálódik a frekvenciacsúcs (tehát a kérdések dallama az f_0 -csúcsig emelkedő), addig a kiegészítendő kérdések esetében a kérdőszón van a frekvenciacsúcs, és a lexikai szerkezet miatt nincs sajátos, csak rá jellemző dallama ennek a kérdéstípusnak, azaz a kérdés végéig ereszkedő alapfrekvenciát várunk (Markó 2017). A 3. és a 4. ábra jól szemlélteti, hogy mind az értelmi fogyatékosok, mind a kontrollszemélyek csoportjában mindkét nemnél ez utóbbi kérdéstípus esetében volt nagyobb a szórás értéke.



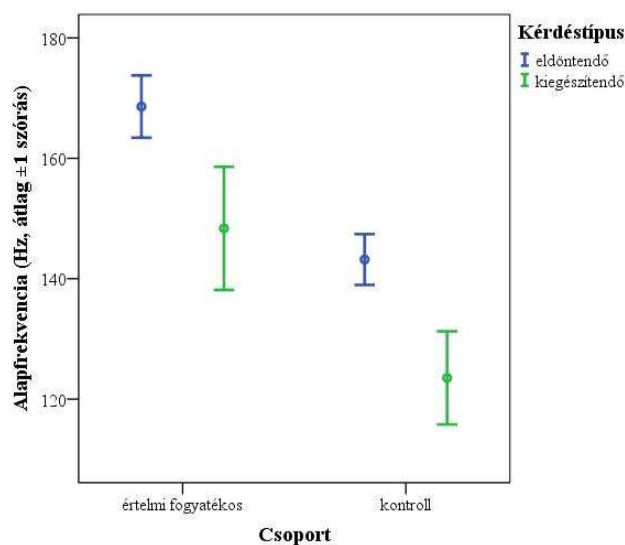
1. ábra. A női adatközlők átlagos f_0 -értéke



2. ábra. A férfi adatközlők átlagos f_0 -értéke



3. ábra. A női adatközlők átlagos f_0 -értéke a kérdéstípus függvényében



4. ábra. A férfi adatközlők átlagos f_0 -értéke a kérdéstípus függvényében

3.2 A kérdő funkciójú megnyilatkozások dallama

Az eldöntendő kérdésekben a frekvenciacsúcs helyét és a megelőző, valamint a követő szótagban az f_0 minimumát, a kiegészítendő kérdésekben (mivel mindkettő kérdés első szava a kérdőszó) csak a követő szótag f_0 -minimumát mértem meg. Ebből számoltam

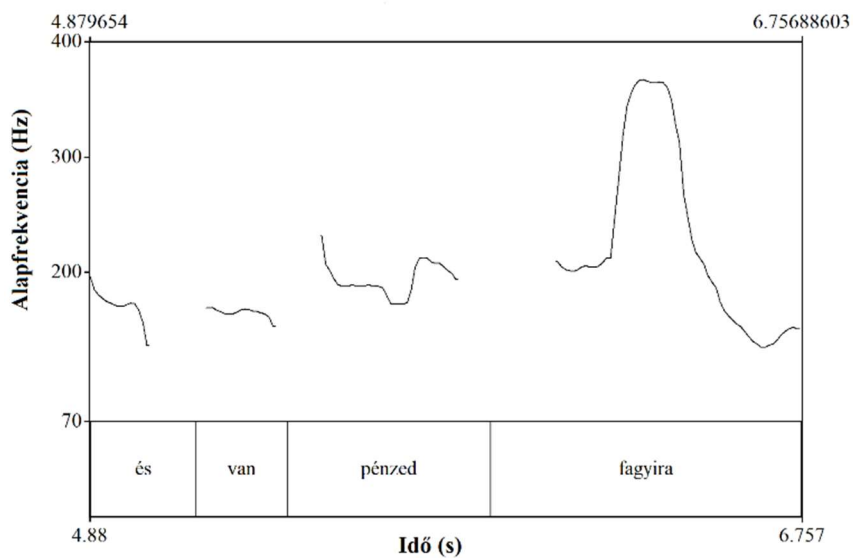
ki a fel- és a lelépés hangközértékét, amelyet a tanulmányban félhangokban adok meg. Abban az esetben, hogyha a frekvenciacsúcsot tartalmazó szótag utáni szótag irregulárisan valósult meg, akkor az irreguláris zöngét követő modális szakasz f_0 -értéket vettem figyelembe. Ahol az irreguláris rész a szótag végéig tartott, nem számoltam ki a lelépés hangközét. A kiegészítendő kérdéseknél EN62 személy 1. felolvasásában a *Hová...?* kezdetű kérdést, az eldöntendő kérdéseknél pedig KF42 személy 3. felolvasásában az *És...?* kezdetű kérdést nem vettem figyelembe, mert mindkettő a leírt szövegtől eltérő szórenddel valósult meg.

Az eldöntendő kérdések esetében a női adatközlőknél az értelmi fogyatékos csoportban a fellépés hangköze 3,4 félhang (szórás 1,1 félhang), a lelépése 6,4 félhang (szórás 2,1 félhang) volt átlagosan. A kontrollcsoport női tagjainál a fellépés mértéke 5,1 félhang (szórás 2,1 félhang), a lelépés mértéke 8,3 félhang (szórás 3,6 félhang) volt átlagosan. Az értelmi fogyatékos nők frekvenciaváltozásának értéke a fellépés és a lelépés során megegyezett a szakirodalom által (Fónagy–Magdics 1967) leírt értékekkel. A szakirodalomban leírt terenyi érték 3–4 félhangnak, a kvartnyi érték pedig 5–6 félhangnak felel meg. A kontrollcsoport női beszélőinek értéke minden esetben meghaladta a szakirodalomban leírt értéket. A kontrollcsoportban a lelépés során az extrém kiugró értékeket KN29-nél (legnagyobb érték: 16,8 félhang; szórás 3,1 félhang) és KN65-nél (legnagyobb érték: 17,8 félhang; szórás 4,0 félhang) adatoltam. Az eldöntendő kérdéseknél az értelmi fogyatékos férfiaknál a fellépés hangköze 7,4 félhangnyi (szórás 2,7 félhang), a lelépése 9,5 félhangnyi (szórás 5,4 félhang) volt átlagosan. A kontrollcsoportban a fellépésnél 5,1 félhangnyi (szórás 3,8 félhang), a lelépésnél 8,6 félhangnyi (szórás 5,5 félhang) átlagértéket mértem. Amint látható, a férfiak mindkét adatközlői csoportban nagyobb frekvenciaingadozással valósították meg a fel- és lelépést. EF41 adatközlő – habár a szakirodalom szerinti dallammal valósította meg a kérdést – esetében az eldöntendő kérdéseknél nagyon nagyok voltak a lelépés hangközértékei (5. ábra). Ennek vélhetően az volt az oka, hogy az adatközlő az általa hallott mintát (az eldöntendő kérdésekre jellemző dallamkontúrt) akarta jól megvalósítani. A második és a harmadik felolvasás *És...?* kezdetű kérdésében 17,5–17,5 félhangnyi volt a lelépés hangközértéke.

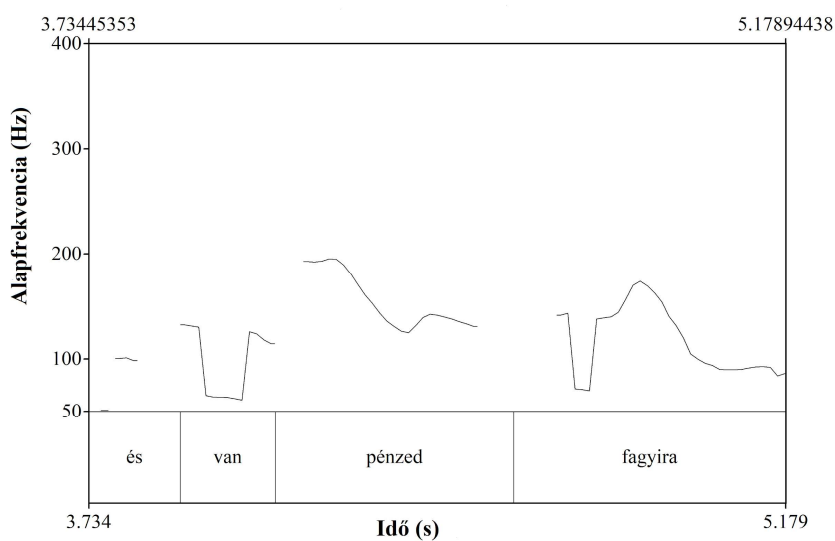
EF37 adatközlő mindhárom *És van...?* kezdetű kérdésében két frázist lehetett elkülöníteni, amelyekben a kérdő intonáció (mindhárom felolvasás esetében) mindkét frázisban előfordult (6. ábra).

A női adatközlőknél az értelmi fogyatékosokkal élők esetében a lelépés átlagértéke 7,5 félhang (szórás 3,0 félhang), a kontrollszemélyeknél pedig 10,8 félhang (szórás 3,7 félhang) volt. A férfi értelmi fogyatékosoknál a lelépés értéke 9,1 félhang (szórás 4,5 félhang), a kontrollszemélyeknél 11,8 félhang (szórás 4,1 félhang) volt.

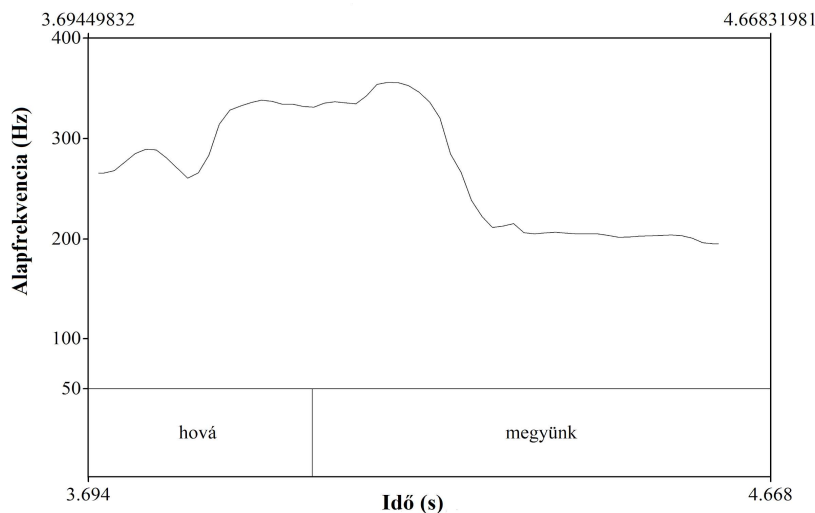
A kérdéstípusok és a felolvasások függvényében megvizsgáltam a kérdések megvalósulását. Ennél a számításnál szintén nem vettem figyelembe azt a két kérdést, amely nem a leírt szórenddel valósult meg. Az értelmi fogyatékos csoportban az 1. felolvasás során az eldöntendő kérdéseknek a 38,5%-a valósult meg szökő-eső zárlattal, a kiegészítendőknél a 91,7%-a valósult meg eső-ereszkedő hanglejtéssel. A 2. felolvasásban az eldöntendők esetében a kérdések 50%-a, a kiegészítendőknél a kérdések 100%-a valósult meg a szakirodalomban leírt realizációs formában. A 3. felolvasásban az eldöntendőknek a 42,9%-a, a kiegészítendőknél pedig szintén a 100%-a valósult meg a szakirodalom által leírt formában. Az 1. felolvasás során EN25 volt az, akinek a *Hová megyünk?* kiegészítendő kérdésében nem a kérdőszó első szótagjában, hanem a *megyünk* szó első szótagjában adatoltam az f_0 -maximumot (7. ábra).



5. ábra. Az *és van pénzed fagyira?* kérdés egy realizációja (EF41, 3. felolvasás)



6. ábra. Az *és van pénzed fagyira?* kérdés irreguláris zöngét tartalmazó realizációja (EF37, 1. felolvasás)



7. ábra. A *Hová megyünk?* kérdés egy realizációja (EN25, 1. felolvasás)

A kontroll adatközlőknél az eldöntendő kérdéseknek az 1. felolvasásban a 64,3%-a, a 2. felolvasásban a 42,9%-a, a 3. felolvasás során pedig a 69,2%-a valósult meg az utolsó két szótagon szökő-eső dallamformával. A kiegészítendő kérdések mindegyike mindhárom felolvasásban a szakirodalom szerint realizálódott. Az eldöntendő kérdések esetében az 1. és a 3. felolvasásban nagyobb százalékban valósultak meg a kérdések a szakirodalom szerinti realizációs formában a mentálisan ép férfiaknál és nőknél. Azonban fontos megjegyezni, hogy nagy volt az egyéni variabilitás mindkét adatközlői csoporton belül.

3.3 A zöngé minősége a kérdő funkciójú megnyilatkozásokban

A kérdésekben manuálisan jelöltem az irregulárisan megvalósuló szakasz hosszát és annak a megnyilatkozásban elfoglalt helyét (a megnyilatkozás első szótagja, a megnyilatkozás belseje, a megnyilatkozás utolsó szótagja). A nem modális részek időtartamát automatizálva mértem.

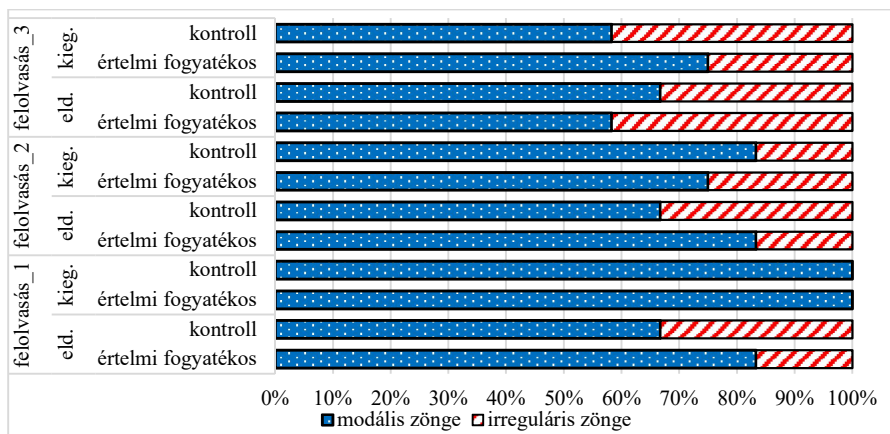
Az értelmi fogyatékkal élő nőknél az 1. felolvasásban az eldöntendő kérdések 16,7%-a valósult meg valamekkora részben irreguláris zöngével. A kiegészítendő kérdések mindegyike csak modális zöngét tartalmazott. A 2. felolvasásban az eldöntendő kérdések 16,7%-a, a kiegészítendőknél a 25%-a valósult meg részben irreguláris zöngével. A 3. felolvasásban az eldöntendőknél ez az arány 41,7%, a kiegészítendőknél 25% volt. A kontrollszemélyek női adatközlőinél az 1. felolvasásban az eldöntendő kérdések 33,3%-a valósult meg részben irreguláris zöngével. Az 1. felolvasás során náluk sem fordult elő olyan kiegészítendő kérdés, amely irreguláris zöngét tartalmazott. A 2. felolvasásban az eldöntendő kérdéseknek szintén a 33,3%-a, a kiegészítendőknél a 16,7%-a valósult meg részben irreguláris fonációval. A 3. felolvasásban az eldöntendőknél a 33,3%-a, a kiegészítendőknél a 41,7%-a valósult meg részben

irreguláris zöngével. A női adatközlők mindegyikénél adatoltam olyan megnyilatkozásrészeket, amelyek irreguláris zöngét tartalmaztak (8. ábra).

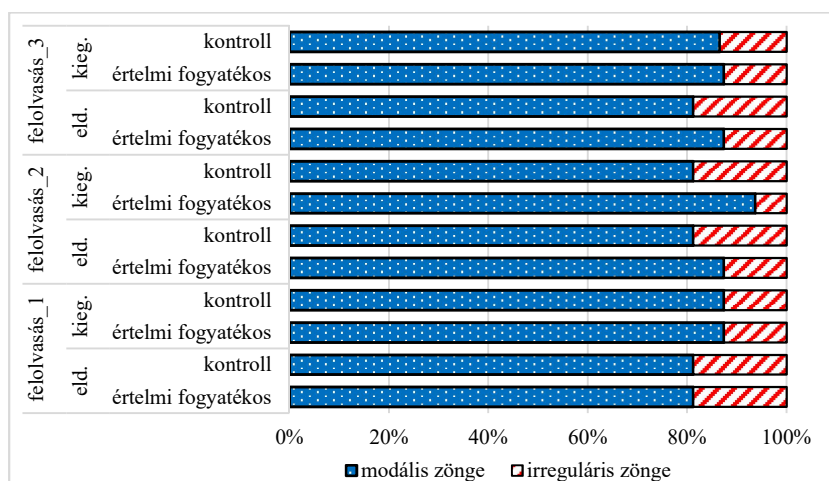
A férfiak esetében az értelmi fogyatékosok közül EF41-nél és EF43-nál, a kontrollszemélyeknél KF42-nél egyáltalán nem fordult elő irreguláris részt tartalmazó kérdő funkciójú megnyilatkozás. Az értelmi fogyatékkal élők esetében az 1. felolvasásban az eldöntendő kérdések 18,8%-a, a kiegészítendő 12,6%-a valósult meg részben irreguláris zöngével. A 2. felolvasásban az eldöntendő kérdések 12,6%-a, a kiegészítendő 6,3%-a tartalmazott irreguláris szakaszokat. A 3. felolvasásban az eldöntendő és a kiegészítendő kérdéseknek is 12,6%-a valósult meg valamely részben irregulárisan. A kontrollszemélyek esetében az 1. felolvasásban ugyanolyan arányban fordultak elő irreguláris részt tartalmazó kérdések mindkét típusban, mint az értelmi fogyatékosoknál (18,8% és 12,6%). A 2. felolvasásban mind az eldöntendő, mind a kiegészítendő kérdéseknek a 18,8%-a valósult meg részben irreguláris zöngével, míg a 3. felolvasás során az eldöntendő kérdéseknek a 18,8%-a, a kiegészítendőkné a 12,6%-a tartalmazott irreguláris zöngét (9. ábra).

A felolvasásokat és a kérdéstípusokat összességében nézve az értelmi fogyatékkal élő női adatközlőknél a kérdések 41,7%-a, míg a mentálisan ép női beszélőknél a kérdések 52,8%-a valósult meg részben irreguláris zöngével. A férfiaknál ez az arány az értelmi fogyatékosoknál 25%, a kontrollszemélyeknél 33,3%. Látható, hogy mindkét nem esetében a kontrollszemélyeknél valósult meg több kérdés irregulárisan, és a két nem közül pedig a nőknél fordult elő több irreguláris zöngét tartalmazó kérdés.

Az irreguláris zöngé az értelmi fogyatékos nőknél a nem modális zöngét tartalmazó kérdések 13,3%-ban a kérdések első és utolsó szótagjában is megjelent, míg a kérdések 86,7%-ban a végén volt észlelhető. A kontrollszemélyek női tagjainál az irreguláris zöngé a kérdések 10,5%-ban a kérdések belsejében, 21%-ban a belsejében és a végén, 5,3%-ban az elején és a végén, 63,2%-ban csak a végén jelent meg. Az értelmi fogyatékos férfiaknál az irreguláris zöngé a nem modális zöngét tartalmazó kérdésekben 15,4%-ban a megnyilatkozás belsejében és végén, 7,7%-ban a belsejében, 69,2%-ban a végén jelent meg. Egy kérdés esetében (7,7%) az irreguláris zöngé a megnyilatkozás első szótagjában, a megnyilatkozás belsejében és az utolsó szótagban is megjelent. A kontrollszemélyek férfi tagjainál 6,3%-ban az irreguláris zöngé a kérdések belsejében és végén, 15,3%-ban az elején és a végén jelent meg. Egyetlen kérdés esetében (6,3%) az irreguláris zöngé csak a megnyilatkozás első szótagjában jelent meg. Az irreguláris zöngét tartalmazó kérdések 75%-ban az irreguláris fonáció csak a kérdések utolsó szótagjában volt észlelhető.



8. ábra. Az irreguláris zöngé előfordulásának aránya a női adatközlőknél



9. ábra. Az irreguláris zöngé előfordulásának aránya a férfi adatközlőknél

4 Következtetések

A jelen kutatásban enyhe és közepesúlyos értelmi fogyatékosággal élők háromszori felolvasásában vizsgáltam a kérdő funkciójú megnyilatkozásokat, és ezeket vettem össze a nemben és korban illesztett kontrollszemélyek felolvasásaiban található kérdéseivel. Tudomásom szerint mindeztidáig sem a nemzetközi, sem a hazai irodalomban nem vizsgálták az értelmi fogyatékosággal élő személyek kérdéseit fonetikai szempontból.

Első hipotézisemet, amely az átlagos alapfrekvencia-értékekre vonatkozott, alátámasztották az eredmények. A nemzetközi irodalom eredményeihez hasonlóan a jelen vizsgálatban is tendenciaszerűen magasabb volt mind a női, mind a férfi értelmi fogyatékosoknál.

tékossággal élők átlagos f_0 -értéke a hozzájuk nemben és korban illesztett kontrollszemélyekhez viszonyítva. Ennek oka az értelmi fogyatékosággal élő személyek (elsősorban a Down-szindrómások) eltérő anatómiai és fiziológiai felépítésében keresendő. A kérdések típusát tekintve mindkét adatközlői csoportban az eldöntendő kérdésekben volt magasabb az átlagos f_0 értéke. A legmagasabb szórásértékeket viszont a kiegészítendő kérdésekben adtam.

Második hipotézisemet, amelyben azt feltételeztem, hogy az értelmi fogyatékosok kérdései realizálódnak változatosabb dallammenettel, a kérdések dallamával, vagyis a frekvenciacsúcsot megelőző és követő szótagok között mérhető hangközértékek kiszámításával elemeztem. Hipotézisemet nem támasztották alá az eredmények. Az értelmi fogyatékos adatközlőknél kisebb volt mind a fellépés, mind a lelépés értéke mindkét kérdéstípusban a kontrollszemélyekhez képest. Ám fontos megjegyezni, hogy mindkét adatközlői csoportban, mind a női, mind a férfi beszélők fellépésének és lelépésének átlagos hangközértéke jóval meghaladta a szakirodalomban leírt értéket. Ennek okára a jelenlegi vizsgálat nem tud magyarázatot adni. A több adatközlővel elvégzett kutatás választ adhat arra a kérdésre, hogy egyéni varianciáról vagy egyéb változásról van-e szó.

Harmadik hipotézisemet egy nemzetközi, Down-szindrómásokkal végzett kutatás eredményeire alapoztam. Feltételeztem, hogy az értelmi fogyatékosággal élő személyek kérdéseiben kevesebb lesz az irreguláris zöngével megvalósuló szakasz. Ezt a feltételezést alátámasztották az eredmények. Mindkét nem esetében az értelmi fogyatékosok kérdéseiben adtam kevesebb, részben irreguláris zöngével megvalósuló szakaszt. A nemek tekintetében mindkét adatközlői csoport esetében a női beszélőknél fordult elő nagyobb arányban olyan kérdés, amely irreguláris zöngét tartalmazott. Ez az eredmény megegyezik Markó (2013) megállapításával. Továbbá a vizsgálat eredményei szerint az irreguláris zöngé mindkét adatközlői csoportban legnagyobb arányban a kérdések utolsó szótagjában fordult elő.

A kutatás anyagát képező kérdő funkciójú megnyilatkozások elemzéséhez fontos adalék lehet a kísérleti személyek olvasási teljesítményének értékelése. Ez túlmutat a tanulmány keretein, ezért ennek elemzése most nem történt meg. Ám fontos megjegyezni, hogy egyik értelmi fogyatékos személynek sem okozott gondot a párbeszéd felolvasása. A kérdésekben előforduló megakadásjelenségek elemzése egy következő tanulmány témáját alapozhatja meg.

A jelen kutatásban részt vevő értelmi fogyatékosággal élő személyek egy kivétellel aktívan játszanak különböző előadásokban. A jövőben érdemes lesz olyan adatközlők beszédmintáit is összehasonlítani, akik nem számítanak képzett beszélőnek, így kiderülhet, hogy ez a tényező mely paraméterekre van hatással. Továbbá fontos kiemelni, hogy a kísérleti személyek között mind az életkor, mind az értelmi fogyatékoság súlyossága és a különböző etiológiájú sérülés miatt nagyok az egyéni különbségek, így a vizsgálat eredményeit és következtetéseit nagy körültekintéssel kell kezelni.

A jelen kutatás eredményei hozzájárulhatnak az enyhe és a közepesúlyos értelmi fogyatékosággal élők beszédbeli sajátosságainak megismeréséhez. A kutatás megválaszolásra váró kérdéseire az adatközlőszám bővítésével kaphatunk válaszokat.

Irodalom

- Albertini, G. – Bonassi, S. – Dall’Armi, V. – Giachetti, I. – Giaquinto, S. – Mignano, M. 2010. Spectral analysis of the voice in Down syndrome. *Research in Developmental Disabilities* 31(5): 995–1001.
- Anilkumar, A. – Suresh, P. A. – Devassy, D. 2017. Speech and language challenges in Indian children with Down syndrome. *Asia Pacific Journal of Research* 50: 1–3.
- Audacity Team 2018. (Version 2.3.1). Elérhető: <https://www.audacityteam.org/>
- Bartók J. 1978. A négyféle kérdő hanglejtés. *Magyar Fonetikai Füzetek* 1: 97–103.
- BNO-10 1995. *A betegségek és az egészséggel kapcsolatos problémák nemzetközi statisztikai osztályozása*. 10. revízió. 1. kötet. Budapest: Népjóléti Minisztérium.
- Boersma, P. – Weenink, D. 2018. *Praat: doing phonetics by computer* (Version 6.0.43). Elérhető: http://www.fon.hum.uva.nl/praat/download_win.html (letöltés ideje: 2018. 09. 01.)
- Bóna J. 2012. A spontán beszéd sajátosságai idősebb, idős és matuzsálemi korban. In: Markó A. (szerk.) *Beszéstudomány. Az anyanyelv-elsajátítástól a zöngékezdési időig*. Budapest: ELTE BTK – MTA Nyelvtudományi Intézet. 100–115.
- Bóna J. 2013. *A spontán beszéd sajátosságai az időskorban*. Budapest: ELTE Eötvös Kiadó. Elérhető: <http://www.elteader.hu/kiadvanyok/a-spontan-beszed-sajatossgai-az-idoskorban/> (letöltés ideje: 2017. 09. 13.)
- Bóna J. – Imre A. – Markó A. – Váradí V. – Gósy M. 2014. GABI – Gyermeknyelvi beszéd-AdatBázis és Információtár. *Beszédkutatás* 2014: 246–251.
- Bóhm T. – Ujváry I. 2008. Az irreguláris fonáció mint egyéni hangjellemző a magyar beszédben. *Beszédkutatás* 2008: 108–120.
- Csákvári J. 2013. *Intellektuális képességszavar és szociális kogníció: A szociális kogníció néhány aspektusának vizsgálata tekintetkövetéses módszerrel mérsékelt intellektuális képességszavarral élő felnőttek körében*. PhD-értekezés. Budapest: ELTE PPK.
- Csákvári J. – Mészáros A. 2012. Értelmi fogyatékos (intellektuális képességszavarral élő) gyermekek, tanulók komplex vizsgálatának diagnosztikus protokollja. In: Torda Á. (szerk.) *Diagnosztikai kézikönyv*. Budapest: Educatio. 1–83.
- Fónagy I. – Magdics K. 1967. *A magyar beszéd dallama*. Budapest: Akadémiai Kiadó.
- Frint T. 1982. Az értelmi fogyatékosok beszédzavarai (dyslogia). In: Frint T. – Surján L. (szerk.) *A hangképzés és zavarai, beszédzavarok. Foniátria*. 2. bővített kiadás. Budapest: Medicina Könyvkiadó. 235–237.
- Gósy M. 1993. A kiegészítő kérdés dallamváltozása. *Magyar Nyelvőr* 117(4): 443–447.
- Gósy M. 2004. *Fonetika, a beszéd tudománya*. Budapest: Osiris Kiadó.
- Gósy, M. – Terken, J. 1994. Question marking in Hungarian: Timing and height of pitch peaks. *Journal of Phonetics* 22: 269–281.
- Hatos Gy. 2000. Az értelmileg akadályozott gyermekek az óvodában és az iskolában. In: Illyés S. (szerk.) *Gyógypedagógiai alapismeretek*. Budapest: ELTE BGGYFK. 409–428.
- Hodapp, R. M. – Dyckens, E. M. 2003. Mental Retardation (Intellectual Disabilities). In: Mash, E. J. – Barkley, R. A. (szerk.) *Child Psychopathology*. 2nd ed. New York: The Guilford Press. 486–519.
- ICD-11 = International Classification of Diseases. 11th Revision. The global standard for diagnostic health information. 2018. Elérhető: <https://icd.who.int/en/> (letöltés ideje: 2019. 05. 29.)
- Jankovics J. 2017. *Intellektuális képességszavarral élő személyek beszédének néhány szupraszegmentális jellemzője*. Elhangzott: Magyar Fonetikai, Foniátriai és Logopédiai Társaság 2017. Kongresszus. ELTE BGGYFK, Budapest, 2017. június 17.
- Jankovics J. 2019. Alaphangjellemzők vizsgálata enyhe és közepes súlyos értelmi fogyatékkal élő felnőttek beszédében. *Beszédkutatás* 2019: 314–330.
- Kent, R. D. – Vorperian, H. K. 2013. Speech impairment in Down syndrome: A review. *Journal of Speech, Language and Hearing Research* 56(1): 178–210.
- Kiefer F. 2000. Jelentésmélet. Budapest: Corvina Kiadó.
- Kugler N. 2000. A mondattan általános kérdései. In: Keszler B. (szerk.) *Magyar grammatika*. Budapest: Nemzeti Tankönyvkiadó. 369–393.
- Lányiné Engelmayer Á. 2012. *Intellektuális képességszavar és pszichés fejlődés*. Budapest: Medicina Könyvkiadó.
- Lee, M. T. – Thorpe, J. – Verhoeven, J. 2009. Intonation and phonation in young adults with Down syndrome. *Journal of Voice* 23(1): 82–87.

- Lezcano, A. – Troncaso, M. V. 1999. A Down-kórosok nyelvhasználatának értékelése és a fejlesztés lehetőségei. *Gyógypedagógiai Szemle* 27(3): 187–196.
- Lukács A. – Kas B. 2014. Nyelvvelsajátítás és értelmi fogyatékoság. In: Pléh Cs. – Lukács Á. (szerk.) *Pszicholingvisztika 1 – 2. Magyar pszicholingvisztikai kézikönyv*. Budapest: Akadémiai Kiadó. 1383–1404.
- Macher M. 2007. Tanulásban akadályozott gyermekek aktív szókincsének vizsgálata. *Beszédkutató 2007*: 142–151.
- Markó A. 2007. Kérdő funkciójú hanglejtésformák a spontán beszédben. *Beszédkutató 2007*: 59–74.
- Markó A. 2013. *Az irreguláris zöngé funkciói a magyar beszédben*. Budapest: Eötvös Kiadó. Elérhető: <http://www.eltereader.hu/kiadvanyok/az-irregularis-zongge-funkcioi-a-magyar-beszedben/> (letöltés ideje: 2019. 03. 17.)
- Markó A. 2015. *A spontán beszéd prozódiai szerkezete. Időzítés és beszéd dallam*. Budapest: Akadémiai Kiadó.
- Markó A. 2017. Hangtan. In: Tolcsvai Nagy G. (főszerk.) *Nyelvtan*. Budapest: Osiris Kiadó. 75–206.
- Mérei F. 1980. Az értelmi fogyatékosok beszédhibái. In: Kovács E. – Mérei F. (összeáll.) *Tanulmányok a logopédia köréből*. Kézirat. Budapest: Nemzeti Tankönyvkiadó. 194–209.
- Olaszy G. 1995. A kérés, a figyelmeztetés, a felszólítás és a kérdés prozódiaja a kijelentő mondat tükrében. *Beszédkutató 2014*: 46–61.
- Olaszy G. 2002. A magyar kérdés dallamformáinak és intenzitás szerkezeteinek fonetikai vizsgálata. *Beszédkutató 2002*: 83–99.
- Olaszy G. 2010. A beszéd szupraszegmentális szerkezete. In: Németh Géza – Olaszy Gábor (szerk.) *A magyar beszéd: Beszéd kutatás, beszéd technológia, beszéd információs rendszerek*. Budapest: Akadémiai Kiadó. 171–205.
- Seiffpanahi, S. – Bakhtiar, M. – Salmalian, T. 2011. Objective vocal parameters in Farsi-speaking adults with Down syndrome. *Folia Phoniatrica et Logopaedica* 63(2): 72–76.
- Stivers, T. 2010. An overview of the question-response system in American English conversation. *Journal of Pragmatics* 42(10): 2772–2781.
- Stoel-Gammon, C. 2001. Down syndrome phonology: Developmental patterns and intervention strategies. *Down Syndrome Research and Practice* 7(3): 93–100.
- Tóth A. 2017. *A spontán beszéd a nem és az életkor függvényében gyermek- és fiatal felnőttkorban*. PhD-értekezés. Budapest: ELTE BTK.
- Varga L. 1994. A hanglejtés. In: Kiefer F. (szerk.) *Strukturális magyar nyelv 2. Fonológia*. Budapest: Akadémiai Kiadó. 468–549.
2011. évi CXCV. törvény a nemzeti köznevelésről. Elérhető: <https://net.jogtar.hu/jogszabaly?docid=A1100190.TV> (letöltés ideje: 2018. 01. 30.)

Határon túli magyar egyetemisták nyelvhasználatának megítélése

Kiss Anita

Debreceni Egyetem BTK Nyelvtudományok Doktori Iskola
kissanita0728@gmail.com

Kivonat: Tanulmányom a nyelvi változatossággal kapcsolatos attitűdöket ismerteti. Határon túli magyar, illetve magyarországi származású egyetemisták vélekedései alapján mutatom be a kétnyelvű és egynyelvű magyar nyelvhasználat közötti különbségeket, valamint a különböző nyelvváltozatokhoz kapcsolódó értékítéleteket. A kutatás során arra a kérdésre kerestem a választ, hogy a tanulmányi céluk miatt mobilis hallgatók nyelvhasználatára milyen hatással van a kétnyelvű magyar közegből az egynyelvű magyar közegbe való beilleszkedés, és ez milyen nyelvi attitűdöket eredményez. A kétnyelvű magyar beszélők nyelvhasználatát eltér az egynyelvű magyarokétól annak okán, hogy a mindennapos kommunikációjukat egyszerre két vagy több nyelv, és annak változatai is befolyásolják. Ebből kifolyólag azt is megvizsgálom, hogy a magyarországi származású egyetemisták egy csoportja, akik a mindennapok során gyakran érintkeznek külföldi magyar hallgatókkal, hogyan vélekednek a határon túli magyarok nyelvhasználatáról, és ők milyen különbségeket fedeznek fel a kétfajta magyar nyelvhasználat között.

1 Bevezetés

A Magyarországgal szomszédos területeken élő kisebbségi magyar közösségeket határok választják el az anyaországtól. Ennek következményeképpen egyrészt korlátozódtott a határon túli magyarok kapcsolata az egynyelvű magyarokkal, másrészt a többségi állam nyelve a kisebbségi magyarok mindennapos kommunikációjának fontos meghatározója lett. Ezek a tényezők hatással vannak a magyar nyelvhasználatra, és különbségeket eredményeznek a kisebbségi magyar, valamint a magyarországi nyelvhasználat között. Ennek egyik jellegzetessége, hogy a határon túli magyar nyelvhasználatban erőteljesebben jelennek meg nyelvjárási sajátosságok, a Magyarországtól távolabbi területekre lassabban jutnak el a nyelvi újítások, és tovább fennmaradnak az archaikus, valamint nyelvjárási formák. A másik jelentős tényező a kétnyelvű helyzetből adódik. A határon túli magyarok mindennapjaiban az anyanyelvük mellett, szükségszerűen jelen van a többségi állam nyelve is, amelyet bizonyos helyzetekben használnak. To-

Kiss Anita: Határon túli magyar egyetemisták nyelvhasználatának megítélése. In Váradi Tamás (sorozatszerkesztő), Ludányi Zsófia, Grácz Tekla Etelka (szerkesztő): *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2019. XIII. Alkalmazott Nyelvészeti Doktoranduszkonferencia*. Budapest: MTA Nyelvtudományi Intézet. 2019. 92–103. DOI: 10.18135/Alknyelvdok.2019.13.7

vábbá a nyelvhasználati színterek nagy részén, ahol Magyarországon a köznyelv használatos, kisebbségi területeken leginkább az államnyelv használata jellemző (Csernicskó–Szabó Mihály 2010: 181–182).

Jelen munkában külhoni magyar és magyarországi származású egyetemisták vélekedései alapján szeretném bemutatni a határon túli magyar és a magyarországi nyelvhasználat közötti eltéréseket. Továbbá az imént bemutatott tényezők figyelembevételével (nyelvjárási vonatkozások, kontaktushatások) vizsgálom a határon túli magyar és magyarországi származású egyetemisták nyelvi attitűdjét.

1.1 Nyelvi attitűd fogalma

Az első nyelvi attitűddel kapcsolatos vizsgálatok a 70-es években kezdődtek el. Nyelvi attitűd alatt a nyelvvel kapcsolatos tudást, vélekedést értjük, amelyben megmutatkoznak egyes emberek vagy embercsoportok viszonya és értékítéletük bizonyos nyelvekre, nyelvváltozatokra, kiejtésváltozatokra, illetve bármely nyelvi jelenségre. Az attitűdök nagyon pozitív vélekedésektől egészen a teljes elutasításig terjedhetnek, amelyek egyes nyelvváltozatok „helyességéről”, értékességéről, beszélői személyes tulajdonságairól alkotott ítéletekben mutatkoznak meg (Grin 2013: 682; Trudgill 1997: 58).

A nyelvi attitűdök elsősorban nem nyelvi, hanem társadalmi eredetűek, tehát nem nyelvi vagy esztétikai minőséget tükröznek, hanem társadalmi szokásokat, viselkedési szabályokat, ide vonatkoztatva az előítéleteket is. A szociolingvisztikai kutatások rávilágítottak, hogy a nyelvi attitűdök a már korábban létrejött nem nyelvi sztereotípiák nyelvi igazolásaként, megerősítéseként alakultak ki (Kiss 2002: 136). Tehát a nyelvi attitűdök befolyással lehetnek a nyelvi változásra, viselkedésre, és nyelvi bizonytalanságot eredményezhetnek. Nyelvi bizonytalanság alatt azt értjük, hogy a beszélők ellenszenvet mutatnak a saját nyelvváltozatuk iránt, vagyis megkérdőjelezzik a nyelvváltozatuk „helyességét”, és megpróbálnak egy magasabb státuszú beszédformát elsajátítani, amely az egyén részéről túlhelyesbítéshez, a társadalmi csoport részéről normatúlteljesítéshez vezethet (Trudgill 1997: 58).

Bizonyos helyzetekben viszont a stigmatizált nyelvváltozatoknak, valamint nyelvi formáknak is van presztízszük, amennyiben egy adott közösség tagjainál pozitív értéket hordoznak, elsősorban az összetartozást, a közösségvállalást mutatják, ami arra ösztönzi a nyelvhasználókat, hogy továbbra is használják azokat a nyelvi formákat. Ezt nevezzük rejtett presztízsnak. Ide vonatkozóan megemlíteném Peter Trudgill norwichi nyelvhasználatról végzett kutatása, amelynek eredményeként leírta a rejtett presztízsz fogalmát. Megállapította, hogy a sztenderd változat magas társadalmi tekintélye mellett, a megvizsgált közösségek, a saját, nemsztenderd nyelvváltozataiknak is presztízst tulajdonítanak, és kötődnek a saját nyelvváltozatukhoz, akkor is, ha amiatt érte már őket negatív megkülönböztetés (Trudgill 1974).

1.2 Határon túli magyar egyetemisták nyelvi attitűdjével foglalkozó kutatások

Nyelvi attitűdöt feltáró kutatásokat végeztek kárpátaljai, romániai, szerbiai és szlovákiai magyar egyetemi/főiskolai hallgatók körében.

T. Károlyi (2002) Kárpátalján, Beregszászban magyar nyelvű képzésben részt vevő tanító szakos hallgatók nyelvi tudatosságát és attitűdjét mérte fel, amelyben azt vizsgálta, hogy a pályakezdő tanítóknak milyen ismereteik vannak az anyanyelvről, annak

változatairól és funkcióiról, tudják-e, hogy milyen különbségek, illetve egyezések vannak a magyarországi magyar és a határon túli magyar nyelvhasználat között. A válaszadók többsége a kárpátaljai magyar nyelv dialektális jellegét hangsúlyozta, és csak 17,5%-a említette meg a másodnyelvi hatásokat. Az eredmények alapján kiderült, hogy az adatközlők a saját nyelvváltozatukhoz inkább pozitívan, míg a magyarországihoz inkább negatívan viszonyulnak, és a többség elfogadhatónak tartja mind a nyelvjárási sajátosságokat, mind a kétnyelvűségi hatásokat. A kommunikáció során a többség a kiejtésre, a hangsúlyozásra és a hangok helyes képzésére ügyel inkább, és csak 9%-uk nevezte meg direkt módon, hogy odafigyel beszéd közben arra, hogy ne használjon nyelvjárási elemeket.

Sándor (2009) Szlovákiában elsőéves magyar szakos nyitrai egyetemisták körében végzett nyelvjárással kapcsolatos attitűdvizsgálatot. A kérdőívet 64-en töltötték ki, ebből az adatközlők több mint kétharmada falun él, és majdnem egyharmada városban. A kitöltők 81,2%-a saját bevallása szerint beszél valamilyen nyelvjárást, és az eredményekből az is kiderült, hogy a szlovákiai magyarok által lakott területeken a nyelvjárást nemcsak a falvakra, hanem a városokra is jellemző. A válaszadók többségét nem érte negatív megkülönböztetés a nyelvhasználat miatt, és szerintük szükség is van a nyelvjárássokra. Egy hasonló vizsgálatban, amit szintén nyitrai elsőéves egyetemisták körében végeztek, a nyelvváltozatok megítélésével kapcsolatban kiderült, hogy a többség negatívan ítéli meg a budapesti, illetve a magyarországi vidéki nyelvváltozatokat, a határon túli magyar nyelvhasználatról pedig semleges véleményeket (Presinszky 2009).

Fazekas (2016) romániai magyar egyetemisták körében azt vizsgálta, hogyan viszonyulnak a hallgatók a nyelvjárássokhoz, a román és a kisebbségi nyelvhez. A nyelvi attitűdöket illetően a válaszadók többsége szerint Erdélyben beszélnek a legszebben magyarul, tehát pozitívan viszonyulnak az anyanyelvükhöz, a román nyelvhez való hozzáállásuk viszont nem mutatott egyöntetű eredményeket, az adatközlők egy része elengedhetetlennek tartja az államnyelv ismeretét, míg mások nem beszélnek megfelelőképp a román nyelvet, amiért az oktatási rendszert teszik felelőssé.

Csincsik (2011) vizsgálatában összehasonlította a Magyarországon, illetve Szerbiában tanuló vajdasági hallgatók nyelvi attitűdjét. A kutatásba harminc szerb, illetve harminc magyar környezetben tanuló hallgatót vontak be. Az eredmények a magyar nyelvhez való pozitív hozzáállást mutatták. Csoportidentitás szempontjából az adatközlők többsége a Vajdasághoz kötődött, és pozitívan értékelték az ottani nyelvváltozatokat. Mindkét csoportnál jellemző volt a vajdasági nyelvváltozatokhoz való pozitívabb viszonyulás a magyarországiakkal szemben.

Az eddigi kutatások azt mutatják, hogy a hallgatók pozitívan ítélik meg a saját településük nyelvváltozatait, és a saját közösségük nyelvhasználatát preferálják a magyarországi nyelvhasználattal szemben. Az említett kutatások majdnem mindegyike határon túli magyar településen felsőoktatási tanulmányokat folytató hallgatók nyelvi attitűdjét vizsgálta. Saját kutatásomban a kisebbségi magyar nyelvterületről részben már eltávolodott, az anyaországi elhelyezkedést célba vett hallgatók nyelvváltozatokhoz való viszonyát vizsgálom a saját településük, illetve a magyarországi nyelvhasználat vonatkozásában.

1.3 Kutatási hipotézisek

1. A kétnyelvű közegben élő kárpátaljai magyarok a mindennapos kommunikáció során három nyelv és annak változatai közül választhatnak. Bizonyos helyzetekben a magyart, máskor az ukrán vagy az orosz nyelvet használják, esetenként pedig vegyesen a magyart és az ukránt (Csernicskó 2004). A kétnyelvű helyzet következtében a határon túli magyar és a magyarországi nyelvváltozatok között különbségek figyelhetők meg. Feltételeztem, hogy a kétnyelvű közegből kilépő külhoni magyar fiatalok észlelik az eltéréseket, és felismerik, hogy ezek a különbségek milyen sajátosságokban mutatkoznak meg.

2. Trudgill (1974) eredményei alapján feltételeztem, hogy a határon túli magyar adatközlők pozitív értékeket társítanak azokhoz a nyelvváltozatokhoz, amelyeket a szülőföldjükön elsődlegesen sajátítottak el, ugyanakkor az új nyelvi környezetben próbálják a nyelvhasználatukat a magyarországiakéhoz igazítani.

2 Módszertan

A kutatásba elsősorban határon túli magyar hallgatókat vontam be, akik a középiskola/gimnázium befejezéséig szülőföldjükön éltek, majd felsőoktatási tanulmányaik elkezdése miatt Budapestre költöztek. Kontrollcsoportként magyarországi egyetemistákkal is beszélgetést folytattam, akik Magyarország különböző településeiről származnak, jelenleg Budapesten folytatják felsőoktatási tanulmányaikat, és a határon túli magyar hallgatókhoz hasonlóan a Budapesti ELTE Márton Áron Szakkollégiumban laknak. A magyarországi adatközlők bevonása egyrészt azzal a céllal történt, hogy megerősítsék vagy cáfolják az interjú azon részeinek eredményeit, amelyben a határon túli magyar adatközlők véleményét kértem arról, hogy a saját észrevételeik szerint a magyarországiak hogyan viszonyulnak a határon túli magyarok nyelvhasználatához. Másrészt, hogy összehasonlíthassam a két csoport véleményét arra vonatkozóan, hogy milyen különbségeket fedeznek fel a határon túli magyar és a magyarországi nyelvhasználat között. Mindehhez azért választottam a Márton Áron Szakkollégium magyar származású hallgatóit, mivel ők napi rendszerességgel kapcsolatban állnak az ott lakó határon túli magyar hallgatókkal, és a gyakori interakciók során megtapasztalhatják a külhoni magyar fiatal hallgatók nyelvhasználati sajátosságait, szokásait, nyelvi alkalmazkodásukat.

A vizsgálat során a következő kérdésekre kerestem a választ:

1. Okozott-e a határon túli magyar fiatalok számára nyelvhasználati nehézségeket a kétnyelvű magyar közegből az egynyelvű magyar közegbe való beilleszkedés?
2. Milyen különbségeket vesznek észre a megkérdezett határon túli magyar és a magyarországi származású adatközlők a kétféle magyar nyelvhasználat között?
3. Milyen ismereteik vannak az adatközlőknek a nyelvi változatosságról, és hogyan ítéli meg a határon túli magyar, valamint a magyarországi csoport a kétféle magyar nyelvhasználatot?

A vizsgálatot a Budapesti (ELTE) Márton Áron Szakkollégiumnak egy olyan emeletén végeztem, ahol különböző határon túli magyar területekről érkezett egyetemisták

mellett magyarországi származású hallgatók is laknak. Félig strukturált interjúk kérdésekkel dolgoztam, amelyek során fény derülhetett az adatközlők szubjektív tapasztalataira. Az egyéni interjúk alkalmával összesen 25 adatközlővel folytattam beszélgetést, közülük 17 hallgató valamely határon túli magyar nyelvterületről érkezett, további 8 hallgató pedig magyarországi származású volt. A beszélgetések az adatközlők beleegyezésével hangfelvételes rögzítésre kerültek.

Az interjúk vizsgálat egy nyelvi attitűddel kapcsolatos kérdőíves felmérés kiegészítéseként készült. A félig strukturált interjú a következő témákat foglalta magában:

1. A határon túli, kisebbségben élő magyarok viszonya a többségi (ukrán, román, szerb, szlovák) nemzethez, illetve az egynyelvű magyarokhoz.
2. A szülőföldön maradás, valamint az anyaország irányába történő mobilitás motivációi.
3. Az adatközlők anyanyelvhasználatra vonatkozó tapasztalatai kisebbségi, illetve egynyelvű magyar nyelvterületen.
4. Kétnyelvűségi hatásokhoz, nyelvjárási jelenségekhez fűződő nyelvi attitűdök.
5. Nyelvi előítéletekkel és nyelvhasználati nehézségekkel kapcsolatos tapasztalatok.

A jelen tanulmányban a kutatási kérdésekhez és hipotézisekhez igazodva csak a negyedik, illetve ötödik ponthoz tartozó válaszok eredményeit ismertetem. Az interjúk beszélgetések névtelenek, a példaként felhozott szövegrészek végén, zárójelben feltüntettem az adatközlő születési évét, valamint származási helyét (ország, település neve), például: 1996, Ukrajna, Beregszász.

3 Eredmények

Szociológiai jellemzőket figyelembe véve a megkérdezettek 60%-a 1995 és 1997 között született. A legtöbben ukrainai/kárpátaljai (36%), illetve magyarországi (32%) származásúak. Ukrajna mellett azért tartottam fontosnak a kárpátaljai jelzést is megjelölni, mivel az interjúk során a legtöbb adatközlő arra a kérdésre, hogy: „hol született?”, azt válaszolta, hogy Kárpátalján, ami a kárpátaljai magyarok számára egy fontos nemzeti identitásjelző tényező (vö. Cserniczkó–Soós 2002: 130–131; Molnár–Orosz 2007: 192). Kisebbségi arányban Szerbiából, Romániából és Szlovákiából érkeztek az adatközlők (1. táblázat). A megkérdezettek tanulmányaira vonatkozó adatok alapján a többség (60%) az ELTE hallgatója (2. táblázat). Az egyetemisták által választott szakok nagyon sokszínűek, ezért tudományterületek szerint csoportosítottam azokat. A legtöbb adatközlő bölcsészettudományi (52%) tanulmányokat folytat (3. táblázat).

Származási hely	Fő
Szerbia	4
Szlovákia	1
Románia	3
Ukrajna	9
Magyarország	8

1. táblázat. Az adatközlők származási helye

Felsőoktatási intézmény	Fő
-------------------------	----

Budapesti Gazdasági Egyetem	1
ELTE	15
Óbudai Egyetem	3
Gábor Dénes Főiskola	1
Károli Gáspár Református Egyetem	1
Széchenyi István Egyetem	1
Pázmány Péter Katolikus Egyetem	1
Semmelweis Egyetem	2

2. táblázat. Az adatközlők felsőoktatási intézménye

Tudományterületek	Fő
gazdaságtudományok	3
természettudományok	2
bölcsészettudományok	13
orvostudományok	2
műszaki tudományok	4
társadalomtudományok	1

3. táblázat. Az adatközlők tanulmányainak besorolása tudományterületek szerint

A rögzített interjúszövegek lejegyzését követően kategorizációs módszerrel végeztem az elemzést. Az interjú beszélgetések egyes részei kerültek elemzésre oly módon, hogy csoportosítottam a kutatási kérdések szempontjából lényeges válaszokat. A határon túli magyar nyelvhasználat és a magyarországi nyelvhasználat közötti eltérések vizsgálatánál rendszereztem mind a külföldi magyar hallgatók, mind a magyarországi egyetemisták válaszait, majd összevettem a két adatközlői csoport megállapításaiból származó eredményeket. A kölcsönzőket tartalmazó megnyilatkozásoknál szintén rendszereztem, hogy az adott kölcsönző mely határon túli magyar nyelvterületről érkezett egyetemistától hangzott el. A nyelvjárási jelenségek megítélésénél összesítve mutatom be azokat a vélekedéseket, észrevételeket, amelyeket az adatközlők többsége megemlített a beszélgetések során.

A határon túli magyar adatközlőket arról kérdeztem, előfordult-e, hogy a magyarországiakkal való kommunikáció során valamilyen nehézségbe ütköztek a nyelvhasználatuk miatt. A válaszok alapján egyik megkérdezettnek sem okozott nyelvhasználati nehézséget a többnyelvű közegből az egynyelvű magyar környezetbe való beilleszkedés. Azt azonban minden válaszadó kiemelte, hogy érzékelik a határon túli magyar és a magyarországi nyelvhasználat közötti különbségeket, de ők ezt nem nyelvhasználati nehézségként értelmezik [1].

- [1] TM: *Kárpátaljai magyarként előfordult már veled, hogy valamilyen nehézségbe ütköztél a nyelvhasználatod miatt egy itteni, magyarországi beszélővel?*

AK: *Igazából nem, hanem vicces szituációkat eredményez, mert hiába magyarul beszélünk otthon is, teljesen más, tehát ugye mi az é betűt nem úgy mondjuk, mint itt, meg az ó is nekünk ilyen óu, kicsit olyan hangzása van, és ha nagyon belelendülök, nem azt mondom, hogy pad, hanem lóuca, de ilyen igazi kárpátaljai lóuca, szóval ilyenek voltak. Például dolgoztam és akkor ott is ilyen viccesen mondták, nem ilyen kötekedően, csak mindig megjegyezték, például nem azt mondtam, hogy felvizezték, hanem felvizezték,*

merthogy az itt mást jelent, de nálunk így mondják (1996, Ukrajna, Beregszász).

A megkérdezett külhoni magyar hallgatók többsége elsősorban a magyar nyelv határon túli változataiban megjelenő kölcsönszók használatában látja a különbséget. A beszélgetések során a következő kölcsönszókat említették:

Kárpátaljai magyar adatközlők: *marsutka* 'kisbusz', *bulocska* 'zsömle', *szvetafor* 'jelzőlámpa', *dekret* 'szülési szabadság', *cirkula* 'körfűrész'.

Szerbiai magyar adatközlők: *ájde* 'oké, igen, gyerünk', *kao* 'mintha', *majica* 'póló', *trénerka* 'tréningruha', *licsna* 'személyi igazolvány', *szemafor* 'jelzőlámpa', *patika* 'tornacipő', *patent* 'rotring', *pruszlik* 'mellény', *bánka* 'befőttesüveg'.

Szlovákiai magyar adatközlők: *tyepláki* 'tréningruha', *párki* 'virslis', *nanuk* 'jégkrém', *zsuvi/zsuvacska* 'rágógumi', *intril/internát* 'koli', *horcsica* 'mustár'.

Romániai magyar adatközlők: *szemafor* 'jelzőlámpa', *murok* 'répa', *buletin* 'személyi igazolvány', *pix* 'golyóstoll', *pungam* 'zacskó', *plásza* 'szatyor'.

A tizenhét megkérdezett határon túli magyar hallgatóból tizenkettő fel tudott eleveníteni olyan esetet, amikor a magyarországiaknak feltűnt, hogy ők valamely határon túli magyarlakta területről érkeztek, és megjegyezték, hogy ez megmutatkozik a nyelvhasználatukban, de ezeket a megjegyzéseket a megkérdezett külhoni magyar hallgatók nem érzik sértőnek.

A magyarországi adatközlőktől is megkérdeztem, hogy mi a véleményük a külhoni magyarok nyelvhasználatáról, és megfigyeltek-e valamilyen különbséget a magyarországi és a határon túli magyar nyelvhasználat között. Szerintük is észrevehető, hogy a határon túli magyarok másképp beszélnek, mint a magyarországiak, de a többségük szerint ezzel nincs semmi probléma, mivel tudják, hogy az egynyelvű magyarok is különbözőképpen beszélnek attól függően, hogy Magyarországnak mely részéről származnak [2].

[2] TM: *Észrevetél valamilyen különbséget a magyarországi magyarok és a külhoni magyarok nyelvhasználatában?*

AK: *Talán az, ahogy kiejtik a szavakat, de hogy igazából ez ugyanolyan, hogy Magyarországon belül is van olyan, hogy Mosonmagyaróváron másképp beszélnek, mint Szegeden, tehát ezt inkább csak úgy mondanám, hogy mint tájegységbeli különbség* (1999, Magyarország, Pély).

A határon túli magyar és a magyarországi nyelvhasználat közötti különbség a magyarországiak többsége szerint a nyelvjárási jelenségek használatában mutatkozik meg, itt kiemelték a tájszólást, ami egyes adatközlők szerint a magyarországiaknak idegen, vagy csak az idősek (nagyapák) beszédére jellemző. Gyakran előforduló példaként felhozták, hogy a külhoni magyarok a *hűtő* helyett azt mondják *fridszider* [3].

[3] AK: *Az itteni legjobb barátom, ő szabadkai, és a szobatársa pedig erdélyi, és én nagyon érdekesnek tartom azt, hogy például a szabadkai barátom nem az, hogy ilyen régiesen beszél, de hogy nagyon sok mindenre olyan szavakat használ, mint mondjuk a nagymamám. Szóval a hűtő helyett fridszider... Akkor például az erdélyi barátomja nagyon sok mindent így románul mond először, és utána kijavítja magyarra. Például a répa helyett murok, azt hiszem. Viszont most már, hogy így egyre többet velünk van egyre jobban érzem, hogy nemhogy veszíti el, de így öntudatlanul átállt ránk, és amikor mindig jön vissza, újra, de állítólag ez velem is így van, pedig aztán*

Szigligeten nincs tájszólás, de mi állítólag a t-t nagyon megnyomjuk (1996, Magyarország, Szigliget).

Ezen kívül jellemző szerintük bizonyos hangtani vonatkozású nyelvjárási jelenségek előfordulása a beszédükben, például az *eredj/menj* helyett azt mondják, *eriggy* [4].

[4] TM: *Észrevettél valamilyen különbséget a magyarországi magyarok és a külhoni magyarok nyelvhasználatában között?*

AK: *Hát egy picit lehet szerintem észrevenni, vagyis az én szintemen lakókon.*

TM: *És ezt hogy érted?*

AK: *Hogy hűtő. Nem hűtő, hanem hűtő. Néhány szóhasználat, ami eltér, például: ami nálunk a vágódeszka, az náluk a lapító, lapítónak mondja az egyik ismerősöm. De szerintem ezek olyan kis aranyos dolgok, mármint, hogy van pár olyan szó, amit így beszéd közben hallok, és elgondolkozom rajta, hogy vajon mit jelenthet, de valószínűleg úgysem azt jelenti, és akkor megyek és megkérdezem, és azért elég sok új szót tudok tőlük tanulni. Viszont a legjobb barátnőm, ő Kárpátaljáról származik, és talán annyit szokott mondani, hogy menj helyett, hogy eriggy vagy talán 2-3 ilyen szót használ csak, ami eltér az enyémtől, szóval annyira nem veszek észre különbségeket. Nagyon szépen tudnak beszélni magyarul, nagyon szépen tudnak mindent megfogalmazni, valamikor sokkal tisztábban ejtik a szavakat, mint mi (1997, Magyarország, Diósd).*

A megkérdezett magyarországi adatközlők maguk is úgy vélekedtek, előfordul, hogy megjegyzést tesznek arra, ha egy külhoni magyar által használt szó jelentését nem ismerik, vagy eltérő hangalakban ejtik ki, és többen említették, hogy időnként viccelnek egymás nyelvhasználatával, azt azonban minden megkérdezett kihangsúlyozta, hogy nem rosszindulatúan teszi.

Összességében a magyarországi adatközlők nagyjából pozitívan vélekedtek azokról a jelenségekről, amelyeket leginkább csak a határon túli magyar társaik nyelvhasználatában fedeznek fel, és többen megjegyezték, hogy sok új szót is tanulhatnak tőlük. Megemlítették például a kárpátaljai magyarok által gyakran használt szót, ami a *bulocska* 'zsömle', illetve a szlovákiai magyarok által használt *zsuvi* 'rágógumi' szót. Többen is megjegyezték, hogy észrevételeik szerint minél több ideje élnek Magyarországon, annál kevésbé jellemzők a beszédükre azok a sajátosságok, amelyeket otthonról hoztak magukkal.

Amikor azt a kérdést tettem fel a határon túli magyar hallgatóknak, hogy: „*egyetértene-e azzal, ha valaki azt mondaná nekik, hogy határon túli magyarként nem beszélnek olyan „jól/helyesen” magyarul, mint a magyarországiak?*”, ezzel egy-két adatközlő kivételével senki sem értett egyet. E kérdésnél szinte minden megkérdezett a nyelvjárási beszédre gondolt, mivel a beszámolók alapján jórészt a tájszólásukra kapták a legtöbb megjegyzést. A nyelvjárási beszédet azonban ők nem vélik helytelennek, és megjegyezték, hogy ez a magyarországiakra is jellemző [5].

[5] TM: *Ha valakitől azt hallanád, hogy a kárpátaljai magyarok nem beszélnek olyan „jól/helyesen” magyarul, mint a magyarországiak, akkor ezzel egyetérteneél?*

AK: *Nem értenék egyet.*

TM: *Ezt meg tudod magyarázni?*

AK: *Magyar iskolába jártunk mi is, magyarul tanultunk, szóval nincs különbség beszéd között, nyilván ugyebár a nyelvhasználat az eltér falunként meg régióként, viszont ha például megfigyeled, akkor a budapestiek sem úgy beszélnek, mint például a győriek (1996, Ukrajna, Zápszony).*

Egy-egy adatközlő kifejezetten felháborítónak tartja, ha valakitől azt hallja, hogy ők nem beszélnek „helyesen” magyarul, sőt úgy gondolják, hogy „szebben” beszélnek, mint a magyarországiak [6]

- [6] AK: *Hát én ki szoktam akadni, mikor mondják, hogy az ó betűt milyen hosszasan ejtjük, és hogy milyen fura, és akkor mindig kiakadok.*
 TM: *Tudsz egy példát erre?*
 AK: *Amikor dolgoztam, és akkor mondtam, hogy milyen jó ez a..., és akkor mondták, hogy olyan furán ejtem az ó-t, hogy ők ezt nem úgy ejtik. Én meg mondtam, hogy pont ti ejtitek rosszul, mert mondom nem lehet észrevenni, hogy rövid ó-t vagy hosszú ó-t mondtok a szóba, ezen ki szoktam akadni, mert szerintem teljesen jól beszélünk (1996, Ukrajna, Tiszapéterfalva).*

Egy-két kivételes válasz is adódott, egy szerbiai magyar adatközlő felidézett egy esetet, amikor egy magyarországi szaktársa megbántotta egy őt sértő kijelentéssel, és ez az interjúalanyt el is bizonytalanította kissé [7].

- [7] AK: *Egy szaktársam volt, aki szőrszálhasogató.*
 TM: *És ő mit mondott?*
 AK: *Hogy nem is tudom, hogy nem beszélünk helyesen, mert van még egy szlovák gyerek, és akkor, hogy mi igazából nagyon hasonlóan beszélünk, de mindketten helytelenül.*
 TM: *És ezzel te egyetértesz?*
 AK: *Hát nem tudom igazából, lehet, hogy persze vannak hibáink, mert hogy nem tudom pontosan vagy valami, de hát igazából szerintem többnyire jól beszélem (1997, Szerbia, Szabadka).*

Továbbá egy kárpátaljai adatközlő részben egyetértett az állítással, mert szerinte az ukrán nyelvből átvett kölcsönszók, nem képezik a magyar nyelv, a magyar szókinés részét [8].

- [8] AK: *Hát abból a szempontból, hogy az a szó nem létezik a magyar nyelvben, igen. Tehát tényleg az, hogy szosziszki ukránul van, és most mondhatom bármilyen magyarosan, nem lesz benne a szótárban, meg akármennyire izé, az soha nem lesz magyar szó, mert ukrán. De hogyha egyszer odajövünk, hogy átveszik rendszeren, följegyzik szótárakba vagy valami, és azt mondják, hogy ez egy helyes magyar szó, akkor igen, de addig nem beszélünk helyesen (1996, Ukrajna, Szürte).*

A nyelvjárási beszédről is kikerem az adatközlők véleményét. Szinte minden válaszban elhangzott, hogy vicces (pozitív értelemben). Sokan elismerték, hogy gyakran tesznek megjegyzéseket a mások beszédében előforduló nyelvjárási jelenségekre, de egyáltalán nem lenézően. A megkérdezettek többsége nem vélekedett negatívan a nyelvjárási beszédről azzal az indokkal, hogy amíg megértik egymás mondanivalóját, addig nem zavaró a számukra [9].

- [9] AK: *Én nagyon nevetek rajtuk, mert vicces szerintem, ahogy beszélnek. Meg a szerbeknél is van egy ilyen érdekes hangsúly, és ha beszél, akkor rögtön*

meg tudod mondani, hogy ő szerb, ő meg szlovák. Annyira nem fura, megértem őket, meg ők is engem, úgyhogy nem zavar (1996, Ukrajna, Tiszapéterfalva).

Egy kárpátaljai magyar adatközlő kihangsúlyozta, hogy szereti, ha valaki nyelvjárásban beszél, mert akkor azt a személyt közelebbinek érzi magához [10].

[10] TM: *Mi a véleményed a nyelvjárásos beszédéről?*

AK: *Először mosolygok, az biztos, mert egyből az jutna eszembe, hogy akkor ő se budapesti, és akkor már olyan rokonléleknek tekinthető. Meg hát nem tudom, szerintem színesíti a magyar nyelvet* (1997, Ukrajna, Beregszász).

A magyarországi adatközlők egy része úgy vélekedett, hogy ők nem beszélnek nyelvjárásban, egy másik része elismerte, hogy előfordul a beszédükben egy-egy nyelvjárás szó, de próbálnak leszokni róla, viszont mások beszédében nem zavarja őket az előfordulásuk [11].

[11] TM: *És akkor inkább pozitív vagy negatív a véleményed a nyelvjárásos beszédéről, szerinted el kellene hagyni mindenkinek?*

AK: *Jaj nem, igazából azt nem mondanám. Én csak azt tudom, meg hát igazából azt tapasztaltam, hogy nálunk otthon odafigyeltek arra, hogy ezt annyira ne lehessen hallani, és én magamnak annyira nem szeretem, hogy használom. De egyáltalán nem zavar, meg igazából sehogy nem hat rám, hogyha más valamilyen nyelvjárással beszél* (1999, Magyarország, Pély).

Volt olyan adatközlő, aki akkor szembesült azzal, hogy valamilyen nyelvjárású jelenséget használ, miután valaki felhívta rá a figyelmét [12].

[12] TM: *És a te településeden van valamilyen jellemző nyelvjárás?*

AK: *Nem igazán, azt szokták mondani, hogy úgy mondjuk, hogy hűttő.*

TM: *Mások mondják nektek?*

AK: *Aha.*

TM: *Neked is mondták már?*

AK: *Volt, volt, hogy mondták, igen.*

TM: *És téged ez hogyan érintett?*

AK: *Ja hát nekem nem tűnt fel, úgyhogy így érdekesnek: komolyan? tényleg? Igen elcsodálkoztam rajta, biztos igazuk van, én nem veszem magamon észre* (1999, Magyarország, Tatabánya).

4. Következtetések

Tanulmányomban különböző határon túli magyar területekről származó egyetemi hallgatók véleményét kértem ki a nyelvi változatosságról. A kutatás tervezésekor felállított hipotéziseket az interjúk alátámasztották. A megkérdezett adatközlők észlelik a magyarországi és a határon túli magyar nyelvhasználat közötti különbségeket (1. hipotézis). A határon túli magyar, és a magyarországi nyelvváltozatok közötti eltéréseket a külhoni magyar adatközlők elsősorban a magyar nyelv határon túli változataiban megjelenő kölcsönszók használatában látják, míg a magyarországi adatközlők főleg a nyelvjárásosságban, tájszóhasználatban, ami szerintük a magyarországiakra sokkal kevésbé jellemző. Alátámasztódott, hogy a megkérdezett határon túli magyar hallgatók

pozitívan vélekednek a saját településük nyelvváltozatairól (2. hipotézis). A határon túli magyarok, illetve a magyarországiak véleménye megegyezett abban, hogy azon határon túli hallgatók nyelvhasználatában, akik újonnan költöztek át Budapestre, sokkal erősebben megfigyelhetők a kétnyelvűségi hatások és a saját településük nyelvjárása, mint azoknál, akik már huzamosabb ideje Budapesten élnek. A legtöbb határon túli magyar hallgató tudatosan alkalmazkodik az új nyelvi környezethez, ami az idő múlásával egyre természetesebbé válik számukra, ugyanakkor nem vélekednek negatívan a saját településük nyelvváltozatairól sem, és hazalátogatáskor visszaváltak az otthoni nyelvhasználatra.

Eredményeimet összevetve a korábbi vonatkozó szakirodalommal megállapítható, hogy a megkérdezett határon túli hallgatók többsége tiszteli annak a közösségnek a nyelvjárását, amelyben felnőtt. Nem próbálják elhagyni azt, csupán igyekeznek alkalmazkodni más nyelvi normákhoz is. A rejtett presztízs tehát nyilvánvalóan megjelenik. A magyarországiak szerint a nyelvjárásiasság a leginkább szembevetendő különbség a magyarországi és a határon túli magyar nyelvváltozatok között. A nyelvjárási beszédhez való viszonyulásuk többnyire pozitív. Akik rendelkeznek bizonyos fokú jártassággal a nyelvi változatosság kapcsán, ők meglehetősen pozitív jelzőkkel illették a nyelvjárási beszélőket, biztatva őket arra, hogy a birtokukban lévő nyelvváltozatok igen értékesek. Mások semlegesén vélekedtek a nyelvi változatosságról, megjegyezve, hogy amíg megértik egymás mondanivalóját, addig nem zavarja őket a nyelvjárási beszéd.

Köszönetnyilvánítás

A tanulmány alapjául szolgáló interjúk a RECENS [MTA TK „Lendület” RECENS kutatócsoport „Szervezeti kommunikáció, kooperáció és reputáció” (ERC/648693)] kutatócsoport egyik kutatási programjának keretében készültek.

Irodalom

- Csernicskó I. 2004. A magyar nyelv kárpátaljai helyzetéről. *Magyar Tudomány* 4: 473–480. Elérhető: http://real.mtak.hu/25146/1/jav_csernicsko_matud.pdf (letöltés ideje: 2018. 11. 19.)
- Csernicskó I. – Soós K. 2002. Gyorsjelentés–Kárpátalja. In: *Mozaik 2001. Magyar fiatalok a Kárpát-medencében*. Budapest: Nemzeti Ifjúságkutató Intézet. 91–135.
- Csernicskó I. – Szabó Mihály G. 2010. Hátrányból előnyt: a magyar nyelvpolitika és nyelvtervezés kihívásairól. In: Bitskey B. (szerk.) *Határon túli magyarság a 21. században*. Budapest: Köztársasági Elnöki Hivatal. 167–198. Elérhető: http://real.mtak.hu/25271/1/alcim_pdf10416.pdf (letöltés ideje: 2018. 12. 23.)
- Csincsik A. 2011. Vajdasági magyar nyelvhasználat egyetemisták körében. *Iskolakultúra* 2–3: 45–58.
- Fazekas N. 2016. Erdélyi magyar egyetemisták nyelvi és nyelvhasználati attitűdjei. In: Kozmács I. – Vančo I. (szerk.) *Sztenderd–nem sztenderd. Variációk egy nyelv változataira*. Lakitelek: Antológia Kiadó. 55–64.
- Grin, F. 2013. Language Policy, Ideology, and Attitudes. In: Bayley, R. – Richard C. – Ceil L. (szerk.) *The Oxford Handbook of Sociolinguistics*. Oxford: University Press. 629–650.
- Kiss J. 2002. *Társadalom és nyelvhasználat*. Budapest: Nemzeti Tankönyvkiadó.
- Molnár E. – Orosz I. 2007. Kárpát Panel–Kárpátalja. Gyorsjelentés. In: Papp Z. A. – Veres V. (szerk.) *A Kárpát-medencei magyarok társadalmi helyzete és perspektívái. Gyorsjelentés*. Budapest: MTA Etnikai–Nemzeti Kisebbségkutató Intézet. 185–243.

- Presinszky K. 2009. Nyelvi attitűdök vizsgálata a nyitrai magyar egyetemisták körében. In: Borbély A. – Vančoné Kremmer I. – Hattyár H. (szerk.) *Nyelvideológiák, attitűdök és sztereotípiák*. Budapest–Dunaszerdahely–Nyitra: Tinta Könyvkiadó. 241–248.
- Sándor A. 2009. A nyelvjárási attitűd vizsgálata a nyitrai magyar szakos egyetemisták körében. In: Borbély A. – Vančoné Kremmer I. – Hattyár H. (szerk.) *Nyelvideológiák, attitűdök és sztereotípiák*. Budapest–Dunaszerdahely–Nyitra: Tinta Könyvkiadó. 231–239.
- T. Károlyi M. 2002. A nyelvi tudatosság és attitűd vizsgálata a beregszászi főiskola hallgatóinak különböző csoportjaiban. In: Hoffmann I. – Juhász D. – Péntek J. (szerk.) *Hungarológia és dimenzionális nyelvszemlélet*. Debrecen: Debreceni Egyetem Magyar Nyelvtudományi Tanszék. 329–338.
- Trudgill, P. 1974. *The social differentiation of English in Norwich*. Cambridge: Cambridge University Press.
- Trudgill, P. 1997. *Bevezetés a nyelv és társadalom tanulmányozásába*. Szeged: JGYTF.

Terminológiai változtatások az Alaptörvény különböző, angol nyelvű változataiban

Klenk Márk

ELTE BTK Nyelvtudományi Doktori Iskola
klenky1990@gmail.com

Kivonat: Kutatásomban Magyarország Alaptörvényének Preambulum és Alapvetések részeiben fellelhető terminológiai változtatásait gyűjtöm össze, néhányat ezek közül részletesen elemezve is bemutatok. A fordításba történő lektori, ellenőrzői beavatkozások általában minőség-növelő céllal történnek meg. Éppen ezért szükségesnek tartom röviden bemutatni a terminológia és lektorálás minőségbiztosító szerepét a fordításokkal kapcsolatban. Jelen pilot kutatás betekintést nyújt doktori értekezéshez készítendő kutatásomba, melynek célja az Alaptörvény két eltérő hatályállapotú fordításának terminológiai szempontú vizsgálata, annak érdekében, hogy átfogó képet kapjunk a közöttük jelentkező terminológiai változtatások mibenlétéről. A teljes szöveg elemzése után egy kategóriarendszer felállításának segítségével, akár az átváltási műveletek mintájára, megalkotható a terminológiai változtatások profilja is.

1 Bevezetés

Az Alaptörvény megjelenése óta különböző műhelyekben számos fordítás látott napvilágot. Tanulmányomban Magyarország Alaptörvényének – a Preambulum és az Alapvetések fejezeteit forrásul véve és elemezve – két hatályállapotban megjelent angol nyelvű fordítását hasonlítom össze annak érdekében, hogy megvizsgáljam, milyen terminológiai változtatások történtek az angol nyelvű szövegben. Kutatásom célja a két eltérő hatályállapotú fordítás terminológiai szempontú vizsgálata, annak érdekében, hogy átfogó képet kapjunk a közöttük jelentkező terminológiai változtatások mibenlétéről. A kutatás érdekessége, hogy az Alaptörvény szövege több olyan helyen is változott, ahol a magyar szövegben nem történt változtatás, tehát a magyar szöveg módosítása nem indokolta azt. Magyar–olasz viszonylatban az Alaptörvény néhány terminológiai változtatását Novák (2013, 2018) már részletesen bemutatta.

Az egyes terminológiai változtatásokat a fogalmi összehasonlítás módszerével elemezve, következtetéseket vonok le arra nézve, hogy a fordítás minősége mennyiben változott meg ezek hatására. Éppen ezért tanulmányom második fejezetében röviden körül járom a minőség fogalmának kérdéskörét, majd ezt leszűkítve a minőség fogalmát határozom meg a fordítás mint szolgáltatás és a fordítás mint produktum vonatkozásában. Mivel a terminológia fontos szerepet tölt be a szövegek minőségértékelésével

Klenk Márk: Terminológiai változtatások az Alaptörvény különböző, angol nyelvű változataiban. In Várad Tamás (szerkesztő), Ludányi Zsófia, Grácsi Tekla Etelka (szerkesztő): *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2019. XIII. Alkalmazott Nyelvészeti Doktoranduszkonferencia*. Budapest: MTA Nyelvtudományi Intézet. 2019. 104–116. DOI: 10.18135/Alknyelvdok.2019.13.8

kapcsolatban, röviden erre is kitérek. A párhuzamos korpuszból származó terminológiai változtatásokat a 4. fejezetben táblázatos formában mutatom be, majd a terminológiában használt fogalmi összehasonlítás módszerét alkalmazva vizsgálom meg a feltárt terminológiai változtatásokat.

2 A fordítás és a minőség fogalma

Az ISO 9000:2015 szabvány minőségdefiníciója szerint a minőség „Annak mértéke, hogy mennyire teljesíti a saját jellemzők egy csoportja a követelményeket”, vagyis a minőségi termék és szolgáltatás folyamatosan és megfelelő árérték arányban nyújtja a vevőnek, amire szüksége van. A fordítás minőségének kérdése már a kezdetektől fogva foglalkoztatja a fordítással és fordítástudománnyal foglalkozó kutatókat és gondolkodókat. Már maga a minőség fogalma is nehezen definiálhatónak bizonyul a fordítással kapcsolatosan. A következő két alfejezetben a fordítás mint produktum, valamint a fordítás mint szolgáltatás minőségének kérdését mutatom be.

2.1 A fordítás mint szolgáltatás minősége

A fordítás mint szolgáltatás minőségének meghatározása nem könnyű. Szemben a termékekkel, a szolgáltatás mérésére, tartósságának, megfelelőségének meghatározására kevés és nehezen alkalmazható eszköz áll rendelkezésre, valamilyen módon mégis mérhetővé kell tennünk (Urbán 2011). Legkézenfekvőbb megoldásnak a mérésre a szabványok kínálkoznak, melyek ugyan önmagukban nem garantálják a tökéletes minőséget, de biztosítják az ahhoz szükséges feltételeket. Ezek közül az éppen aktuális és érvényben lévő legfrissebb az ISO 17100:2015, a fordítási szolgáltatások minőségére vonatkozó szabvány. Ez a szabvány követelményeket határoz meg a fordítási folyamat valamennyi, a fordítási szolgáltatás minőségét és teljesítést érintő elemével kapcsolatban. Előírásokat tartalmaz a fordításslétszolgálatók számára az alapvető folyamatok irányításával, a képzési követelményekkel, az erőforrások rendelkezésre állásával és kezelésével, valamint további, a minőségi szolgáltatásokhoz szükséges tevékenységekkel kapcsolatban. A szabvány a fordításslétszolgálatók részére készült, függetlenül azok méretétől. A szabványnak való megfelelés akkor lehetséges, ha a nemzetközi szabvány valamennyi rendelkezése teljesül, de a megvalósítás módja eltérhet annak alapján, mekkora a szervezet, és mennyire bonyolult a felépítése, illetve bizonyos esetekben függhet a fordításslétszolgálatótól megrendelt fordítási szolgáltatás volumenétől és komplexitásától. A szabvány tehát a fordításslétszolgálatás egészére fókuszál, nem csupán a fordításra mint végtermékre.

2.2 A fordítás mint produktum minősége

Gile (2009) szerint létezik egy általános konszenzus a jó fordítási minőség összetevőivel kapcsolatban. Ez alapján a jó fordítási minőség ismérvei közé tartozik a gondolatmenet követhetősége, a nyelvi elfogadhatóság, a terminológiai pontosság és elfogadhatóság, a szöveghűség, valamint a professzionális viselkedés (Gile 2009: 45). Gouadec szerint a minőségi fordítás kritériumai a pontosság, a kifejezőkészség, a hozzáférhető-

ség, vagyis, hogy az üzenet legyen érthető a végfelhasználó számára, továbbá a felhasználóbarátság és a hatékonyság, a célnyelv nyelvészeti és kulturális normáinak, jogszabályainak, szabványainak, fizikai és funkcionális szabályainak való megfelelés, végül pedig a megbízó érdekeihez való alkalmazkodás (Gouadec 2009: 6–7).

Dróth (2011) rámutat arra, hogy az egyetemeken az oktatás keretén belül, valamint a fordítói gyakorlatban az értékelés szempontjai egyedileg eltérnek egymástól a különböző elvárások és súlyozás következtében. A képzés során az oktatás aktuális céljának és tartalmának megfelelő értékelési helyzetek váltják egymást, az irodai gyakorlatban azonban számos egyéb, a minőség ellen ható követelmény merül fel. Ilyen többek között a határidők szűkössége, az alacsony vállalási ár és a megrendelők eltérő igényei. Heltai (2005) véleménye szerint a fordítás minőségét a nyelvtani és mondatszerkesztési normák, valamint a helyesírási szabályok követése önmagában nem biztosítja. Nyelvi szempontból jelentős mértékben függ a szövegminőség a szövegalkotási normák betartásától, a fordítás minőségét azonban a nyelvi szempontokon kívüli egyéb összetevők is befolyásolják. Ilyen például a forrásnyelvi szöveghez való viszony vagy az információ átadásának pontossága (Heltai 2005).

2.3 A terminológia szerepe a minőségbiztosításban

A fordítások minőségértékelésének egyik kulcseleme a megfelelő terminológiahasználat. Dróth (2011) tanulmányában összegyűjtötte és csoportosította a 2011. január 21-én megrendezett gödöllői szakfordítói napon résztvevő fordítóirodák, valamint fordítással foglalkozó szervezetek értékelési szempontjait, továbbá a hazai fordítóképzésben leggyakrabban alkalmazott értékelési szempontokat. Mindkét csoport esetében a legfontosabb értékelési kritériumok között szerepelt a pontos terminológiahasználat, a megfelelő terminológia kiválasztása, felkutatása. Heltai kiemeli, hogy a szakfordításon belül kiemelkedő jelentőséggel bír a terminológia ismerete. Az adott terület terminológiájának ismerete nemcsak a szöveg minőségét befolyásolja, hanem az információközvetítés pontosságát is (Heltai 2005: 168).

A pontos terminológiahasználat elősegítése – ezáltal a megfelelő minőség elérése – érdekében a terminológiamenedzsment-folyamatoknak fordítási folyamatba való alkalmazása előnyökkel jár. Nem áll rendelkezésre egyetlen szabvány sem, amely teljes egészében lefedné a terminológia-menedzsmenttel kapcsolatos jó gyakorlatokat és modelleket, de számos szabvány hasznosnak bizonyulhat ebben a kontextusban is. Így az ISO 704:2009, amely a terminológiai munka alapelveit és módszereit határozza meg, vagy az ISO 860:2007, amely a terminusok és fogalmak harmonizációjáról szól. Továbbá hasznosnak bizonyulhat a megfelelő terminológiamenedzsment-folyamatok kialakításában az ISO 12620:2009 számú szabvány, amelynek átdolgozása és frissítése jelenleg is folyamatban van. Ebben a szabványban több mint kétszáz féle adatmezőt definiálnak, melyek egy terminológiai adatbázis részeként szolgálhatnak. Ugyanígy az ISO 12616:2002 számú szabvány is segítségre lehet, amely többek között az adatkategóriákkal is foglalkozik a fordításorientált terminográfiái megközelítés szemszögéből. A terminológiai szabványok ismertetésével Sermann (2013) részletesen foglalkozik doktori disszertációjában.

A terminológiamenedzsment-folyamat kiépítése a fordítási folyamatot támogatja. A terminológia-menedzsment magába foglalja azokat az eljárásokat, modelleket és eszközöket, melyek globális szinten segítik a terminológia hathatós használatát, valamint

az adott fordítási feladatokhoz szükséges terminológia létrehozását. A fordításslálgálatók szempontjából szintén fontos, hogy megfelelő és konzisztens fordításokat biztosítsanak megrendelőik számára, ebben nyújt segítséget a terminológiamenedzsment-folyamatok kialakítása. Egy ilyen terminológiamenedzsment-infrastruktúra részei:

- i) az adatbázis, ahol a terminusokat, valamint a kapcsolódó metaadatokat tárolhatjuk;
- ii) egy CAT eszköz, melynek segítségével hatékonyabban férhetünk hozzá a terminológiai adatbázishoz a fordítási folyamat során;
- iii) egy terminológiakezelő eszköz, mely egyrészlől az adatbázisban való keresést biztosítja, másrészlől lehetőséget nyújt új terminusok terminusjelöltként való felvételére, terminológiamenedzsment-feladatok elvégzésére;
- iv) végül pedig egy QA modul, amely segítségével ellenőrizhetjük, hogy a megfelelő terminológia került-e alkalmazásra (Popiolek 2015: 347).

A pontos terminológiahasználat elősegíti a megértést. Jogszabályok fordítása esetén a fordításnak több, különálló célja is lehet. Honosító szemlélet alkalmazása esetén a cél az, hogy a jogszabály a célnyelvi kultúrában is értelmezhető legyen, míg idegenítő stratégia alkalmazása esetén a cél az, hogy a magyar jogszabály tartalmát, a magyar jogi környezetet írják le és fogalmazzák meg idegen nyelven. Ezek a stratégiák a terminusok kiválasztásának és a terminológiai változtatások szintjén is megjelennek.

2.4 Lektorálás mint minőségbiztosítási eszköz

A terminológia diszciplínáján belül, a terminológiai változtatásokkal foglalkozó szakirodalom igen szűkös, a kutatók leginkább egy-egy terminus kapcsán foglalkoznak a változás részletes bemutatásával. A lektorálás szakirodalmában azonban találhatunk támpontokat a terminológiai változtatások vizsgálatához kapcsolódóan.

A lektorálás célja, hogy a lektorált fordítás magasabb színvonalú legyen, mint amelyet a fordítás önmagában elért (Horváth 2011: 95). Horváth szerint a lektorálás a fordítás ellenőrzése és szükség szerinti kijavítása (Horváth 2011: 233). Mossop definíciója szerint a lektorálás a nyersfordítás átolvasásának folyamata abból a célból, hogy az esetleges meglévő hibákat felismerjük és kijavítsuk (Mossop 2014: 228). Robin definíciója szerint a lektorálás a fordító által célnyelvre ültetett szöveg lektor általi ellenőrzését és módosítását jelenti, azzal a céllal, hogy tartalmilag pontos, nyelvileg helyes, az eredetivel egyenértékű, a célnyelvi olvasó számára feldolgozható szöveg jöjjön létre (Robin 2018: 74). Mindhárom szerző a lektorálás javító, ellenőrző funkcióját emeli ki, melybe többek között a terminológiai pontatlanságok helyreigazítása is beletartozik. Ebből a szempontból a lektorálás is tekinthető terminológiai minőségbiztosító eszköznek. Terminológiai segédeszközök nélkül a lektor feladata a fogalmak szintjén való vizsgálat, a fogalmi szinten való eltérésből eredő változtatás szükségének felismerése.

3 A kutatás bemutatása

Tanulmányomban a fogalmi összevetés módszerét alkalmazva elemzem az Alaptörvény két, angol nyelvű fordításában megvalósult terminológiai változtatásokat. A terminológiai vizsgálatok során fontos a rendszeralapú szemlélet, mivel a fogalmak nem

önmagukban állnak, hanem valamely fogalmi rendszer részét képezik. Az egyedi fogalmak leírását nagymértékben megkönnyíti az adott fogalmi rendszer ismerete. A terminussal jelölt fogalom a dologi ismeretekből absztrakció során alakult ki úgy, hogy az absztrakciós folyamat során a közös jegyeket emelik ki (Fóris 2013: 165).

A terminológiai szemléletmód alapját képezi a fogalomalapú megközelítési mód, melynek értelmében a terminusok elemzésekor először fel kell térképezni a fogalmak jellemzőit, definíciót alkotni, majd csak ezt követően rendelhetünk hozzá a fogalmakhoz terminusokat. A fogalmak a definíciókban jelennek meg, vagyis elsőként a megfelelő definíció felkutatására vagy megalkotására van szükség. Éppen ezért a szaktárgyi tudás fontos szerepet játszik a terminológiában (Tamás 2017a: 19). Az összehasonlítás elvégzése során a terminusokat nem fordítjuk, hanem megfeleltetjük azokat a célnyelvi megfelelővel. Összehasonlítás során érdemes megkeresni azt az egy vagy több megkülönböztető fogalmi jellemzőt, amelyek alapján a fogalmak jól elhatárolhatók egymástól. A fogalmi összevetés korántsem egyszerű feladat. Annak érdekében, hogy megfelelő és pontos eredményt kapjunk, megfelelő forrásból kell definíciót választanunk, esetleg tárgy szakértővel is érdemes konzultálnunk (Tamás 2017a: 20).

3.1 A korpusz

A jelen kutatásom korpuszát az Alaptörvény első változatának, valamint a hetedik módosított változatának angol és magyar nyelven is elérhető Preambulum és Alapvetések fejezetei képezik. Az eredeti szövegrészlet 1445 szó és 11295 karakter hosszúságú, a hetedik módosításra ez 2260 szóra és 18366 karakterre változott. Az eredeti szöveg a Nemzeti Könyvtár oldalról származik, az első változat itt található angol nyelvű szövege a Magyar Közlöny angol nyelvű különszámaként jelent meg. A hetedik módosítás szövege a Nemzeti Jogszabálytár honlapján elérhető, a dokumentum fejlécében jelzik, hogy hivatalos fordításnak nem tekinthető, mégis, mivel az Igazságügyi Minisztérium gondozásában megjelent, ellenőrzött szövegről van szó, így vizsgálatomhoz megfelelő alapot biztosít.

Az Alaptörvényt 2011. április 18-án fogadták el, és 2012. január 1-jén lépett a korábbi Alkotmány (1949. évi XX. törvény) helyébe. Az elmúlt hét év során a törvény több módosításon esett keresztül. Kutatásom jelenlegi fázisában csak a módosítással nem érintett szövegrészeket vizsgáltam annak érdekében, hogy feltárjam, melyek voltak azok a terminológiai változtatások, amelyek ezekben a szövegrészekben előfordultak.

3.2 A korpusz feldolgozása

Először a magyar nyelvű Word-dokumentumokat vettem egybe annak érdekében, hogy megtudjam, melyek azok a szövegrészek, amelyek újonnan kerültek be az Alaptörvénybe. Ezután az angol nyelvű fordításokat a Word szövegösszehasonlító funkciójának segítségével vettem össze. Ennek elvégzése érdekében először egy OCR programmal átalakítottam a .pdf fájlokat .doc fájlkká, majd a forrásnyelvi magyar szövegek mintájára szerkesztettem őket. Az 1. ábra az összehasonlítás eredményeként kapott angol nyelvű szöveget szemlélteti.

Az angol nyelvű összehasonlított szövegből kézi ellenőrzéssel kiválasztottam azokat az eseteket, amelyekben konkrét terminológiai változtatás történt. A terminusok elemzéséhez a terminológiában használt fogalmi összehasonlítás módszerét választottam. Különböző szótárak, jogszabályok és internetes források bevonásával megvizsgáltam az eltérő terminusokhoz tartozó fogalmakat, és összehasonlítottam azokat.

Article-N)¶

(1) Hungary shall ~~enforce~~observe the principle of balanced, transparent and sustainable budget management.¶

(2) ~~Parliament~~The National Assembly and the Government shall have primary responsibility for the ~~enforcement~~observance of the principle ~~set out~~referred to in ~~Paragraph~~paragraph (1).¶

(3) ~~In the course of~~ performing their duties, the Constitutional Court, courts, local governments and other state organs shall be obliged to respect the principle ~~set out~~referred to in ~~Paragraph~~paragraph (1).¹⁴¶

Article-O)¶

~~Every person~~Everyone shall be responsible for ~~his~~him or herself, and shall be obliged to contribute to the performance of state and community tasks according to the best of his or her abilities and ~~potential~~possibilities.¶

Article-P)¶

~~All natural~~(1) Natural resources, ~~especially agricultural in particular arable~~ land, forests and ~~drinking the reserves of water supplies~~; biodiversity; in particular native plant and animal species; and cultural ~~assets~~artefacts, shall form ~~part of the nation's~~ common heritage, ~~and of the nation, it shall be the obligation of~~ the State and ~~every person shall be obliged~~everyone to protect, ~~sustain and maintain them~~, and to preserve them for future generations.¶

(2) ~~The limits and conditions for acquisition of ownership and for use of arable land and forests that are necessary for achieving the objectives referred to in paragraph (1), as well as the rules concerning the organisation of integrated agricultural production and concerning family farms and other agricultural holdings, shall be laid down in a cardinal Act.~~¶

Article-Q)¶

(1) In order to create and maintain peace and security, and to achieve the sustainable development of humanity, Hungary shall strive for cooperation with ~~every national~~all the peoples and ~~country~~countries of the world.¶

(2) ~~Hungary shall ensure harmony between international law and Hungarian law in~~ In order to ~~fulfil~~comply with its obligations under international law, ~~Hungary shall ensure that Hungarian law is in conformity with international law.~~¶

(3) Hungary shall accept the generally recognised rules of international law. Other sources of international law shall become part of the Hungarian legal system by ~~publication in the form of legislation~~promulgation in laws.¶

1. ábra. Az összehasonlítás eredményeként kapott szöveg

4 Eredmények

A vizsgált rész terminológiai változtatásait az 1. táblázatban gyűjtöttem össze. Az első oszlopban az Alaptörvény magyar nyelvű változatában szereplő terminusokat láthatjuk. A táblázat második oszlopában ezen terminusok angol nyelvű megfelelői szerepelnek az Alaptörvény fordításának első változatából, míg a harmadik oszlopban a hetedik módosításban alkalmazott terminológiát láthatjuk.

A 3. fejezetben bemutatott fogalmi összehasonlítás módszerével röviden szemléltetem a terminológiai változtatással érintett fogalmakat. A terminusok elemzésekor először

a magyar és idegen nyelvű terminusokat határozom meg definíciók segítségével, majd fogalmi összehasonlítást végezve megvizsgálom, hogy a megváltoztatott terminusok jobban megfelelnek-e az első módosításban fellelhető alternatíváiknál. Egyes terminusok esetében (például: *Országgyűlés*, *Országgyűlés elnöke*, *országgyűlési képviselő*) az összetételekben csak az egyik tagban történt változtatás (*Országgyűlés* terminus), így ezeket a terminusokat egy elemzési pontban vizsgálom meg. A táblázatban szereplő terminusok közül a terjedelmi korlátok okán az alábbiakat fejtem ki részletesen.

Alaptörvény magyar nyelvű változata	Alaptörvény angol nyelvű első fordítás	Alaptörvény angol nyelvű hetedik módosítás
nemzetiség	nationality	national minority
Országgyűlés	Parliament	National Assembly
hatalom megosztásának elve	principle of separation of powers	principle of the division of powers
jogszabály	legislation	law
termőföld	agricultural land	arable land
jogszabályban történő kihirdetés	publication in the form of legalisation	promulgation in law
Országgyűlés elnöke	Speaker of the House	Speaker of the National Assembly
kihirdetés	publication	promulgation
önkormányzati rendelet	local ordinance	local government decree
törvény	Act of Parliament	Act
miniszterelnöki rendelet	order by the Prime Minister	prime ministerial decree
rendelet	order	decree
hivatalos lap	Official Gazette	official gazette
Magyar Nemzeti Bank elnökének rendelete	orders by the Governor of the National Bank of Hungary	decrees of the Governor of the National Bank of Hungary
önálló szabályozó szerv	autonomous regulatory body	independent regulatory organ
rendeletet ad ki	issue an order	adopt a decree
országgyűlési képviselő	Member of the Parliament	Member of the National Assembly
jogállam	state governed by the rule-of-law	rule-of-law State
tagállam	member state	Member State
nemzetközi szerződés	international agreement	international treaty

1. táblázat. Változtatással érintett terminusok

4.1 Az *Országgyűlés* terminus és annak összetételei

A vizsgált szövegrészletben az *Országgyűlés* terminus az alábbi összetételekben szerepelt: *Országgyűlés*, *Országgyűlés elnöke* és *országgyűlési képviselő*. Itt kétféle változtatásnak lehetünk tanúi, az egyik a *Parliament*<*National Assembly* és a másik a *House*<*National Assembly*.

A magyar *Országgyűlés* terminus fogalmát jól szemlélteti az alábbi, a Parlament honlapjáról származó információkból összerakott definíció: Az Országgyűlés Magyarország legfőbb népképviselői szerve, az országgyűlési képviselők összessége alkotja. Megalkotja az egyes életviszonyokat szabályozó törvényeket és határozatokat, ellenőrzi a végrehajtó hatalom (a kormányzat) munkáját, megválasztja a közjogi tisztviselőket, valamint elfogadja az éves költségvetési javaslatot. Jogállását és működését az Alaptörvény, valamint az Országgyűlésről szóló 2012. évi XXXVI. törvény szabályozza, belső működési folyamatait és szervezetét pedig a házszabályi rendelkezések biztosítják (W1 alapján).

Az első fordításban használt *Parliament* terminus a Black's jogi szótárban az alábbi definícióval szerepel: „1. The supreme legislative body of some countries; esp. (*cap.*), in the United Kingdom, the national legislature consisting of the monarch, the House of Lords and the House of Commons. 2. The period during which a parliament meets” (Garner 2014: 1290). Az Oxford Dictionary of Law szótárában pedig az alábbi módon szerepel: „The legislature of the UK, consisting of the sovereign, the House of Lords, and the House of Commons” (Law 2015: 444). Mindkét definíció alapján elmondható, hogy a *Parliament* terminus valamilyen törvényhozó testületet jelöl, amely országonként eltérően, más-más felépítéssel működik. A második vizsgált fordításban megjelenő *National Assembly* terminus definíciója a Black's szótár szerint:

„(*historical*) In France, the first of the revolutionary assemblies, existing from 1789-1791. – Its chief work was the formation of the constitution, as a result of which it was also known as the „Constitutional Assembly.” The legislatures organized in France in 1848 (after the February revolution) and in 1871 (after the overthrow of the second empire) were also known as National Assemblies” (Garner 2014: 1184).

A szótár archaikus kifejezésként közli az adott szót, azonban, ha megvizsgáljuk az egyes országok törvényhozó és törvényalkotó testületeinek angol nyelvű elnevezéseit, azzal szembesülhetünk, hogy ha egykamarás parlamentről van szó, ezeket inkább nevezik angol nyelven *National Assembly*-nek (pl. National Assembly for Wales, National Assembly of Albania, National Assembly of Greece, National Assembly of Serbia). Így a két terminusnál, bár részben egymás szinonimájának tekinthető, a fogalmi eltérés leginkább abban fogható meg, hogy egy- vagy többkamarás törvényhozó testületről beszélünk-e. A magyar *parlament* és *Országgyűlés* terminusokat megvizsgálva arra figyelhetünk fel, hogy mind a Házszabályi Rendelkezésekben, mind az országgyűlésről szóló 2012. évi XXXVI. törvényben a magyar *Parlament* terminus csak az épület összefüggésében, valamint az Európai Parlamenttel kapcsolatban jelenik meg. A terminológiai változtatás motivációja lehet az, hogy az *Országgyűlés* és *parlament* terminusok szinonimák. Amikor ma parlamentről beszélünk, mindig a törvényhozó szervet értjük alatta. A magyar parlament 1945-től hivatalos elnevezése pedig a nagybetűvel írt Országgyűlés. Az elnevezés annak érdekében alakult így, hogy megkülönböztethető legyen a korábbi, történelmi rendi országgyűlésektől (Trócsányi–Schanda 2014: 116). Ha a magyar *Parlament* terminus az épületet jelöli, ahol a törvényhozás zajlik, abban az esetben angol nyelven a *Parliament* terminust használjuk, viszont, ha a legfőbb népképviselői szervet, az *Országgyűlést*, akkor angol nyelven a *National Assembly* terminust alkalmazzuk. A *House* < *National Assembly* terminológiai változtatást pedig feltehetően a konzisztens terminológiahasználat igazolja, vagyis az, hogy egyes összetételekben (pl. *Member of the National Assembly*) konzisztensen kerüljön alkalmazásra a megváltoztatott szakkifejezés.

4.2 A *nemzetiség* terminus

A *nemzetiség* terminus elemzésekor fontos a nézőpont, az, hogy honnan szemléljük a fogalmat. A magyar *nemzetiség* terminus polyszém szó, egyrészt jelenthet állampolgárságot, nemzeti hovatartozást, másrészt pedig, szűkebb értelemben, valamely Magyarországon elismert nemzetiségi kisebbséghez való tartozást is. A nemzetiségek jogairól szóló 2011. évi CLXXIX. törvény definíciója értelmében nemzetiségnek tekinthető:

1. § [...] nemzetiség minden olyan – Magyarország területén legalább egy évszázada honos – népcsoport, amely az állam lakossága körében számszerű kisebbségben van, a lakosság többi részétől saját nyelve, kultúrája és hagyományai különböztetik meg, egyben olyan összetartozás-tudatról tesz bizonyosságot, amely mindezek megőrzésére, történelmileg kialakult közösségeik érdekeinek kifejezésére és védelmére irányul (2011. évi CLXXIX. törvény, 1. §).

A magyar definíció jól láthatóan kiemeli azt, hogy a nemzetiségi helyzet (számszerű) kisebbségi helyzetet jelent az ország területén. Sem a Black's angol jogi szótár, sem az Oxford Dictionary of Law nem tartalmazza a *nationality* terminus nemzetiségi megfeleltetését, csak állampolgárság értelemben írják le a használatát. A második vizsgált fordításban használt *national minority* terminus definíciója szintén hiányzik a fent említett szótárakból. A Black's szótár a *minority* szócikk alatt az alábbi definíciót közli: „[...] 3. A group that is different in some respect (such as race or religious belief) from the majority and that is sometimes threatened differently as a result; a member of such a group” (Garner 2014: 1147). Az Emberi Jogok Európai egyezményében megjelenik a *national minority* terminus, ahol az alábbi definícióval szerepel:

For the purposes of this convention the expression "national minority" refers to a group of persons in a state who (a) reside on the territory on that state and are citizens thereof, (b) maintain long standing, firm and lasting ties with that state, (c) display distinctive ethnic, cultural, religious or linguistic characteristics, (d) are sufficiently representative, although smaller in number than the rest of the population of that state or of a region of that state, (e) are motivated by a concern to preserve their culture, their traditions, their religion or their language (W2).

Unió szövegekben a *national minority* terminus a *nemzeti kisebbség* ekvivalenseként jelenik meg, a magyar jogszabályi, valamint az angol nyelvű definíciót összehasonlítva ugyanakkor több közös jegyet is felfedezhetünk (pl. számszerű kisebbségi helyzet, eltérő nyelv, kultúra, hagyományok és azok megőrzésére való törekvés). Egészen 2011-ig magyar nyelven is hasonló, *nemzeti és etnikai kisebbség* terminussal nevezték meg a kisebbségeket, az Alaptörvény cserélte le a fenti terminust az új, *nemzetiség* terminusra. A kisebbség terminus még fellelhető bizonyos kontextusban (pl. Regionális vagy Kisebbségi Nyelvek Európai Kartája). Ezek a dokumentumok egyfelől korábban készültek, másfelől a nemzetközi jogi terminológiát alkalmazzák (Tamás 2017a: 51).

4.3 A *hatalom megosztásának elve* terminus

A hatalommegosztás elvéből kifolyólag a hatalmi ágak egymástól elkülönülnek, ezzel is biztosítják a fékek és ellensúlyok rendszerét és a hatalommal való visszaélés gátjaként funkcionál. Lényege az, hogy a törvényhozó hatalomnak, a végrehajtó hatalomnak, valamint az igazságszolgáltatásnak hatáskörben, intézményi és személyi vonatkozásban is el kell válnia egymástól.

A jogtudományi szakirodalomban a hatalmi ágak megkülönböztetésére mind horizontális, mind vertikális módon is lehetőség van. A horizontális felosztás a három hatalmi ág szétválasztását jelenti, míg a vertikális hatalommegosztás az amerikai jogelméletből indult ki, ahol a központi és helyi (eltérő területi alapon szervezett) szervek (pl. állami és szövetségi szervek) egymás döntési szabadságát felügyelik, korlátozzák (Trócsányi–Schanda 2014: 201). Az első fordításban a *principle of separation of powers* terminussal jelölték a fogalmat, amely a Black's szótárban fellelhető definíció szerint a hatalmi ágak elválasztásának horizontális szemléletét jelöli.

„1. The division of governmental authority into three branches of government – legislative, executive, and judicial – each with specified duties on which neither of the other branches can encroach. [...] 2. The doctrine that such a division of governmental authority is the most desirable form of government because it established checks and balances designed to protect the people against tyranny” (Garner 2014: 1572).

Ezzel szemben a Black's *division of powers* terminusra vonatkozó definíciója a vertikális hatalommegosztás elvét szemlélteti: „The allocation of power between the national government and the state. – Under the Tenth Amendment, powers not delegated to the federal government are reserved to the state or the people” (Garner 2014: 582). A terminológiai változtatás oka így az lehet, hogy jogi értelemben jelen esetben a hatalom vertikális és nem horizontális megosztásáról van szó.

4.4 A *jogszabály* terminus és annak összetételei

Az Alaptörvényben fellelhető *jogszabály* (és annak jelen kutatás keretében vizsgált összetétele, a *jogszabályban történő kihirdetés*) terminus változásával Tamás már több cikkében is részletesen foglalkozott (Tamás 2017a; Tamás 2017b), így jelen tanulmányomban nem törekszem a terminológiai változtatás részletes bemutatására. A magyar *jogszabály* terminus az első vizsgált fordításban a *legislation*, a második vizsgált fordításban pedig a *law* terminussal feleltették meg. Tamás (2017b) arra a következtetésre jutott, hogy mind a *law*, mind a *legislation* terminusok kontextusfüggők. A *law* terminus a jogszabályok összességét, illetve jogszabályt is jelölhet. A fontos fogalmi különbség azonban az, hogy a *law* terminus tágabb fogalmat takar, mint a magyar *jogszabály* terminus, mivel eltérően a kontinentális jogrendtől, az esetjogban bírói döntéseket, szokásjogot és bíróságon alkalmazott elveket is magába foglal (Tamás 2017b).

4.5 A *termőföld* terminus

A termőföld védelméről szóló 2007. évi CXXIX. törvény szerint a termőföld „[...] az a földrészlet, amely a település külterületén fekszik, és az ingatlan-nyilvántartásban szántó, szőlő, gyümölcsös, kert, rét, legelő (gyep), nádas vagy fásított terület művelési ágban van nyilvántartva, kivéve, ha a földrészlet az Evt.-ben meghatározott erdőnek minősül” (2007. évi CXXIX. törvény, 2. §). A magyar definíciót megvizsgálva jól látszik, hogy a termőföld és szántóföld fogalmak jól elválaszthatók egymástól. A definícióból kiindulva a szántóföld a termőföld fajfogalmának, a termőföld egyik alkategóriájának tekinthető. A KSH fogalomtára szerint szántónak tekinthető „minden olyan terület, amely rendszeres szántóföldi művelés alatt áll, tekintet nélkül arra, hogy a talajmunkát milyen módon végzik, vagy a területen átmenetileg növénytermelést nem

végeznek (ugaroltatják). A parlagterületek is idetartoznak, amit pihentetés céljából vontak ki a művelésből” (W3). Az első vizsgált fordításban a *termőföld* magyar terminus megfeleltetéseként az *agricultural land* terminussal találkozhatunk, melyet a második fordításban az *arable land* terminusra cseréltek. Az angol nyelvű definíciókat megvizsgálva arra a megállapításra juthatunk, hogy ezek a terminusok országonként eltérően, más-más definícióval rendelkeznek. Az eltérő definíciók oka, hogy a művelési ágak és földhasználati kategóriák más-más felosztásban szerepelnek.

Az Eurostat fogalomtárában az *arable land* terminus a „Land worked (ploughed or tilled) regularly, generally under a system of crop rotation.” (W4) definícióval szerepel, *agricultural land* terminussal itt nem találkozhatunk. Az *agricultural land* helyett az *agricultural area* terminus jelenik meg, mely definíciója szerint „area already used for farming, or that could be brought back into cultivation using the resources normally available on an agricultural holding” (W5). Az *agricultural area* terminus magába foglalja az *arable land*, *permanent grassland*, *permanent crops*, *kitchen garden*, *unutilised agricultural area* és a *special holding area* fogalmait is. A megkülönböztető jegyeket és az alá- és fölérendeltségi viszonyokat jól szemlélteti a FAO, az ENSZ Élelmiségügyi és Mezőgazdasági Szervezetének glosszáriumában szereplő magyarázat, melyet táblázatos formában a 2. táblázatban mutatok be. Bár eltérő terminológiát használ a két forrás, az *agricultural land* terminus a fogalmi jegyek összehasonlítása után megfeleltethetőnek látszik az EU-s *agricultural area* terminusnak. A megfeleltetést erősíti az is, hogy mindkét terminus hasonló fogalmakat foglal magába, és az erdős-fás területeket mindkettő kizárja.

Ha a magyar *termőföld* fogalommal hasonlítjuk össze, arra a megállapításra juthatunk, hogy a fogalmi jegyek összehasonlítása után az első fordításban szereplő *agricultural land* változat sokkal inkább fedi a magyar fogalmat. Éppen ezért a terminus cseréje szükségtelen volt, nem járt a szöveg minőségének emelkedésével.

I	II	III	IV
Agricultural land	Cropland	Arable land	Land under temporary crops
			Land under temporary meadows
			Land under temporary fallow
		Land under permanent crops	
	Permanent meadows and pastures		
Forest or other wooded lands			
Other land			

2. táblázat. A földhasználat kategóriái a FAO-nál

4.6 A *kihirdetés* terminus és annak összetételei

A vizsgált szövegrészletben a *kihirdetés* terminus és annak *jogszabályban történő kihirdetés* összetétele fordult elő. A *jogszabály* terminus változtatását már a korábbiakban bemutattam, így jelen részben csak a *kihirdetés* terminus változásával foglalkozom. Ezt az első fordításban a *publication* terminussal adták vissza, a második fordításban pedig a *promulgation* terminust alkalmazták. A terminus magyar definíciója szerint jogszabály kihirdetésén az írott jogszabálynak azt az első hivatalos közreadását értjük, amely a kötelezettek és jogosítottak tudomására hozza és általuk megismerhetővé teszi az elfogadott és aláírt szöveget. A jogszabályok kihirdetése Magyarország hivatalos lapjában, a Magyar Közlönyben történik (Trócsányi–Schanda 2014: 71–72).

A Black's angol jogi szótárban az első fordításban használt *publication* szerzői jogi értelmezése merül csak fel, ezzel szemben a második fordításban használt *promulgation* terminus az alábbi definícióval szerepel: „The official publication of a new law or regulation, by which it is put into effect” (Garner 2014: 1408). A magyar és angol nyelvű definíciók összevetéséből egyértelműen kitűnik, hogy a második vizsgált fordításban használt terminus fedi inkább a magyar jogi fogalmat, így ebben az esetben a terminológiai változás a szöveg terminológiai minőségének javulásával járt együtt.

5 Összegzés

Tanulmányomban tíz terminológiai változtatást mutattam be részletesen, melyek forrását az Alaptörvény Preambulum és Alapvetések fejezetei képezték. A vizsgálat elvégzéséhez a fogalmi összehasonlítás módszerét alkalmaztam. Az Alaptörvény a magyarországi jogszabályi hierarchia csúcsán helyezkedik el, az általa definiált fogalmak, közjogi tisztségek, számos egyéb jogszabályban megjelennek, jogszabályok alapját képezik. A terminológiai minőségbiztosítás szerepe többek között ezért az, hogy az itt lefektetett alapvetések, fogalmak és megnevezések a jogszabályfordítások esetén egységesen kerüljenek alkalmazásra, a megnevezések fordítása a különböző jogszabályokban egységesen történjen.

A kutatás célja a terminológiai változtatások vizsgálata volt a két eltérő hatályállapotú szöveg között. Az elemzések alapján arra a következtetésre juthatunk, hogy a terminológiai változások megvalósulhatnak:

- konkretizációval vagy honosítással (lásd az *Országgyűlés* terminus),
- köznyelvi szó terminusra váltásával (szakterülethez igazodás) (lásd a *kihirdetés* terminus),
- eltérő kategorizációból eredő változtatás (lásd *hatalom megosztásának elve* terminus),
- (a minőséget negatívan befolyásoló) szükségtelen változtatásként (lásd *termőföld* terminus).

A minőség kérdésköre összetett, a terminusok további elemzésére és összehasonlítására van szükség. A szöveg egészének részletes, fogalmi összehasonlításra alapuló vizsgálata után a fent felsorolt kategóriák még bővíthetnek. A kutatás alapot biztosíthat a terminológiai változtatások kategorizációjának megalapozásához, mely mind az oktatásban, mind a fordítói munkában hasznosnak bizonyulhat.

Források

2007. évi CXXIX. törvény
 2011. évi CLXXIX. törvény
 Az Alaptörvény első fordítása. Elérhető: <http://nemzetikonyvtar.kormany.hu/kiadvanyok>
 (letöltés ideje: 2019. 03. 10.)
 Az Alaptörvény hetedik módosítása. Elérhető:
http://www.njt.hu/translated/doc/TheFundamentalLawofHungary_20180629_FIN.pdf
 (letöltés ideje: 2019. 03. 10.)
 W1 = <http://www.parlament.hu/bevezetes-jogi-szabalyzas1> (letöltés ideje: 2019. 03. 10.)
 W2 = <http://assembly.coe.int/nw/xml/XRef/Xref-XML2HTML-en.asp?fileid=15235>
 (letöltés ideje: 2019. 03. 10.)
 W3 = <https://www.ksh.hu/docs/hun/agrar/html/fogalomtar.html> (letöltés ideje: 2019. 03. 10.)
 W4 = https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:_Arable_land
 (letöltés ideje: 2019. 03. 10.)
 W5 = [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:_Agricultural_area_\(AA\)](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:_Agricultural_area_(AA)) (letöltés ideje: 2019. 03. 10.)

Irodalom

- Dróth J. 2011. A fordítások értékelése a szakfordítóképzésben és a fordítói munka világában. *Fordítástudomány* 13(2): 5–36.
- Fóris A. 2013. A szaknyelvek szociolingvisztikai és terminológiai megközelítése. In: Kontra M. – Németh B. – Sinkovics B. (szerk.) *Elmélet és empiria a szociolingvisztikában (Válogatás a 17. Élőnyelvi Konferencia – Szeged, 2012. augusztus 30.-szeptember 31. – előadásából)*. Budapest: Gondolat Kiadó. 163–173.
- Garner, B. A. (szerk.) 2014. *Black's Law Dictionary*. St. Paul: Thomson Reuters.
- Gile, A. 2009. *Basic Concepts and Models for Interpreter and Translator Training*. Amsterdam: John Benjamins.
- Gouadec, D. 2009. *Translation as a profession*. Amsterdam: Benjamins.
- Heltai P. 2005. A fordítás és a nyelvi norma III. *Magyar Nyelvőr* 129(2): 165–172.
- Horváth P. I. 2011. *A szakfordítások lektorálása. Elmélet és gyakorlat*. Budapest: Tinta Könyvkiadó.
- Law, J. 2015. *Dictionary of Law*. Oxford: Oxford University Press.
- Mossop, B. 2014. *Editing and Revising for translators*. London/New York: Routledge.
- Novák B. 2013. Magyarország új Alaptörvényének olasz nyelvű fordítása közben felmerülő terminológiai kérdések. In: Váradai T. (szerk.) *Alknyelvdok7*. Budapest: MTA Nyelvtudományi Intézet. 159–167.
- Novák B. 2018. *A terminológiai munkafolyamat a minőségi jogalkotásban. Magyar-olasz összehasonlító vizsgálat az alkotmányjogi terminológia területén*. PhD-értekezés. PTE–AJK
- Popiołek, M. 2015. Terminology management within a translation quality assurance process. In: Kockaert, H. J. – Frieda, S. (szerk.). *Handbook of Terminology, vol. 1*. Amsterdam: John Benjamins. 341–359.
- Robin E. 2018. *Fordítási univerzálék és lektorálás*. Fordítástudományi értekezések III. Budapest: ELTE Eötvös Kiadó.
- Sermann E. 2013. *A terminológiai szabványosítás és a terminológiai harmonizáció fordítási vonatkozásai*. PhD-értekezés. Budapest: ELTE BTK, Kézirat.
- Tamás D. M. 2017a. *Bevezetés a jogi terminológiába a terminológus szemüvegén át*. Budapest: ELTE Eötvös Kiadó.
- Tamás D. M. 2017b. *A jogszabály terminus megfeleltetésének kihívásairól az Alaptörvény angol nyelvű fordításaiban*. Elérhető: <http://www.jogiforum.hu/blog/jogi-szaknyelv/188> (letöltés ideje: 2018.12.15.)
- Trócsányi L. – Schanda B. (szerk.) 2014. *Bevezetés az alkotmányjogba. Az Alaptörvény és Magyarország alkotmányos intézményei*. Budapest: HVG-ORAC Lap- és Könyvkiadó Kft.
- Urbán M. 2011. Minőségbiztosítás és értékelési technikák a fordítóiroda gyakorlatában. In: Dróth J. (szerk.). *Szaknyelv és Szakfordítás*. Szent István Egyetem, Gödöllő, Gazdaság- és Társadalomtudományi Kar. 7–19.

Az elérhetőségi elmélet névmási anaforafeloldásra gyakorolt hatása

Kovács Viktória

SZTE BTK Nyelvtudományi Doktori Iskola
viktoria.kovacs12@gmail.com

Kivonat: A névmási anaforafeloldás egy szövegben megtalálható összes visszautaló névmás és a hozzá tartozó legközelebbi antecedens azonosítását jelenti. A cikkben bemutatott gépi tanítási kísérletek az elérhetőségi elmélet (Ariel 2014) névmási anaforafeloldásra gyakorolt hatását mutatják meg két osztályozó kiértékelésén keresztül. Ehhez egy blogbejegyzésekből álló saját korpuszt használok fel, amelyben kizárólag a névmási visszautalások vannak kézzel annotálva. Az első kísérlethez a lehető legkevesebb előelemzési lépéssel, morfológiai és szintaktikai információkat felhasználva építettem a tanítófájlt, míg a második kísérlethez ezekhez az információkhoz hozzáadtam az elérhetőségi elméletben megfogalmazott, automatikusan is kinyerhető elveket.

1 Bevezetés

A névmási anaforafeloldás az anaforafeloldás és a koreferenciafeloldás részfeladata. Ezeknek a viszonyoknak a pontos felismeréséhez hozzájárulnak a morfológiai, a szintaktikai, a szemantikai és a pragmatikai információk is. Az anaforafeloldással kapcsolatos leggyakoribb probléma a referenciális többértelműsége alapul, azaz a kifejezés több antecedensre is visszautalhat a szövegben. Ennek ellenére a kommunikáció során megértik egymást a kommunikációs partnerek, ugyanis az antecedens felismeréséhez segítségül hívhatják a megnyilatkozás kontextusát, azaz a fizikai teret és időt, a diskurzusban korábban elhangzott információkat és a mentális enciklopédiájukat. További problémát jelentenek a zéró névmások, amelyek a szövegben ugyan nem jelennek meg, de visszautalhatnak egy korábbi objektumra. Ez utóbbi probléma megoldásához olyan, a felszíni szerkezetből is kinyerhető információkat használhatunk fel, amelyeknek a segítségével következtethetünk a zéró névmás jelenlétére. A szöveg kontextusára, illetve a szövegalkotó és a címzett mentális állapotára is következtethetünk bizonyos felszíni szerkezeti jegyek alapján, azonban ezeknek a jegyeknek a száma csekély, ezért nagy kihívást jelent az anaforafeloldás mind az elméleti, mind a számítógépes nyelvészet területén. A magyar nyelvvel kapcsolatban számos kezdeményezésről olvashatunk, amelyek egyre jobb eredményeket produkálnak a koreferencia és anaforafeloldás tekintetében (Lejtovicz–Kardkovács 2006; Miháltz 2012; Varasdi et al. 2007), azonban egyik kutatásnak sem kifejezetten a névmási anaforafeloldás áll a fókuszában.

Kovács Viktória: Az elérhetőségi elmélet névmási anaforafeloldásra gyakorolt hatása. In Váradi Tamás (sorozatszerkesztő), Ludányi Zsófia, Grácsi Tekla Etelka (szerkesztő): *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2019. XIII. Alkalmazott Nyelvészeti Doktoranduszkonferencia*. Budapest: MTA Nyelvtudományi Intézet. 2019. 117–125. DOI: 10.18135/Alknyelvdok.2019.13.9

Az anaforafeloldás történet szabály alapú rendszerek segítségével, de emellett lehetőségünk van gépi tanítási kísérleteket is végezni pozitív és negatív példák alapján, ez pedig egyszerűbbé teheti a feladat megoldását. A cikkben is egy ilyen gépi tanítási módszert mutatok be, amely Soon, Ng és Lim 2001-es munkáját veszi alapul (Soon et al. 2001), ennek a rendszernek egy, a feladathoz átalakított változata. A gépi tanítási kísérlet célja egyrészt az, hogy megvizsgáljam mennyire eredményes egy olyan osztályozó, amely a lehető legkevesebb előelemző lépéssel és manuális annotációból származó információk nélkül működik, másrészt pedig az, hogy az elérhetőségi elmélet (Ariel 2014) automatikusan is kezelhető elveinek tanításra gyakorolt hatását megvizsgáljam.

1.1 Elérhetőségi elmélet

Az elérhetőségi elmélet (Ariel 2014) a kifejezések formáját hozza összefüggésbe a referensük mentális elérhetőségével. Az elérhetőségi elmélet szerint a beszélő olyan kifejezést választ, amiről feltételezi, hogy a hallgató értelmezni tudja, tehát a kiválasztás során figyelembe veszi, hogy a hallgató mentális állapotában éppen mennyire van központi pozícióban az adott objektum, amire utalni akar. Az elméletben a különböző fokú mentális elérhetőséget mutató kifejezéseket Ariel egy skálán helyezi el, így a kifejezések a formájuk és nyelvtani tulajdonságaik alapján összehasonlíthatók az elérhetőség szempontjából. Azok a kifejezések, amelyek kevés nyelvi információt tartalmaznak, ilyen a névmás is, magas elérhetőséget mutatnak, azaz a referensük a hallgató mentális állapotának középpontjában helyezkedik el.

Az Ariel (2014) által legfontosabbnak vélt jellemzők az informativitás mértéke, a kifejezés rigidsége és a hossza. Az informativitás mértéke arra vonatkozik, hogy az adott kifejezés mennyire jellemzi hiánytalanul az adott dolgot, amire utal. A rigidség azt mutatja, hogy az adott kifejezés mennyire merev jelölő, a hossz pedig a nyelvi forma hossza, amely a kifejezés írott méretére és fonológiai méretére is vonatkozik. Minél informatívabb, rigidebb és hosszabb egy nyelvi kifejezés, annál alacsonyabb az értéke az elérhetőségi skálán. Ez azt jelenti, hogy ha a beszélő azt feltételezi, hogy a hallgató mentális állapotában nehezen elérhető a referens, akkor nagyon specifikusnak kell lennie a kifejezésnek ahhoz, hogy a hallgató értelmezni tudja az információt. Ezzel szemben, ha könnyen elérhető a címzett számára a referens, elegendő kevésbé rigid és rövidebb kifejezéssel utalni rá, mivel kevesebb erőfeszítésre van szüksége a hallgatónak ahhoz, hogy azonosítsa a kifejezés referensét. Az, hogy egy objektum a beszélgetésben a figyelmi állapot középpontjába kerüljön, és elegendő legyen névmással utalni rá, számos módon kiváltható. Elérhetőbbé teszi az objektumot az, ha jelen van a fizikai kontextusban, vagy ha korábban már szó volt róla. Szintén magas az elérhetősége azoknak az objektumoknak, amelyek szorosan illeszkednek a diskurzusuniverzumba, vagy épp feltűnően nem illeszkednek oda. További tényező még, hogy az adott objektum élőlény vagy tárgy, az elmélet szerint ugyanis az élőlények egyszerűbben elérhetőek.

Tehát amikor a visszautaló névmáshoz tartozó antecedenst keressük, olyan kifejezést keresünk, amely már önmagán hordozza az elérhetőség jeleit. A két kifejezés közötti elérhetőséget külön, relációként jellemzi Ariel (2014), melyben szerepet játszik a távolság is. Azt a megállapítást tette, hogy abban az esetben, ha az antecedens is magas elérhetőségű, messzebből is vissza lehet rá utalni. Az elméleti keret szerint minél közelebb van az antecedens, annál elérhetőbb, annál egyszerűbb azonosítani.

2 Korpusz

A gépi tanítás egy interneten található blogbejegyzésekből, rövid cikkekből álló korpusz segítségével történt. A korpusz szövegei a következő blogokról származnak:

A Prohardver Fujitsu blogjáról (W1), a Websztán blogról (W2), a Könyvkritikák blogról (W3), a Filmvilág blogról (W4), a Városikonyha blogról (W5), az Egyedikutya blogról (W6), a Játéknapló blogról (W7), a Nesze!szer blogról (W8), az Otthonédes blogról (W9). A korpuszban található szövegek választását az indokolta, hogy ezek a szövegek egyszerűen hozzáférhetőek és kevésbé bonyolult a szerkezetük. A választás másik oka az volt, hogy egyre több az online tartalom, ezért egy ilyen rendszer felhasználása is valószínűleg hasonló szövegen történne. A szövegek tartalmi szempontból változatos képet mutatnak. A korpuszban minden egyes névmási visszautalás teljesen manuálisan lett annotálva az MMAX2 annotációs szoftver (Müller–Strube 2006) segítségével. A visszautalásokon kívül azonban semmilyen információ nem lett kézzel a szövegekhez hozzáadva.

A korpuszban található szövegek a magyarlánc (Zsibrita et al. 2013) parse moduljával lettek előelemezve, ezeknek az információknak a segítségével generálja le egy, az adatokat megfelelő formátumra alakító algoritmus a tanítófájlt. A tanítófájlból minden visszautaló névmás és hozzá tartozó antecedensjelölt pár megtalálható, a párhoz hozzárendelt információkkal együtt.

A korpusz 60 db szöveget tartalmazott, összesen 430 névmási visszautalást, ebből 216 vonatkozó névmási, 126 személyes névmási, 88 mutató névmási visszautalás volt.

2.1 A tanításhoz használt párok

A gépi tanításhoz lehetséges visszautaló névmásokból és a hozzájuk tartozó lehetséges antecedensjelöltekből álló párokra van szükség. Ezeket a párokat a magyarlánc parse moduljának segítségével határoztam meg. A magyarlánc parse moduljának kimenete tartalmazza a szóalakot, a lemmát, az MSD-kódot, a POS taget, a morfológiai információkat, valamint a dependencia és konstituens elemzést. A konstituens elemzés segítségével a főnévi csoportok és mondatok, valamint a hozzájuk tartozó elemzések kinyerhetők az elemzett fájlból. A párok első eleme olyan főnévi csoport, amelyhez az elemző a PRON címkét rendeli. Kivételt képeznek ez alól azok a névmások, amelyeknek PronType címkéje *Art* (névelők), *Ind* (határozatlan), *Int* (kérdő), *Neg* (negatív), *Tot* (általános), mivel ezek a névmások nem referálnak egy konkrét objektumra a kontextusban. Kivételt képeztek továbbá azok a személyes névmások, amelyeknek a Person címkéje 1 vagy 2 a morfológiai elemzés során, mivel ezek a névmások a mindenkori beszélőre illetve címzetre referálnak. Tehát az ilyen típusú névmásokhoz nem tudunk antecedenssel rendelni a szövegben. Az antecedensjelöltek a névmásokat a szövegben megelőző főnévi csoportok (NP) és teljes propozíciók (CP). Minden egyes lehetséges visszautaló névmás párt alkot az öt megelőző NP-vel és CP-vel a kézzel is annotált valódi antecedensével bezárólag. Tehát minden esetben annyi pár jön létre, ahány NP és CP található a névmás és az antecedense között, plusz egy, maga az antecedense.

2.2 A tanításhoz felhasznált tulajdonságok

A névmásokból és antecedensjelöltekből álló párokhoz hozzárendeltem még a két kifejezésre vonatkozó morfológiai, szintaktikai, illetve szemantikai és pragmatikai tudásból adódó információkat. A kísérlet célja az volt, hogy viszonylag egyszerűen, manuális annotáció nélkül is kinyerhető információkat használjon fel az osztályozó. Jelenleg a tanítófájlban a névmási anafora és antecedens párokat 21 tényező jellemzi, ezek mind a magyarlánc előelemzésének kimenetéből, illetve a nyers szövegből automatikusan kinyert információk. A tanítófájlba maguk a kifejezések nem kerültek bele, kizárólag az őket jellemző tulajdonságok.

A tanítófájlban a párok a következő módon vannak jellemezve:

```
[antLen, CPdist, NPdist, wDist, antCat, antPOS, anaTyp, anaCas,
antCas, casAgr, anaNum, antNum, numAgr, anaPer, antPer, perAgr,
antNom, antPrp, antPron, antDef, ant3, anaphoric]
```

A jellemzők két fő csoportra oszthatók. Az első csoportba morfológiai és szintaktikai jellemzők kerültek, amelyeket közvetlenül a magyarlánc által a szavakhoz rendelt címkékből nyertem ki. A másik nagyobb csoportba azok a jellemzők tartoznak, amelyeket szintén a magyarlánc kimenetéből, de nem az előelemzés segítségével nyertem ki. Az utolsó jellemző pedig a manuális annotációból származó adat, amely azt mutatja, hogy a pár anaforikus-e vagy sem. A következő két fejezetben ezen jellemzők részletes ismertetése található.

2.2.1 A tanításhoz felhasznált morfológiai és szintaktikai tulajdonságok

Az *antCat* jellemző a konstituens elemzésből származó címke, azt mutatja meg, hogy az antecedensjelölt CP vagy NP címkét kapott-e a konstituens elemzés során, tehát az antecedensjelölt teljes proposíció vagy főnévi csoport. Ennek megfelelően a *CP* vagy az *NP* értéket veheti fel.

Az *anaTyp* jellemző a névmás típusát jelöli. A magyarlánc morfológiai elemzőjének *PronType* típusú címkéit veheti fel értéként. A párokba rendezés során kizárt címkéken kívül ezért a *Dem* (mutató) a *Prs* (személyes) a *Rel* (vonatkozó) és *Rcp* (visszaható) értékeket veheti fel.

Az antecedensjelölt esete *antCas*, száma *antNum*, és személye *antPer* három különböző jellemzőként jelenik meg a tanítófájlban, és szintén a morfológiai elemzésből származik. A főnévi csoportnál a csoport fejéhez rendelt Case típusú morfológiai címkével egyezik meg az érték, a mondatoknál pedig hiányzó értéket jelez.

Az anaforikus névmás esete *anaCas*, száma *anaNum*, személye *anaPer* szintén három különböző jellemző, amelyek a morfológiai elemzésből származnak.

Az egyeztetés eset szerint *casAgr*, szám szerint *numAgr*, személy szerint *perAgr* jellemzők azt vizsgálják meg, hogy a névmáshoz rendelt eset, szám és személy címke és az antecedenshez rendelt eset, szám és személy címke megegyezik-e. Abban az esetben, ha megegyezik a jellemző, az 1-es értéket veszi fel, ha pedig nem, a 0-t.

Az *antPrp* jellemző azt jelöli, hogy az antecedensjelölt tulajdonnév-e. Abban az esetben, ha az antecedens a PROPON POS taggel rendelkezik, az érték 1, ha nem, az érték 0.

Az *antPron* jellemző azt jelöli, hogy az antecedensjelölt névmás-e. Abban az esetben, ha az antecedens a PRON címkével rendelkezik, az érték 1, ha nem, az érték 0.

Az *antPOS* jellemző a magyarlánc elemzésében az antecedensjelölt fejéhez rendelt POS taget írja ki értéként.

2.2.2 A tanításhoz felhasznált további tulajdonságok

A szintaktikai és morfológiai jellemzőkön kívül további tulajdonságokat is figyelembe vett az osztályozó, amelyeket szintén az előelemzés kimenetéből nyertem ki. Az elérhetőségi elméletben (Ariel 2014) találhatóak olyan alapelvek, amelyek a névmáshoz tartozó antecedens azonosítására vonatkoznak, és automatikusan is meghatározhatók.

Az *antNom* jellemző két értéket vehet fel, az értéke 1, ha alany esetű az antecedens, és 0, ha egyéb. Ennek a jellemzőnek az alapja az elérhetőségi elméletben megfogalmazott egyik alapelv: az elérhetőségi elv alapján az alany esetű főnévi csoportok könnyebb elérhetőséget mutatnak a mentális állapotban, ezért valószínűbb, hogy a későbbiekben is visszautalunk rájuk.

Az *antDef* jellemző szintén két értéket vehet fel attól függően, hogy az antecedensjelölt főnévi csoport a morfológiai elemzés során a Definite tulajdonság szerint milyen értéket kapott. *Def* értéket vesz fel, ha határozott, és *Ind*, ha határozatlan. Abban az esetben, ha a morfológiai elemzésben nincs a határozottságra utaló címke, hiányzó értéket jelöl a rendszer. A határozatlan névelős kifejezések gyakran diskurzusreferenst szoktak bevezetni, a frissen bevezetett diskurzusreferens által jelölt dolog pedig elérhetőbb, mint a már bevezetettek, ezért ez az információ is befolyásolhatja az osztályozó eredményességét.

Az *antLen* jellemző azt mutatja, hogy az antecedensjelölt hány szóból áll, tehát ez egy skaláris érték. Az elérhetőségi elmélet szerint a kifejezések hossza is mutatja az általuk jelölt dolog mentális elérhetőségét, a hosszabb kifejezések az általuk jelölt dolog nehezebb elérhetőségét mutatják. Ariel (2001) szerint éppen ezért az antecedens hossza és a visszautaló névmás és az antecedense közötti távolság nagysága között összefüggés van.

Az *ant3* jellemző két értéket vehet fel: ha több mint három szóból áll az antecedensjelölt, akkor 1, ha nem, akkor 0. Ez a jellemző kapcsolódik az *antLen* jellemzőhöz, Ariel (20014) elméleti keretében a három vagy annál több tartalmas szóból álló főnévi csoportok alacsony mentális elérhetőséget mutatnak.

Az antecedensjelölt és a névmás közötti távolságot három érték alapján is figyelembe veszi a rendszer.

A *CPdist* jellemző azt mutatja, hogy a névmástól számított hányadik tagmondatban helyezkedik el az antecedensjelölt, tehát ha azonos tagmondatban vannak, 0 az értéke, ha szomszédos tagmondatban akkor 1. Minden tagmondati határátlépés eggyel növeli az értéket.

Az *NPdist* jellemző azt mutatja meg, hogy a névmás és az antecedensjelölt között hány darab főnévi csoport helyezkedik el.

A *wDist* azt mutatja meg, hogy a névmás és az antecedensjelölt között hány szó található. Ebből az értékből következtethetünk az antecedens és a névmás között elhelyezkedő főnévi csoportok hosszára, tehát arra, hogy milyen elérhetőségű főneveket ítélnet nem anaforikusnak az osztályozó.

2.3 A gépi tanítási kísérletek

A meghatározott jellemzők segítségével a tanulófájlon több osztályozót is építettem a Weka szoftver (Eibe et al. 2016) segítségével, a legeredményesebb ezek közül a Random Forest (Breiman 2001) volt. Ez az algoritmus nem tipikusan ehhez a feladathoz készült, azonban a baszk nyelvvel kapcsolatban is hasonlóan ez az osztályozó ért el jobb eredményeket (Arregi et al. 2010).

2.3.1 Az osztályozó tesztelése

A két osztályozó teszteléséhez az alacsony számú visszautalás miatt a keresztvalidálás módszerét alkalmaztam. A korpuszt a szövegek alapján tíz részre osztottam, kilenc részből készült el a tanítófájl, és egy részből a tesztfájl, ezt a módszert pedig tízszer megismételtem, a végleges kiértékeléshez pedig az egyes tesztek átlagát használtam fel.

A tesztfájlokban a párok úgy jöttek létre, hogy a névmást közvetlen megelőző főnévi csoport volt az első lehetséges antecedensjelölt, amelyet a névmáshoz rendeltem. Egészen a névmást tartalmazó tagmondatot megelőző tagmondat határáig a főnévi csoportokat rendeltem a névmáshoz, ezután a következő antecedensjelölt a névmást tartalmazó tagmondatot megelőző tagmondat volt. Ez alól kivételt képeztek azok az esetek, amikor a mondat első szava névmás volt, ebben az esetben a megelőző tagmondat volt az első antecedensjelölt. Tehát a főnévi csoportok és a tagmondatok felváltva a névmástól számított távolságuk alapján lettek a névmáshoz rendelve egészen a szöveg elejéig. A párok sorbarendezésére azért volt szükség, mert amikor az osztályozó egy névmáshoz és a hozzá rendelt lehetséges antecedensjelölt párhoz az anaforikus viszonyt ítélte, továbblépett a következő névmásra, így egy névmáshoz csak egy jelöltet talált.

3 Eredmények

Ahhoz, hogy megtudjam, hogy az általam épített osztályozó mennyire eredményes, több tanítási és tesztelési kísérletet is elvégeztem. A Base Line-ként használt legegyszerűbb antecedensmeghatározáshoz egyszerűen figyelmen kívül hagytam a fent említett jellemzőket, és minden névmáshoz az első lehetséges, öt megelőző antecedensjelöltet rendeltem valódi antecedensként, ennek az esetnek az eredményei a következő táblázatban a *firstmatch* oszlopban találhatóak. Az összes többi tanítási kísérletet ehhez hasonlítottam, így mutatva meg, hogy a teljes tanítófájl és jellemzőkészlet mennyire hatékony.

A pusztán morfológiai és szintaktikai jellemzőket felhasználó osztályozó eredményeit mutatja a *withoutAccessibility(w/oA)* oszlop. Ehhez építettem a tanítófájlok segítségével olyan osztályozókat, amelyek nem vették figyelembe az *antLen*, *antDef*, *ant3*, *antNom* *antPron* és *wDist* jellemzőket.

Az elérhetőségi elmélet automatikusan is kezelhető elveinek a tanításra gyakorolt hatását mutatják a teljes jellemzőkészletet és a hozzájuk rendelhető összes értéket felhasználó osztályozók, ezek eredményei a *trainFull* oszlopban találhatóak.

A tíz teszt átlagát tekintve tehát az elérhetőségi elmélet számítógéppel automatikusan is kezelhető elvei a pontosságon (precision) 8,09%-ot, a fedésen (recall) 1,18%-ot, az F-mértéket tekintve pedig 2,51%-ot javítottak. A tesztek eredményeit a 1. táblázat mutatja.

Az elérhetőségi elmélet alapján megfogalmazott jellemzők gépi tanításra gyakorolt hatását a Szeged korpusz koreferenciaannotált részében (Csendes et al. 2005) található névmási visszautalások tekintetében is teszteltem, mivel ez a korpusz több névmási visszautalást tartalmaz. A korpuszban nincsenek jelölve a tagmondatra történő visszautalások, ezért az osztályozás előtt kizárólag a főnévi csoportokat rendeztem párba a névmásokkal. Az eredményeket a 2. táblázat mutatja.

A tíz teszt átlagát tekintve tehát az elérhetőségi elmélet számítógéppel automatikusan is kezelhető elvei a névmáshoz tartozó antecedens azonosítását tekintve a Szeged

korpuszon a pontosságon (precision) javítottak 4,92%-ot, a fedésen (recall) rontottak 2,1%-ot, az F-mértéket tekintve pedig javítottak 1,92%-ot.

	<i>firstmatch</i>			<i>w/oA</i>			<i>trainFull</i>		
	P	R	F	P	R	F	P	R	F
TEST1	4,04	26,67	7,02	31,58	13,33	18,75	38,89	15,56	22,23
TEST2	2,99	16,67	5,07	52,38	18,33	27,16	68,75	18,33	28,94
TEST3	3,42	14,58	5,54	55,00	22,92	32,36	63,13	25,00	35,82
TEST4	3,11	14,29	5,11	53,58	25,00	34,09	63,63	32,56	43,08
TEST5*	4,84	13,04	7,06	75,00	39,13	51,43	75,00	39,13	51,43
TEST6	0,99	12,50	1,83	50,00	27,50	35,48	61,11	27,50	37,93
TEST7**	0,85	9,52	1,56	70,59	28,57	40,68	68,75	26,19	37,93
TEST8	4,17	14,71	6,50	57,90	32,35	41,51	71,43	29,41	41,67
TEST9	0,57	6,45	1,05	55,56	32,26	40,82	57,90	35,48	44,00
TEST10	0,90	9,80	1,65	54,55	23,53	32,88	68,42	25,49	37,14
ÁTLAG	2,59	13,82	4,24	55,61	26,29	35,51	63,70	27,47	38,02

1. táblázat. A tanítási kísérletek adatai (P = precision, R = recall, F = F-measure)

	<i>w/oA</i>			<i>trainFull</i>		
	P	R	F	P	R	F
TEST1	21,93	34,25	26,74	24,77	36,99	29,67
TEST2	27,73	47,14	34,92	36,96	48,57	41,98
TEST3	28,33	44,74	34,69	35,42	44,74	39,54
TEST4	40,74	45,21	42,86	45,21	45,21	45,21
TEST5	37,17	53,85	43,98	42,55	51,28	46,51
TEST6	31,65	39,68	35,21	30,12	39,68	34,25
TEST7	40,00	59,70	47,90	58,33	41,79	48,69
TEST8	41,77	49,25	45,20	40,79	46,27	43,36
TEST9	32,80	37,84	35,14	34,74	44,60	39,06
TEST10	39,64	53,00	45,36	42,10	44,58	43,30
ÁTLAG	34,18	46,47	39,20	39,10	44,37	41,16

2. táblázat. A Szeged korpuszon végzett tanítási kísérletek adatai (P = precision, R = recall, F = F-mérték)

3.1 Kiértékelés, hibaelemzés

Az internetes blogkorpusznál tíz tesztből nyolc esetében javítottak az elérhetőségi elmélet alapján megfogalmazott jellemzők, egy esetben nem változtattak az eredményen (*), egy esetben pedig rontottak (**). Az egyes tesztek értékei nagy eltéréseket mutatnak egymástól. Ennek az az oka, hogy a tesztfájlokban egymástól eltérő mennyiségű visszautalás volt, de egyik teszt sem tartalmazott 60-nál több visszautalást, tehát már egy anaforikus kapcsolat is több százalékos eltérést jelent a kiértékelésben.

Egyedül a TEST7 esetében mondható el, hogy az elérhetőségi elmélet alapján megfogalmazott elvek rontottak a tanítás sikerességén. A tesztfájlból 42 db névmási visszautalás volt, ebből 18 személyes, 9 mutató és 15 vonatkozó névmási. A személyes

névmási visszautalások közül egyiket sem azonosította az osztályozó, és a mutatónévmási visszautalások közül is mindössze kettőt. Ennek oka egyrészt az, hogy magában a tanítófájlban is kevesebb az ilyen típusú visszautalás, ezért a tesztelés során is átlagosan kisebb a felismert visszautalások aránya. Másrészt a személyes névmási visszautalás esetében a visszautalás távolsága átlagosan nagyobb, így az osztályozó tévesen egy közelebbi főnévi csoportot jelöl meg antecedensnek. Abban az esetben, ha az osztályozó nem csak az első antecedensnek azonosított főnévi csoportig vagy tagmondatig fut, azonosítja a névmás valódi antecedensét is.

Általános problémát jelent azoknak a névmásoknak a kezelése is, amelyeknek nem szükséges a szövegben antecedenset keresni. Jelenleg az osztályozó a deiktikus névmásokhoz is sorra vizsgálja a lehetséges jelölteket. Az ebből a problémából fakadó hibákat úgy lehetne orvosolni, hogy a tanítófájlhoz olyan negatív példákat adok, amelyek deiktikus névmásokat tartalmaznak, illetve előszűröm a névmásokat a tesztfájlban.

Az osztályozó jelenleg csak a szövegben megjelenő névmásokat és a hozzájuk rendelhető antecedensjelölteket vizsgálta meg. Abban az esetben, ha a zérónévmásokhoz is lehetne antecedensjelöltet rendelni szintén egy megelőző lépés segítségével, kevesebb antecedensjelöltet kellene az osztályozónak átvizsgálnia, ez lehetséges, hogy javítana az eredményen.

A Szeged korpuszon végzett teszt eredményeinek értékelése során figyelembe kell venni, hogy a Szeged korpusz koreferenciaannotált változatában csak a főnévi csoportra történő visszautalások vannak annotálva, tehát a visszautalások száma alacsonyabb, ezáltal a pozitív tanítópéldák száma is kevesebb. Ennek ellenére összességében az F-mértéket tekintve így is javítottak az elérhetőségi elmélet alapján megfogalmazott elvek a tanítás sikerességén.

4 Konklúzió

A gépi tanítási kísérlet célja egyrészt az volt, hogy megvizsgáljam mennyire eredményes egy olyan osztályozó, amely a lehető legkevesebb előelemző lépéssel és manuális annotációból származó információk nélkül működik, másrészt pedig, hogy az elérhetőségi elmélet automatikusan is kezelhető elveinek tanításra gyakorolt hatását ellenőrizzem.

Az elérhetőségi elméletben megfogalmazott, automatikusan is kezelhető elvek növelték a tanítás hatékonyságát, leginkább a fals pozitív esetek szűrésében hatékonyak, tehát érdemes további kognitív nyelvészeti és szemantikai, pragmatikai elveket is figyelembe venni a gépi tanítás során.

A tesztek alapján az a következtetés vonható le, hogy a fals negatív példák száma okozza a legnagyobb problémát. Ezen elsősorban a tanítófájl növelésével lehet javítani. Egy másik megoldási lehetőség, hogy az osztályozás során nem csak az első antecedensnek értékelt főnévi csoportot, illetve tagmondatot vesszük figyelembe, hanem többet, és egy második lépésben ezeken a jelölteken újabb szűrést hajtunk végre.

Azonban ezzel együtt is további előelemző lépésekre van szükség, amelyek pontosan azonosítják azokat a névmásokat a szövegben, amelyekhez ténylegesen szükséges antecedenset keresni.

Források

- W1 = [https://prohardver.hu/fooldal/rovat/fujitsu_blog]
W2 = [http://webisztan.blog]
W3 = [http://konyvkritikak.blog.hu]
W4 = [http://filmvilag.blog]
W5 = [https://varosikonyha.blog.hu]
W6 = [https://www.egyedikutya.hu/egyedi-kutya-blog]
W7 = [https://jateknaplo.blog.hu]
W8 = [https://neszeszer.blog.hu]
W9 = [https://otthonedes.blog.hu]

Irodalom

- Ariel, M. 2001. Accessibility theory. An overview. In: Sanders, T. – Schilperoord, J. – Wilbert S. (szerk.) *Text Representation: Linguistic and Psycholinguistic Aspects*. Amsterdam: John Benjamins Publishing Company. 29–87.
- Ariel, M. 2014. *Accessing noun-phrase antecedents*. London: Routledge.
- Arregi, O. – Ceberio, K. – Diaz de Illaraza, A. – Goenaga, I. – Sierra, B. – Zelaia, A. 2010. A First Machine Learning Approach to Pronominal Anaphora Resolution in Basque. In: Kuri Morales, A. – Simari, G. R. (szerk.) *Advances in Artificial Intelligence*. Berlin, Heidelberg: Springer. 234–243.
- Breiman, L. 2001. Random Forest. *Machine Learning* 45(1): 5–32.
- Csendes, D. – Csirik, J. – Gyimóthy, T. – Kocsor, A. 2005. The Szeged Treebank. In: Matoušek, V. – Mautner, P. – Pavelka, T. (szerk.) *Proceedings of the 8th International Conference on Text, Speech and Dialogue (TSD 2005)*. Karlovy Vary, Czech Republic: Springer. 123–131.
- Eibe, F. – Hall, M. A. – Witten, I. H. 2016. *The WEKA Workbench. Online Appendix for „Data Mining: Practical Machine Learning Tools and Techniques”*. Fourth Edition. Morgan Kaufmann.
- Lejtovicz K. E. – Kardkovács Z. T. 2006. Anaforafeloldás magyar nyelvű szövegekben. In: Alexin Z. – Csendes D. (szerk.) *IV. Magyar Számítógépes Nyelvészeti Konferencia: MSZNY 2006*. Szeged: Szegedi Tudományegyetem. 362–363.
- Miháltz M. 2012. Tudásalapú koreferencia- és birtokosviszony-feloldás magyar szövegekben. In: Kenesei I. – Prószték G. – Várad T. (szerk.) *Általános Nyelvészeti Tanulmányok 24*: 151–166.
- Müller, C. – Strube, M. 2006. Multilevel annotation of linguistic data with MMAX2. In: Braun, S. – Kohn, K. – Mukherjee, J. (szerk.) *Corpus Technology and Language Pedagogy. New Resources, New Tools, New Methods*. (English Corpus Linguistics 3) Frankfurt: Peter Lang. 197–214.
- Soon, W. M. – Ng, H. T. – Lim, D. C. Y. 2001. A machine learning approach to coreference resolution of noun phrases. *Computational Linguistics* 27(4): 521–544.
- Varasdi K. – Vajda P. – Miháltz M. – Naszódi M. 2007. NP-koreferenciák feloldása magyar szövegekben a Magyar WordNet ontológia segítségével. In: Tanács A. – Csendes D. (szerk.) *V. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged: Szegedi Tudományegyetem. 138–146
- Zsibrita J. – Vincze V. – Farkas R. 2013. magyarlanc: A Toolkit for Morphological and Dependency Parsing of Hungarian. In: Angelova, G. – Bontcheva, K. – Mítkov, R. (szerk.) *Proceedings of RANLP 2013*. Hissar, Bulgaria: Shoumen. 763–771.

Névelem-felismerés magyar nyelvű jogi szövegeken

Üveges István

SZTE BTK Nyelvtudományi Doktori Iskola
uvegesistvan898@gmail.com

Kivonat: A jelen tanulmányban a névelem-felismerés hatékonyságának elemzésére tesztek kísérletet jogi szövegek területén. A vizsgálat során részletesebb elemzésnek vettem alá a két elemzőt: a magyarul nyelvű elemzőt és a szintén az MTA-SzTE Mesterséges Intelligencia Kutatócsoport fejlesztette korábbi tulajdonnév-felismerő kimenetét. Elsőként röviden ismertetem a jelen elemzés szempontjából lényeges szakirodalmi hátteret. Ezt követően a vizsgálat tárgyát képező adatok kvantitatív elemzésére térek ki bővebben. A tanulmány következő részében néhány reprezentatív, problémás esetet és ezekre vonatkozó megoldási javaslatot ismertetek, amelyeket végül a további kutatási irányok meghatározása követ.

1 Bevezetés

A névelem-felismerés a szövegbányászat, ezen belül is az információkinyerési feladatok egyik alterülete. A szövegbányászat maga „szöveges adatokon végzett feldolgozási és elemzési tevékenység, melynek célja a dokumentumokban rejtetten meglévő új információk feltárása, azonosítása és elemzése” (Tikk et al. 2006: 22). Ezzel összefüggésben információkinyerésen „a szövegbányászati feladatok egy speciális esetét értjük, ahol a cél az adott feladat szempontjából fontos szövegrészek (információk, tények) kigyűjtése a dokumentumokból, azaz strukturálatlan szövegekből strukturált információ előállítását” (Tikk et al. 2006: 81). A névelem-felismerés a számítógépes nyelvészetben egészen a korai 1990-es évektől kezdve van jelen, és azóta is fontos feladatnak és megoldandó problémának számít. A jelen tanulmányban a hangsúly a jogi szövegekben fellelhető névelemek felismertetésén van, ahol a dokumentumok (főleg) automatikus anonimizációja, valamint az informatívabb és hatékonyabb keresőeszközök kifejlesztése iránt napjainkban is folyamatos az érdeklődés. A jogi doménben névelemek alatt nem pusztán emberek vagy szervezetek neveinek említéseit érthetjük, de számításba kell vennünk például törvények neveit is.

A nemzetközi szakirodalomban számos olyan projekttel találkozhatunk, amelyek a névelemek felismerését tűzték ki célul az angolszász jogrendszer dokumentumaiban (pl. Lenci et al. 2009; Quaresma–Gonçalves 2010; Surdeanu et al. 2010), de a névelemek felismerésének kérdése megjelenik a hazai szakirodalomban is (Móra et al. 2011; Simon 2008; Simon 2017; Vincze–Farkas 2012 stb.).

Gyakorlati szempontból a névelemek felismerésével és kategorizálásával a jogi információkinyerés elősegíthető mind a jogászok, bíróságok, valamint egyéb kormányzati szervek, mind pedig a joghoz laikusok számára.

2 Módszertan

Ebben a fejezetben röviden ismertetem a vizsgálat alapjául szolgáló korpuszt, illetve a kutatás metodológiáját.

A magyar nyelvre jelenleg három módon lehetséges a tulajdonnevek automatikus annotációja: tulajdonnév-felismerő algoritmussal, szófaji címke szintjén történő megkülönböztetéssel, valamint szintaktikai szintű címkézéssel. Utóbbi kettőre példa a jelen tanulmányban is elemzett magyarlanc működése, mivel ez szófaji szinten megkülönbözteti a tulajdonneveket, a szintaxis szintjén pedig jelöli a többtagúakat.

A tulajdonnév-felismerő algoritmusok megkeresik az adott szövegben a tulajdonneveket, majd azokat valamilyen kategóriába sorolják. Az [1] esetében például az algoritmus személynévi (PER) címkével látta el az adott tokent.

[1] Péter I-PER

A [2] a) egy névelem szófaji megkülönböztetését, a [2] b) pedig a többtagú névelemek szintaktikai elemzését szemlélteti. Utóbbi esetében a névelem tagjait összekötő NE (Named Entity) él mutatja meg az érintett tokeneket.

[2]

- | | | | | | | | |
|----|---|-------|-------|-------|---------------------------------|---|-----|
| a) | 7 | Dóm | Dóm | PROPN | Case=Nom Number=Sing | 8 | ATT |
| b) | 6 | Dóm | Dóm | PROPN | Case=Nom Number=Sing | 7 | NE |
| | 7 | utcai | utcai | ADJ | Case=Nom Degree=Pos Number=Sing | 8 | ATT |

A továbbiakban elsőként a vizsgált adatokra térek ki, majd a fent említett három eljárás általam elvégzett vizsgálatát ismertetem röviden.

2.1 Adatok

A vizsgálatot a Miskolc Jogi Korpuszon (Vincze 2018) végeztem, amely jogász, nyelvész és informatikus szakértők közreműködésével készült el annak érdekében, hogy megnyissa az utat a jogi nyelv (akár korpusznyelvészeti) egzakt tanulmányozása előtt. Összeállítása során fontos célkitűzés volt, hogy a létrejövő szövegállomány a magyar jogi nyelv minél szélesebb szegmensét lefedje, így hat különböző forrásból tartalmaz szövegeket (vö. Vincze 2018):

- 5 magyar törvény teljes szövege (a továbbiakban: Törvények),
- jogi témájú fórumok szövegei (a továbbiakban: Fórumok),
- bírósági tárgyalások és rendőrségi kihallgatások átiratai (a továbbiakban: Átiratok),
- jogszabályok miniszteri indoklásai és jogi egyetemek tankönyveinek szövegei,
- bírósági és törvényszéki ítéletek szövegei,
- törvényekből és jogszabályokból kiválasztott szövegrészek.

A jelen elemzés bázisát a Törvények, Fórumok és Átiratok részkorpusz első, megközelítőleg 6000 tokenje adta (1. táblázat).

A vizsgált szövegtípusok kiválasztásakor a fő szempont a minél inkább heterogén tipográfiai megjelenés (tagolás, központosítás, bekezdések stb.), valamint a jogi nyelv intuitívan és korábbi doménhasonlósági vizsgálatok alapján (vö. Vincze 2018) is leginkább eltérő aspektusainak reprezentálása volt. Ennek megfelelően a metanyelvi szövegeket a fórumbejegyzések, a „klasszikus” jogi szövegeket a törvények, míg a beszélt nyelvi jogi témájú szövegeket az átiratok reprezentálták.

Részkorpusz	Tokenszám	Szószám
Törvények	6014	4660
Fórumok	6041	4718
Átiratok	6010	4594

1. táblázat. A vizsgált részkorpuszok jellemzése

Az elemzés során fontos kérdés volt, hogy a jogi nyelv eltérő forrásai eltérő kezelést igényelnek-e a névelemek felismerése szempontjából. Ennek megválaszolására a kvantitatív vizsgálat három szinten valósult meg:

- Az első lépést a kiválasztott szövegrészek manuális annotálása jelentette a bennük előforduló névelemekre.
- Ezt követően a standardként szolgáló manuális annotációs eredmények összevetése történt meg a magyarul címkézésének eredményével, valamint a tulajdonnév-felismerő kimenetével.
- Végül a magyarul függőségi nyelvtani elemzésében megjelenő (vagy éppen hiányzó) NÉ (Named Entity) címkék felmérése következett a többszavas névelemek esetében.

Az elemzés kvalitatív részében az automata elemzők leggyakoribb hibaforrásainak felmérése történt meg. Az így szerzett adatok betekintést engednek a jogi szövegek néhány olyan sajátosságába, amelyek a kiválasztott szövegek esetében megnehezítették az automatikus elemzést, így segíthetnek a jövőbeni tulajdonnév-felismerő elemzők pontosságának javításában is.

2.2 Manuális annotáció

A kontroll adatokat a kézi annotálás adta, amely nagyban támaszkodott a HunNER korpusz annotálása során alkalmazott irányelvekre (Simon et al. 2006).

Ezekhez képest a lényegi eltérést a tag-for-tagging elv alkalmazása jelentette, azaz, hogy a keresett kifejezések nem az aktuális, szövegeli szerepük, hanem lexikális jelentésük alapján lettek kategorizálva, mintegy egyszerűsítve ezzel az automatikus elemzők kimenetének kiértékelését. A névelemek annotálandó kategóriáihoz az ACE 2006 annotálási útmutató szolgált alapul (Linguistic Data Consortium 2006), mindazonáltal az ott felsorolt jelentős számú kategória közül csak a név szerinti említések (name mentions), a helyek nevei (locations) és a szervezetek nevei (organizations) kerültek be a végleges, annotálandó kategóriák közé. Így tehát az annotálás során keresett három alapkategória a személynevek (PER), helynevek (LOC) és szervezetek (ORG) megnevezései voltak.

Emellett említést érdemelnek még a törvények, rendeletek nevei is (pl. Ptk., Tht.) amelyek szintén jelölésre kerültek. Az annotált kifejezéseket a 2. táblázat mutatja be részletesen.

Fontos még említést tenni az alapvető különbségről az angol névelem-felismerés (Named Entity Recognition – NER) és a magyar tulajdonnév-felismerés között. A névelem-felismerés egy alapvetően tágabb vizsgálati tartományt jelöl ki, amelybe beletartoznak numerikus kifejezések, dátumok és minden olyan kifejezés (akár azonosítók, telefonszámok és e-mail címek is), amely a világ valamely entitására egyedi módon (unikálisan) referál (vö. Tikk et al. 2006: 90–98). A szűkebb értelemben vett tulajdonnév-felismerés, ahogyan a neve is sugallja, kizárólag a tulajdonnevekre koncentrálnak. Többben is írtak már arról, hogy a számítógépes nyelvészet szempontjából pontosan mi tekinthető tulajdonnévnek (pl. Vincze–Farkas 2012). A jelen tanulmányban azonban csak a fentebb említett négy „típust” (személynevek, helynevek, szervezetek nevei és törvények, rendeletek nevei) tekintetem annotálandónak.

2.3 Automatikus névelem-felismerés

A kiválasztott szövegeknek a magyarlanccal és a tulajdonnév-felismerővel való elemztetése után a következő lépés annak eldöntése volt, hogy az egyes tokenekhez megtörtént-e a megfelelő címkék hozzárendelése.

A magyarlanccal esetében az elvárt szófaji címke a tulajdonnév (PROPN) volt, míg a tulajdonnév-felismerő esetében egy bármilyen tulajdonnévi kategória (PER, ORG stb.) jelenléte. Utóbbinál a kategóriák szerinti besorolás helyessége nem képezte a vizsgálat tárgyát.

Az eredmények ismertetése előtt fontos röviden kitérni az elemzők eredeti tanító-korpuszára, mivel a névelem-felismerés erősen doménspecifikus feladat, az elemzők betanítása során használt szövegek tehát erősen kihatnak azok későbbi eredményeire is.

Korpusz	Névelemként annotált tokenek száma	Névelemek száma	Többszavas névelemek	Kategória
Átiratok	56	29	23	PER
	41	22	11	LOC
	69	33	22	ORG
	25	11	7	REG ¹
	191	95	63	A részkorpuszban
Fórumok	95	51	19	PER
	4	4	0	LOC
	6	5	1	ORG
	6	6	0	REG
	111	66	20	A részkorpuszban
Törvény-szövegek	0	0	0	PER
	5	3	2	LOC
	14	8	4	ORG
	6	1	1	REG
	25	12	7	A részkorpuszban
Összesen:	327	173	90	

2. táblázat. Annotált kifejezések (¹REG – Regulations: törvények, rendeletek nevei)

A magyarlanc eredeti tanítókorpusza (a Szeged Treebank) hat eltérő részből épült fel; célja szerint a lehető legkülönbözőbb tematikájú szövegtípusokat volt hivatott reprezentálni (Csendes et al. 2004). A tanításhoz használt korpusz ez esetben tartalmazott jogi szövegeket, bár azoknak csak egy speciális esetét; törvények szövegeit. A tulajdonnév-felismerő, bár ugyanezen a korpuszon lett betanítva, de nem a teljes szövegteszten, csak annak egy részhalmazán: az üzleti rövidhíreken. Ugyanezért tehát a jogi szövegek a tanítókorpusznak nem képezték részét. Ennek az elemzőnek az eredetileg mért pontossága az akkor vizsgált négy kategóriára (PER, ORG, LOC, MISC) együttesen 94,77% volt (Szarvas et al. 2006).

A fentiek tükrében az előzetes várakozás szerint a magyarlanc nagyobb eséllyel volt hivatott megfelelően címkézni a névelemeket a jogi domén szövegeiben. Ezen belül is a legpontosabb eredményeket a törvénytörvények címkézése során vártam.

3 Eredmények

A 3. táblázat a fontosabb tokenszintű mérőszámokat ismerteti. Az adatok számítási alapját a magyarlancról PROPEN, a tulajdonnév-felismerőtől pedig I-PER, I-ORG, I-LOC vagy I-MISC címkét kapott tokenek adták. Nem volt ugyanakkor kritérium, hogy egy adott token mind a két elemző által felismert legyen, a két szoftver kimenetét tehát ebből a szempontból egymástól függetlenül értékeltem.

Az adatokból leolvasható, hogy az elemzők közül a tulajdonnév-felismerő konzisztensen jobb eredményeket ért el a mért számok mindegyikében, valamennyi részkorpusz elemzése során. Az előzetes várakozásokkal ellentétben a törvénytörvények bizonyultak a legkevésbé pontosan elemzettnek, míg a skála másik végpontját az Átiratok adták.

A lényegesebb jellemző hibaforrásokat a továbbiakban az egyes részkorpuszokra lebontva tárgyalom.

Részkorpusz		Névelem- felismerő	magyarlanc morfológia
Fórumok	Pontosság	83,10	69,51
	Fedés	51,75	50,00
	F-érték	63,78	58,16
Átiratok	Pontosság	94,48	63,22
	Fedés	70,26	56,41
	F-érték	80,59	59,62
Törvénytörvények	Pontosság	63,33	26,67
	Fedés	73,08	61,54
	F-érték	67,86	37,21

3. táblázat. Az elemzők által elért tokenszintű eredmények

3.1 Fórumok

Az internetes jogi fórumok bejegyzéseiben előforduló névelemek közül a nicknevek bizonyultak a leginkább problémásnak, mivel alakjuk szerint (kis- vagy nagybetűs írásmód, kiterjedés stb.) nehezen megjósolhatók, éppen ezért potenciálisan nagy kihatásuk lehet az elemzők pontosságára.

A [3] néhány tipikus előfordulást szemléltet olyan nicknevekből, amelyek egy tokent tartalmaznak, a [4] pedig néhány olyat, amelyek több tokenné lettek szegmentálva az automatikus elemzés során.

Habár a fenti példák esetében a besorolás nem minden esetben volt megfelelő, azt fontos hangsúlyozni, hogy az elemzők eredeti tanítókorpuszai nem tartalmaztak olyan szövegeket, amelyekből azok elsajátíthatták volna a nicknevek felismeréséhez szükséges mintákat.

A nickneveknek névelemként való kezelése nyelvészeti szempontból is érdekes probléma. Egyrészt az informális szövegek (jelen esetben a fórumbejegyzések) automatikus feldolgozása már kiindulásként is sokkal összetettebb feladat, hiszen azok gyakran nem tesznek eleget minden szigorú grammatikai konvenciónak, több nyelvtani és helyesírási hibát tartalmazhatnak, ami megnehezíti az elemzők munkáját (vö. Einat et al. 2005). Jogi fórumok esetében pedig, mivel a célközönség erősen limitált, nagy eséllyel fordulnak elő csak a szaknyelvre jellemző terminusok, fordulatok, rövidítések (bár ez utóbbi mind a három vizsgált részkorpuszról elmondható).

[3]

Token	Magyarlanc által társított szófaji címke	Tulajdonnév-felismerő címkéje
55teki55	PROP	O
heidi1115	NUM	O
ObudaFan	PROP	I-ORG

[4]

Token	Magyarlanc által társított szófaji címke	Tulajdonnév-felismerő címkéje
Dr.	NOUN	I-PER
Attika	NOUN	I-PER
Kovács	PROP	I-PER
–	X	I-PER
Béla	PROP	I-PER
–	X	I-PER
Sándor	PROP	I-PER

Másrészt a nicknevek nem tartoznak a klasszikus értelemben vett tulajdonnevek közé. A névelem-felismerési feladatoknak lehetnek alanyai, hiszen ez egy sokkal tágabb definíció, amely éppen ezért megengedőbb is a tárgykörébe tartozó, felismerendő kifejezések körével kapcsolatban. A nemzetközi szakirodalomban például előfordul a márkanevek mellett akár konkrét termékek azonosítási kísérletével foglalkozó cikk is (Yangjie–Aixin 2016), így ebbe a tárgykörbe a nicknevek könnyen beletartozhatnak.

A tulajdonnevekkel szemben támasztott egyik fontos követelmény az identifikáló funkció, vagyis, hogy egy adott név különböző nyelvekben egyformán szerepelhet, egyformán azonosíthatja jelöltjét, nem fordítható (Farkas 2007: 167). A nicknevek a Kripke által megfogalmazott merev jelölő definíciójának (egy merev jelölő kifejezés minden lehetséges világban ugyanazt a dolgot jelöli, amennyiben az a dolog létezik a kérdéses lehetséges világban) is megfelelnek (vö. Kripke 1980), még ha a lehetséges világ terminusa nehezen is hozható összefüggésbe a konkrét fórumbejegyzésekkel.

A legfőbb indok azonban, ami miatt a nicknevek tulajdonnévként azonosíthatók, azt főként az a fajta használatuk, hogy a fórumok alkotta környezetben tulajdonneveket helyettesítenek, azok szerepét töltik be. Használatuk célja, hogy az online környezetben azonosítsák viselőjüket, annak valós nevét helyettesítsék, amely szintén tulajdonnév. Ilyen értelemben tehát maguk is unikusan, egyedi módon referálnak a világ valamely entitására.

Egy másik kérdéses pont lehet, hogy miképpen ítéljük meg a korpuszban az egyes szervezetek név szerinti említéseit. Számos olyan eset fordult elő a szövegben, ahol az ugyanarra a szervezetre referáló kifejezés kétféleképpen fordult elő; egy olyan változatban, ahol a név kezdőbetűje kisbetű, és egy olyanban, ahol nagy kezdőbetűs írásmód érvényesült, például:

- [5] ...ez volt a legfőbb érve a **törvényszéknek**, hogy szabálytalanul lett kézbe-sítve az idézés.
- [6] ...a végzés ellen fellebbezést nyújtsak be a várossal egy megyében található **Törvényszéknek** címezve 3 példányban.

Az ilyen esetekben a két megjelenési forma közül kizárólag a nagy kezdőbetűs írásmóddal rendelkezőt tekintetem tulajdonnévnek (a [6] esetében például konkrétan a Szegedi Törvényszékre történő hivatkozásnak), a kisbetűvel írt változatot az intézmény köznevesült említéseként tartottam számon, tehát nem is került annotálásra felismerendő kifejezéseként, nem lett része a statisztikai adatoknak.

Ennek alapját az a feltételezés adta, hogy a „beszélői” szándék szerint, amennyiben a fórumozó konkrét intézményt említ meg, akkor annak teljes nevét, vagy legalábbis nagy kezdőbetűs írásmódját alkalmazza, amennyiben viszont csak az adott intézmény típusára, szerepkörére (iskola, bíróság, rendőrkapitányság stb.) akar utalni, azt kis kezdőbetűs írásmód alkalmazásával teszi.

3.2 Átiratok

A tárgyalások, kihallgatások írott változatai esetében a legjellemzőbb problémák a mondatkezdő pozícióhoz voltak köthetők a magyarlanc kimenetében. Ezen belül két típushiba volt a legjellemzőbb.

A leiratokban a diskurzusok szegmentálásának fontos eszköze a megszólaló személyének rögzítése minden beszélőváltást követően. Ezek a jelölések tipikusan a megszólalások elején helyezkednek el, továbbá az adott beszélőnek a konkrét perben vagy eljárásban betöltött szerepét rögzítik. Ilyen rövidítés volt a korpuszban például a *V.*, amely a vádlott helyett állt, a *B.* amely a bírói szerepkörre utalt, illetve az *Ü.* vagy éppen *Ü / Ügyv.* amely az ügyvéd megjelöléseként került alkalmazásra. A [7] tipikus és rövid példája annak, amikor a rövidítés nem megfelelően került elemzésre.

A magyarlanc által a részkorpuszban tévesen PROPN címkével ellátott (fals pozitív) esetek jelentős többsége (60.93%, ami 39 esetet fedett le az összes 64-ből) ebből a speciális esetből következett.

A további, nem helyesen megjósolt címkék változatosabb okokra voltak visszavezethetők. Ezek között előfordultak könnyebben magyarázhatók, mint például a szintén mondatkezdő pozícióban elhelyezkedő *Bíró*, amely ugyan a magyar vezetéknevvel való analóg formái megjelenése miatt szintén PROPN címkét kapott, habár a jelen esetben csak az illető perben betöltött szerepére utalt vele a leiratozást végző. Ugyanakkor néhány nehezebben interpretálható címke is felbukkant, mint például a [8] esetében.

A tulajdonnév-felismerő fals pozitív címkéit megvizsgálva szintén változatos esetekkel találkozunk [9]. A [9] a) esetében a téves címke egyértelmű, a [9] b) és c) példák érdekesebbek néhány szempontból.

[7]

1	Ü	Ü	PROPN	Case=Nom Number=Sing	0	ROOT
2	/	/	PUNCT		1	PUNCT
3	Ügyv	Ügyv	PROPN	Case=Nom Number=Sing	1	COORD
4	:	:	PUNCT		1	PUNCT
5	Nem	nem	ADV	PronType=Neg	1	NEG

[8]

1	Ööő	Ööő	PROPN	Case=Nom Number=Sing	6	SUBJ
2	amikor	amikor	ADV	PronType=Rel	4	TLOCY
3	azt	az	PRON	Case=Acc Number=Sing Person=3 PronType=Dem	4	OBJ
4	felvetettünk	felvet	VERB	Definite=Ind Mood=Ind Number=Plur Person=1 Tense=Past VerbForm=Fin Voice=Act	6	ATT

[9]

a)	.	I-ORG
b)	Urat	I-PER
c)	Interneten	I-ORG

Az *internet* esetében az angolban, ahonnan a szó ered, sokáig két írásmódja volt használatos, a kisbetűs *internet* egyszerűen számítógépek, vagy egyéb informatikai eszközök belső hálózatát jelentette, míg a nagybetűs *Internet* a mai értelemben a World Wide Web megfelelője volt, arra mint egyedi, elvont fogalomra utalt (Simpson–Weiner 1989). A magyar helyesírásban a szó eredetileg mindkét változatban elfogadott volt, azaz tulajdonnévi és köznévi használata is „megengedettnek” számított (Deme et al. 1999). Ezt később egyértelműsítették (Laczkó–Mártonfi 2004), és már csak kisbetűs, köznévi használata volt megengedett. Ahogyan ezt Vincze–Farkas (2012: 100–101) több más példával is alátámasztva részletesen kifejti, ez jól mutatja, hogy a tulajdonnévség kérdése túlmutat a formai vagy helyesírási kérdéseken. Emellett arra is jó példa, hogy a nyelv különböző diakrón állapotai során is eltérhet egy-egy konkrét szó tulajdonnévi vagy köznévi megítélése (pl. köznevesülés).

Hasonlóan problémás eset lehet a [9] b), ahol az *urat* megjelölés a tulajdonnév részének tekinthető, bizonyos körülmények között. Ilyen eset lehet például, ha egy terem-

ben több Kovács vezetéknevű ember is tartózkodik. Ilyenkor a név jelentése önmagában homályos lehet, szükség van pontosításra, hogy a jelölet egyértelműen azonosítható legyen. Ha a jelenlevők közül csak egy férfi van, és őt akarjuk megszólítani, az *úr* mint titulus akár tekinthető a tulajdonnév részének.

3.3 Törvények

Habár az átiratokban is megjelent, a törvényszövegek részkorpuszára volt igazán jellemző a jogszabályhelyek tételes említése, mint például:

- *Btk. 236 § (1)* (2012. évi C. törvény a Büntető Törvénykönyvről)
- *Ptk 6: 494§ (2)* (2013. évi V. törvény a Polgári Törvénykönyvről)
- *Tht 1§ (2)* (2003. évi CXXXIII. törvény a társasházakról)

Ezekben az esetekben az adott törvény, rendelet vagy jogszabály neve (*Btk.* stb.) volt tulajdonnévként jelölve, a pontos szakasz, bekezdés, pont stb. nem, mivel az csak a jogszabályi hivatkozások belső struktúrája szempontjából fontos részinformáció.

A részkorpuszon elért alacsonyabb F-érték a magyarlanccal esetében két fő hibatípusra vezethető vissza. Az egyik hibatípus a jogszabályok tagolása során alkalmazott megoldások téves besorolása volt, amelynek a [10] jellemző esete.

A névelem-felismerő esetében a legszembeötlőbb hiba egyértelműen az *1952. évi III. Törvény a Polgári perrendtartásról* esete volt, ahol egyetlen tokent sem jelölt névelemként a szoftver. Habár az elnevezés szerkezete meglehetősen tipikus a törvények nevezéktanában (YYYY évi RÓMAI SZÁM törvény a ... -ról/ról), amely akár reguláris kifejezéssel is leírható, ám ha a rendszer eredeti tanítókorpuszában nem szerepelt hasonló adat, akkor nem bejósolható.

[10] „(3a) A (2) bekezdés szerinti nyilvántartás a (3) bekezdés c) – i) pontja szerinti adatok...”

1 ((PUNCT	_	0 PUNCT
2 3a 3a	PROP	Case=Nom Number=Sing	8 ATT
3))	PUNCT	_	2 PUNCT
4 A A	PROP	Case=Nom Number=Sing	2 COORD
5 ((PUNCT	_	6 PUNCT

3.4 Függőségi nyelvtani elemzés

A teljes kép érdekében az utolsó tervezett szint a magyarlanccal függőségi nyelvtani címkézésének vizsgálata volt.

A 4. táblázat a szintaktikai viszonyokban fellelhető NE élécímke jelenléte szerint számolt tokenszintű értékeket mutatja. Ugyanakkor fontos megjegyezni, hogy a fenti értékek nem tekinthetők reprezentatívnak, mivel a kézi annotálás során a vártnál kevesebb, több tokenből álló névelemet sikerült csak találni a jelen tanulmányhoz elemzett korpuszrészben. Az 5. táblázat a kézi annotációban megjelenő névelemeket összegzi.

A második oszlop a magyarlanccal által NE élécímkevel összekapcsolt, több tokenből álló névelemeket jelöli, a harmadik oszlop a kézzel annotáltakat.

Az adatok ritkasága (mind a manuális, mind az automatikus annotáció alapján) egyelőre nem teszi alkalmassá azokat egy pontos statisztikai elemzésre, így e téren további vizsgálatok szükségesek, a szintaktikai elemzés hatékonysága nem ítéhető meg pontosabban.

Részkorpusz		
Fórumok	Pontosság	80,75
	Fedés	70,00
	F-érték	74,99
Átiratok	Pontosság	76,92
	Fedés	66,67
	F-érték	71,43
Törvényszövegek	Pontosság	100,00
	Fedés	57,14
	F-érték	72,73

4. táblázat. A magyarlanc szintaktikai annotációjának eredményessége a tulajdonnév-felismerésben

Részkorpusz	A magyarlanc szintaktikai elemzésében megjelenő névelemek	Manuálisan jelölt több tokenes névelemek
Átiratok	23	63
Fórumok	15	20
Törvényszövegek	7	7

5. táblázat. Referenciaadatok a szintaktikai elemzés szintjéről

4 Következtetések

A jelen tanulmányban a névelem-felismerés hatékonyságát vizsgáltam meg magyar nyelvű jogi szövegek esetében. Ennek érdekében egy tulajdonnév-felismerő szoftver és a magyarlanc kimenetét hasonlítottam össze manuális annotációval a Miskolc Jogi Korpusz kiválasztott részletein. A megvizsgált két szoftver kimenetében azonosítottam a legtipikusabb hibaforrásokat a megfelelő szófaji címkézés (magyarlanc) és a névelem-felismerés (tulajdonnév-felismerő) szempontjából.

A feladat doménspecifikusságával kapcsolatban a vizsgálatok alátámasztották, hogy mindhárom elemzett részkorpusz/szövegtípus esetében vannak olyan sajátosságok, amelyek kezelése még nem megoldott a vizsgált rendszerekben. A feltárt hibaforrások alapján a szükséges megszorítások visszaillesztésével mindkét rendszer hatékonyabbá és pontosabbá tehető, valamint esetleges későbbi névelem-felismerő rendszerek határfoka is javítható az eredmények figyelembevételével.

A vizsgálat során kiderült, hogy a több tokenes névelemek sokkal alacsonyabb arányban vannak jelen a kiválasztott szövegekben, mint ami egy pontos statisztikai elemzéshez szükséges. Ennek megoldásához szükséges a vizsgálat megismétlése egy nagyobb szövegtesten.

Távlati cél az elemzés kiterjesztése a Miskolc Jogi Korpusz mind a hat részkorpuszára annak érdekében, hogy a jogi szövegeken végzett névelem-felismerés azok minél szélesebb spektrumán tesztelhető legyen.

Irodalom

- Csendes, D. – Csirik, J. – Gyimóthy, T. 2004. The Szeged Corpus: A POS Tagged and Syntactically Annotated Hungarian Natural Language Corpus. In: Sojka, P. – Kopeček, I. – Pala, K. (szerk.). *Lecture Notes in Artificial Intelligence* (Subseries of Lecture Notes in Computer Science) 3206: 41–47.
- Deme L. – Fábrián P. – Tóth E. (szerk.) 1999. *Magyar helyesírási szótár*. Budapest: Akadémiai Kiadó.
- Einat, M. – Richard, C. W. – William, W. C. 2005. Extracting personal names from emails: Applying named entity recognition to informal text. *Human Language Technology / Empirical Methods in Natural Language Processing*: 443–450.
- Farkas T. 2007. A tulajdonnevek fordíthatóságáról és napjaink fordítási hibáiról, közsók és tulajdonnevek példáján. *Névtani Értesítő* 29: 167–188.
- Kripke, S. 1980. *Naming and Necessity*. Cambridge, Massachusetts: Harvard University Press.
- Laczkó K. – Mártonfi A. 2004. *Helyesírás*. Budapest: Osiris.
- Lenci, A. – Montemagni, S. – Pirrelli, V. – Venturi, G. 2009. Ontology learning from Italian legal texts. In: *Proceeding of the 2009 Conference on Law, ontologies and the Semantic Web: Channelling the Legal information Flood*: 75–94.
- Linguistic Data Consortium. ACE (automatic content extraction) English annotation guidelines for entities. 2006. Elérhető: <https://www ldc.upenn.edu/collaborations/past-projects/ace>, Version 5.6.6 2006.08.01. (letöltés ideje: 2019. 01. 10.)
- Móra Gy. – Vincze V. – Zsibrita J. 2011. Szófaji kódok és névelemek együttes osztályozása. In: Tanács A. – Vincze V. (szerk.) *VIII. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged: Szegedi Tudományegyetem. 131–142.
- Quaresma, P. – Gonçalves, T. 2010. Using Linguistic Information and Machine Learning Techniques to Identify Entities from Juridical Documents. In: Francesconi, E. – Montemagni, S. – Peters, W. – Tiscornia, D. (szerk.) *Number 6036 in Lecture Notes in AI*. Springer-Verlag. 44–59.
- Simon E. – Farkas R. – Halácsy P. – Sass B. – Szarvas Gy. – Varga D. 2006. A HunNER korpusz. In: Tanács A. – Vincze V. (szerk.) *IV. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged: Szegedi Tudományegyetem. 373–376.
- Simon E. 2008. Nyelvészeti problémák a tulajdonnév-felismerés területén. In: Sinkovics B. (szerk.) *LingDok 7. Nyelvész-doktoranduszok dolgozatai*. Szeged: Szegedi Tudományegyetem, Nyelvtudományi Doktori Iskola.
- Simon, E. 2017. The Definition of Named Entities. In: Gyuris, B. – Mády, K. – Recski, G. (szerk.) *K + K = 120. Papers dedicated to László Kálmán and András Kornai on the occasion of their 60th birthdays*. Budapest: Research Institute for Linguistics, Hungarian Academy of Sciences (RIL HAS): 1–18.
- Simpson, J. A. – Weiner, E. S. C. 1989. *The Oxford English Dictionary*. Oxford: Clarendon Press.
- Surdeanu, M. – Nallapati, R. – Manning, C. D. 2010. Legal claim identification: Information extraction with hierarchically labeled data. In: *Proceedings of the LREC 2010 Workshop on the Semantic Processing of Legal Texts (SPLeT)*. Malta, May 2010
- Szarvas Gy. – Farkas R. – Kocsor A. 2006. A Multilingual Named Entity Recognition System Using Boosting and C4.5 Decision Tree Learning Algorithms. In: *The Ninth International Conference on Discovery Science LNAI 4265*: 267–278.
- Tikk D. – Farkas R. – Kardkovács Zs. T. – Kovács L. – Répási T. – Szarvas Gy. – Szaszko S. – Vázsonyi M. 2006. *Szövegbányászat*. Budapest: Typotex.
- Vincze V. – Farkas R. 2012. Tulajdonnevek a számítógépes nyelvészetben. *Általános Nyelvészeti Tanulmányok XXIV*: 97–119.
- Vincze V. 2018. A Miskolc Jogi Korpusz nyelvi jellemzői. In: Szabó M. (szerk.) *A törvény szavai*. Miskolc: Bíbor Kiadó. 9–36.
- Yangjie, Y. – Aixin, S. 2016. Mobile phone name extraction from internet forums: a semi-supervised approach. *World Wide Web* 19(5): 783–805.

Egyet nem értések fonetikai jellemzői középiskolások beszédében

Weidl Zsófia

ELTE BTK Nyelvtudományi Doktori Iskola
weidlzsofi@gmail.com

Kivonat: A kutatás célja a verbális egyet nem értések fonetikai és pragmatikai elemzése középiskolások spontán beszédében, vitahelyzetben. Az egyetértés általában a preferált, míg az egyet nem értés inkább a nem preferált beszédcselekvések közé sorolódik (Schegloff et al. 1977). Georgakopoulou (2001) szerint azonban a ritualizált vagy barátságos konfliktus gyakran megjelenik az egymással közeli kapcsolatban álló, egyenrangú beszédpartnerek interakciói során.

A vizsgálat során egyetértő és egyet nem értő beszélgetéseket vetek össze. A 10 társalgás elemzése az egyszerre beszélés fonetikai paramétereit mellett kitér a tempó és az alaphérvencia értékeire is. A diákok egyet nem értései során az egyetértésekben megjelenő értékektől eltérő alaphérvenciával és artikulációs tempóval nyilatkoztak meg.

1 Bevezetés

Az emberek közötti szóbeli kommunikáció során a beszédpartnerek véleménye nem mindig egyezik, s ezt gyakran egymás tudtára is adják. Az ilyen típusú egyet nem értéket meghatározhatjuk egy állítás igazságtartalmának megcáfolásaként, egy szándékkal vagy egy cselekedettel kapcsolatos állásfoglalásként, értékítéletként (Bándli 2009). Mindegyikben közös a beszédpartnertől eltérő vélemény kifejezése. Koczogh (2012) meghatározása alapján a verbális egyet nem értés olyan vélemény kifejezése, melynek tartalma vagy illokúciós ereje részben vagy teljesen ellentmond egy azt megelőző megnyilatkozásnak.

Az egyet nem értés jelenségét sok kutatás vizsgálta már a pragmatika felől. Elsőként Goffman (1967) arculatelméletében jelent meg. Az arculat olyan énkép, amelyet társadalmilag jóváhagyott tulajdonságok körvonalaznak. Ezen jelenséghez járul az arculatmunka, amely bármely olyan cselekedet, amit valaki annak érdekében végez, hogy összhangban maradjon az általa mutatott arculattal. A találkozásainkat igyekszünk úgy alakítani, hogy mind a saját, mind a partnerünk önbecsülése és az ehhez kapcsolódóan kialakított arculat megmaradjon. Tehát az arculatóvás az a tevékenység, amellyel fenntartjuk másokban azt a benyomást, hogy nem veszítettük el az arculatunkat. Ugyanakkor a kialakított arculat folyamatos óvása alapvető társadalmi korlátozás, hiszen nem

Weidl Zsófia: Egyet nem értések fonetikai jellemzői középiskolások beszédében. In Váradi Tamás (sorozatszerkesztő), Ludányi Zsófia, Grácz Tekla Etelka (szerkesztő): *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2019. XIII. Alkalmazott Nyelvészeti Doktoranduszkonferencia*. Budapest: MTA Nyelvtudományi Intézet. 2019. 138–147. DOI: 10.18135/Alknyelvdok.2019.13.11

csupán saját arculatunk óvása a cél, hanem a beszélgetőpartner arculatvesztésének elkerülése is. Ezekből a társadalmi szabályokból az következik, hogy a másik fél véleményét inkább jóváhagyjuk, s a kapcsolatfenntartás érdekében szóban inkább elismerjük azt is, amivel alapvetően nem értünk egyet. Tehát az arculatvédő tevékenységek során nem mondunk ki olyan tényeket, amelyek ellentmondhatnak a másik állításainak, így kerüljük az egyet nem értéseket (Goffman 1967).

A következő, az egyet nem értések vizsgálatával kapcsolatos fontos elmélet, Brown és Levinson (1987) udvariasságelmélete. Ebben meghatározzák, hogy a tartózkodó arculat kialakítása a másoktól való függetlenségünket szimbolizálja, míg a közelítő arculat kialakításakor a beszédpartnerrel való egyetértésre törekszünk. Megállapítják, hogy a zavartalan interakció feltétele egymás arculatának elfogadása és megőrzése, és minden illokúciós aktus arculatfenyegető aktusként értelmezhető. Ezekből fakadóan az egyet nem értés kifejezése az arculatfenyegetés egy típusa, amikor a másik fél arculatáról negatív véleménnyel vagyunk. Tehát az arculatfenyegető aktusok elkerülésének egyik stratégiája az egyetértésre való törekvés (Brown–Levinson 1987).

A konverzációelemzés keretei közt is végeztek vizsgálatokat az egyet nem értésekre vonatkozóan. Sacks (1973) megállapította, hogy az egyetértésre való törekvés alapvető társalgási szabály, amelyet egyet nem értésünk kifejezésekor megszegünk, s az interakciók során a beszédpartnerek jellemzően arra törekednek, hogy ezen szabályt betartsák. Tehát társadalmilag inkább támogatott az egyet nem értések elkerülése, mint valós vélekedésünk nyílt megfogalmazása. Preferált és nem preferált beszédcselekvések elkülönítésekor az egyetértések általánosan preferáltak, míg az egyet nem értések a leggyakrabban nem preferáltak. A válaszlépések tekintetében az egyetértések preferáltak, az egyet nem értés pedig sokszor késleltetve, a válaszlépés belsejében vagy végén jelenik meg (Sacks 1973).

Szintén a preferált és nem preferált válaszlépések elkülönítésével kapcsolatban Schegloff és munkatársai (1977) megállapították, hogy a preferált válaszlépések szerkezetileg egyszerűek, míg a nem preferált válaszlépéseket strukturális komplexitás jellemzi, például hosszabb szünet, hezitáló bevezetés, kérdő formájú ismétlés; tehát bírnak fonetikailag is mérhető jegyekkel.

A későbbi egyet nem értéseket érintő vizsgálatok nem anyanyelvi beszélgetésekben figyelték meg a jelenséget. Pearson (1986) kutatásában az egyet nem értések kifejezésére szolgáló pragmatikai eszközöket vizsgálta az angol második nyelvként beszélők körében. Megállapította, hogy ezek a beszélők sokszor explicit módon fejezik ki véleményüket olyan helyzetekben is, amelyekben ez természetellenesnek hathat egy anyanyelvi beszélő számára, valamint gyakran nem adnak hangot egyet nem értésüknek olyan helyzetekben, amelyekben anyanyelvükön hangot adtak volna (Pearson 1986).

García 1989-es kutatása az egyet nem értések típusait vizsgálta, valamint már kulturális különbségeket is igyekezett megállapítani a jelenséggel kapcsolatban. Adatközlői amerikaiak és venezuelaiak voltak, az ő angol nyelvű egyet nem értéseiket vizsgálta. A feladat különböző arculatfenyegető aktusok eljátszása volt, mert így elő tudta idézni az alapvetően nem preferált egyet nem értést. Következtetései között szintén szerepelt az a megállapítás, hogy az egyet nem értések során komplexebb stratégiákat alkalmaztak azok a beszélők, akik anyanyelvükön nyilatkoztak meg. A nem anyanyelvi beszélők megnyilatkozásaira a szerkezeti egyszerűség volt jellemző (García 1989).

Mivel ezen kísérletek eredményeiből fakadóan megállapítható, hogy az egyet nem értésünket nem mindig adjuk könnyen a beszélőpartner tudtára, ezért olyan szituációk kell létrehozni, amelyben ez mégis megtörténik. A további kutatásokban az egyet nem

értés jelenségét szintén olyan módon vizsgálták, hogy igyekeztek olyan szituációkat előállítani, amikor az egyet nem értés válik a preferált beszédcselekvéssé.

Pomerantz (1984) az egyet nem értések típusait vizsgálta együttműködő beszélgetésekben. Ő is megállapította, hogy alapvetően az egyetértés a preferált beszédaktus, amely egyszerű szerkezetű; míg az egyet nem értés nem preferált, és strukturális komplexitás jellemzi, valamint gyakran homályos megfogalmazású és késleltetett. Ugyanakkor megállapította azt is, hogy az egyet nem értés bizonyos kontextusokban preferált megnyilvánulás is lehet, például olyan megjegyzések után, amikor a beszédpartner negatív állítást fogalmaz meg önmagáról. Amikor azonban az egyet nem értés válik preferált aktussá, a látszólagos egyetértés az egyet nem értés jele lehet (Pomerantz 1984).

Kotthoff (1993) az ellentmondás kifejezésének jellemzőit vizsgálta vitahelyzetben. Megállapítása szerint a barátságos beszélgetések egyik fő szervezőereje a konszenzusra való törekvés, azonban vitahelyzetben megváltoznak a preferenciaviszonyok, és az ellentmondás lesz az elvárt válaszlépés, ilyenkor fontos gyorsan és koherens módon ellentmondani.

2000-ben Rees-Miller egyetemi szituációkban vizsgálta az oktatók és a hallgatók egyet nem értési stratégiáit; az egyet nem értések megvalósítási módját a társadalmi távolság és az egyet nem értés mértékének függvényében. Az egyet nem értésekről megállapította, hogy a beszélő a beszédpartner által kimondott vagy feltételezett pozíciót nem tartja igaznak, és olyan megnyilatkozással válaszol, amelynek proposíció tartalma más. A pedagógiai kontextusban sokszor magának a specifikus helyzetnek, valamint a cselekvés (tanítás-tanulás) céljának nagyobb hatása van az egyet nem értés nyelvi megformálására (Rees-Miller 2000).

Kakava (2002) szintén pragmatikai szempontból vizsgálta az egyet nem értéseket görög nyelvű beszélők körében. Megállapításai szerint az egyet nem értés mindig válaszlépés valamire, és az egyet nem értések kifejezőmódjára hatással van a beszélők kora, neme és a köztük levő ismertség mértéke. Mindezek mellett kulturális különbségek is felfedezhetők a jelenséggel kapcsolatban, léteznek olyan kultúrák, ahol a vitatkozás kifejezetten elvárt beszédaktus. Valamint fontos megállapításai közé tartozik, hogy habár alapvetően az egyet nem értések nem preferáltak, az egymással közeli kapcsolatban álló beszélők között az ellentmondás megengedett jelenség, preferálttá válhat.

Georgakopoulou (2001) kutatási kérdései közé tartozott, hogy az egyet nem értések mennyiben nem preferáltak és arculatfenyegetők, illetve milyen módon fejezik ki a beszélők. Adatközlői egymást jól ismerő fiatalok voltak, akiknek a vitaszituációkban elhangzó beszélgetéseit vizsgálta. Megállapította, hogy a ritualizált vagy barátságos konfliktus gyakran megjelenik az egymással közeli kapcsolatban álló, egyenrangú beszédpartnerek – jellemzően kamasz baráti társaságok – interakciói során. Ezen túlmenően azt is kimutatta, hogy az intimitás mértéke hatással van az ellentmondás hatásának megítélésére: a magas fokú intimitás a nem preferált tartományból a preferált tartományba helyezheti az egyet nem értést. Ugyanakkor azt is kijelenti, hogy a diskurzuszstratégiák közösség szintű általánosítása túlságosan egyszerűsítő, inkább az egyéni különbségek jellemzőek.

Edstrom 2004-es kutatása során megállapította, hogy vannak kulturálisan meghatározott hasonlóságok is a beszédaktusok megvalósításában, például a preferált és nem preferált lépések közti válogatáskor. Ő is leszögezi azonban, hogy nem tekinthetünk el az egyéni variációktól sem, amelyek személyes preferenciákon alapulnak.

A magyar nyelvű kutatások az egyet nem értésekre vonatkozóan szintén pragmatikai szempontból tettek megállapításokat. Bándli (2011) vizsgálatában ugyancsak egyet nem értéseket kiváltó szituációkban figyelte meg adatközlőit. Eredményei alapján az

egyét nem értékek megvalósítását szolgáló stratégiák választása nagy mértékben szituációfüggő, s ezen stratégiaválasztás összefügg a társadalmi távolsággal.

Koczogh 2012-es kutatásában egyetemisták verbális egyet nem értéseit vizsgálta szociopragmatikai szempontból. Eredményei alapján elmondható, hogy az egyet nem értékek vizsgálatakor nem egyszerű általános következtetéseket levonni, ugyanis az egyéni különbségek, jellemzők igencsak meghatározók. Ezekon kívül fontos változóként funkcionál a kor, a nem és az ismertségi fok.

Kutatásomban középiskolás fiatalok verbális egyet nem értéseinek pragmafonetikai elemzését végeztem el. Hipotéziseim szerint 1. a diákok egyet nem értékeik során az egyetértésekben megjelenő paraméterektől eltérő alaphangmagassággal és artikulációs tempóval nyilatkoznak meg; 2. az egyszerre beszélés aránya nagyobb az egyet nem értékek során.

2 Módszertan

A kutatás résztvevője 10 fő volt, 15 és 19 év közötti középiskolás fiúk. Mindannyian ugyanazon gimnáziumba járnak, s szociális helyzetük megközelítőleg azonos, szociális távolság nincs köztük.

A kísérletet egy magyar nyelvórába ágyazott retorikai feladatként valósítottam meg. A gyerekek előre meghatározott módon párokat alkottak. A diákok párokba állításához előzetesen szociometriát (Mérei 1971/2006) készítettem a kutatásban részt vevő diákok osztályaiban, hogy a köztük levő viszonyokat pontosabban megismerjem. A szociometria elkészülte után olyan diákokat választottam vitapartnereknek, akik közel megegyezően magas szimpátiaviszonyban, ismertségi fokon állnak, tehát barátok.

Ezután minden pár kapott két témát, amelyekről vitázniuk kellett. A vitaszituáció megteremtésével idéztem elő az egyet nem értő szituációkat. Az adatközlőknek az egyik vita során ellenkező, míg a másik során azonos álláspontot kellett képviselniük, s az egyetértő beszélgetések szolgálták a vizsgálat során kontrollhelyzetként, ezekhez viszonyítva elemeztem az egyet nem értékeket. Az egyes témák meghatározása után mindkét fél utasítást kapott arról is, hogy pontosan mely álláspont mellett kell érvelniük, s egyikük sem tudta, hogy mi a másik feladata. Az irányított állásfoglalás után a feleknek volt idejük arra, hogy érveket dolgozzanak ki a megadott álláspont mellett. Miután a vitapartnerek megkapták a témájukat, a következő instrukció hangzott el: *Csendben gyűjts össze 5-6 érvet az állítás mellett/ellen, majd utána beszélgetés során győzd meg velük a párod!*

A témák meghatározásához előzetesen kérdőíves felmérést végeztem a diákok körében. Ezen kérdőívben azt térképeztem fel, hogyan vélekednek több, a köztudatban jelenlevő, megosztó témáról, valamint olyan témákat is felvettem, amelyek az iskolai élethez kapcsolódnak. A kérdőívekben Likert-skálán (Likert 1932) 1-től 5-ig kellett jelölniük a gyerekeknek, hogy mely állítással milyen mértékben értenek egyet. Egy-egy főbb témához több állítás is tartozott, így a biztosabb eredmény érdekében többször kellett véleményt nyilvánítaniuk ugyanarról a témakörrel a diákoknak. Ennek segítségével kiszűrtem azokat a témákat, amelyek nem váltottak ki érdeklődést a diákokból, és az eredmények alapján választottam ki a főbb vitatémákat. A kérdőív alapján a kiválasztott vitatémák közé egyaránt kerültek olyanok, amelyek közéleti, megosztó témakörök, s olyanok is, bár kisebb számban, amelyek inkább az iskolai élethez kapcsolódnak. A vitapartnerek kialakítása során arra törekedtem, hogy mindenki a valós álláspontját képviselje, hogy minél eredményesebben nyilvánítsanak véleményt.

A vitákról hangfelvételt készítettem, melyeket omnidirekcionális kondenzátoros fej-mikrofonnal és kétsatornás külső hangkárttyával vettem fel. Ez lehetővé tette, hogy az egyes beszélők hanganyagát külön is elemezzem, és bár a másik résztvevő beszéde is behallatszott az adott beszélő felvételébe, annak intenzitása jóval kisebb volt, ezáltal az egyszerre beszélések pontosabb elemzése is lehetővé vált.

Az elkészült felvételeket a Praat szoftver (Boersma–Weenink 2012) segítségével beszédszakasz- és szószinten annotáltam. A hanganyagokon elsősorban a szupraszegmentális szerkezetet vizsgáltam meg, valamint az egyszerre beszélés jelenségét.

Elsőként az egyetértéseket tartalmazó hanganyagokat elemeztem; megállapítottam a beszélők átlagos alaphangfrekvenciáját, valamint kiszámítottam a beszélők átlagos artikulációs tempóját. Ezen értékekhez viszonyítva mindezt megvizsgáltam az egyet nem értéseket tartalmazó beszélgetésekben is, továbbá azt is elemeztem, hogy ezek milyen összefüggésben vannak az egyszerre beszélés jelenségével.

3 Eredmények

Az eredmények bemutatása során ismertetem a hanganyagok időtartamát, a beszélők artikulációs tempóját és alaphangfrekvenciáját, ezeken túl az egyszerre beszélés jelenségéről teszek megállapításokat.

3.1 A viták időtartama

Az egyetértő és az egyet nem értő beszélgetések hossza, valamint az egyszerre beszélések időtartama látható az 1. ábrán. Egymás mellett láthatóak az egy pártól származó egyetértő és egyet nem értő hanganyagok adatai.

A viták, beszélgetések tiszta beszédidejét tekintve az egyet nem értést tartalmazó beszélgetések három pár esetében hosszabb időtartamúak, két pár esetében rövidebbek voltak. Az első és második beszélőpárnál voltak a leghosszabb időtartamú viták, beszélgetések, mindkettő pár esetében az egyet nem értő beszélgetések voltak megközelítőleg 20%-kal hosszabb időtartamúak. A harmadik és a negyedik pár esetében a beszélgetések rövidebb időtartamúak voltak, mint a többi párnál, s az egyetértő és az egyet nem értő viták hossza közel megegyező hosszúságú volt; a harmadik párnál minimálisan az egyet nem értések voltak a hosszabbak, míg a negyedik párnál minimális különbséggel az egyetértések voltak hosszabb időtartamúak. Az ötödik pár esetében az egyetértő beszélgetések voltak hosszabb időtartamúak. Tehát nem lehet egyértelműen kijelenteni minden pár esetében, hogy az egyetértések során könnyebben, gyorsabban lezárták volna a beszélgetést, ez inkább az adott témától is függött.

Az egyszerre beszéléseket tekintve megállapítható, hogy a negyedik pár esetében nem volt ilyen jelenség, míg a harmadik párnál csak az egyetértésekkor jelent meg. Ugyanakkor az is elmondható, hogy a párok produkálták a legrövidebb időtartamú vitákat, tehát ezen eredmény nem meglepő. Jellemző volt rájuk mindkét kondícióban, hogy kivárták, amíg beszédpartnerük befejezi mondandóját, nem vágta egymás szavába. A negyedik párnál felmerülő rövid egyszerre beszélés is egy hosszabb szünet után következett, amikor mindketten hosszabban arra vártak, hogy a másik beszélő szólaljon meg, s végül egyszerre kezdték el a beszédet, majd el is hallgattak.

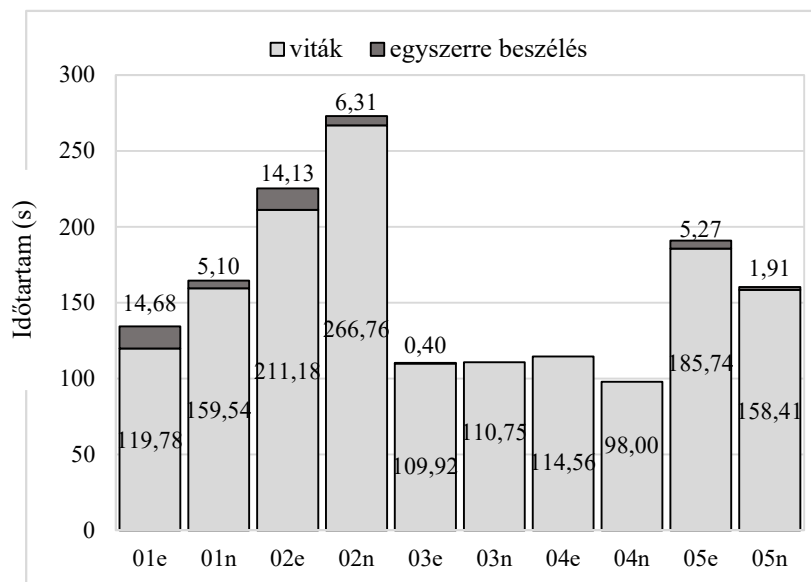
Az egyszerre beszélések aránya az első pár esetében volt a legnagyobb mindkét kondícióban, 10,92% egyetértésekkor és 3,1% egyet nem értésekkor. A többi pár esetében

ezen értékek alacsonyabbak voltak. Általában helyeslő vagy ellenkező közbevetések voltak mindkét beszédtypusnál, illetve egyetértésekkor gyakran egymás gondolatmenetének folytatása, egy elkezdett történet befejezése során jelent meg az egyszerre beszélés jelensége.

A második párnál az egyszerre beszélés aránya 6,27% volt egyetértésekkor és 2,31% egyet nem értésekkor. Ezen arányok nem túl magasak, tekintve, hogy e pár tiszta beszédideje volt a leghosszabb. Az ő esetükben is általában helyeslő megnyilatkozásokkor jelent meg az egyszerre beszélés.

Az ötödik pár esetében az egyszerre beszélés aránya 2,76% volt egyetértésekkor, és ennél alacsonyabb, 1,19% egyet nem értésekkor. Esetükben is általában helyeslő megnyilatkozásokkor vagy a másik mondandójának kiegészítésekor jelent meg az egyszerre beszélés.

A jelenség időtartamaránya minden pár esetében az egyetértő beszélgetéseknél volt magasabb. Ezek általában helyeslő megnyilatkozások, míg az egymás szavába vágás ritkább az egyet nem értéseknél; a negyedik párnál kifejezetten jellemző volt az, hogy megvárták a beszélők, amíg a másik befejezi a gondolatmenetét, érvét.



1. ábra. A viták időtartama és az egyszerre beszélés aránya

3.2 Artikulációs tempó

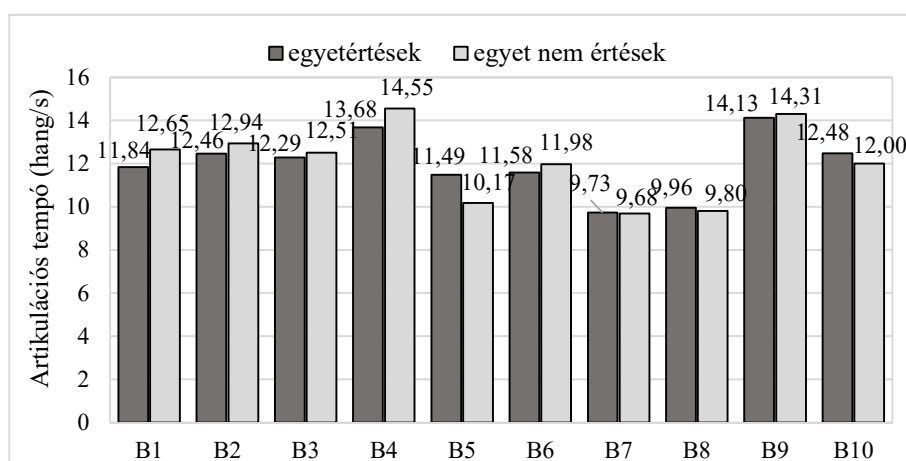
A 2. ábrán látható a 10 beszélő artikulációs tempója egyetértéseik és egyet nem értéseik esetén.

A beszélők átlagos artikulációs tempója 9,73 hang/s és 14,13 hang/s között helyezkedett el egyetértések során, míg egyet nem értésekkor 9,68 hang/s és 14,55 hang/s között. Hat beszélő esetében az artikulációs tempó gyorsabb volt az egyet nem értéseknél, mint az egyetértéseknél, négy beszélő esetében ez fordított volt. Ezen különbségek

egy-egy pár esetében nem voltak oly eltérők, inkább beszélőnkénti eltéréseket fedezhetünk fel az átlagos artikulációs tempo tekintetében.

A B9 és B4 beszélők artikulációs tempója volt a legmagasabb mind az egyetértések-nél, mind az egyet nem értékek-nél, 14 hang/s körüliek, míg beszédpartnereiké 12 hang/s körüliek mindkét kondícióban. A B7 és B8 beszélők artikulációs tempója volt a legalacsonyabb, közel egyenlők értékek; ők vitapartnerek voltak. Feltételezhető, hogy valamelyik beszélő alkalmazkodott a párjához. Hasonló alkalmazkodás feltételezhető a B1–B2 pár esetében, valamint a B5–B6 pár esetében, közel egyenlő artikulációs tempo volt megfigyelhető esetükben is.

Továbbá elmondható még, hogy az egyetértések és az egyet nem értő viták esetében is a párok ugyanazon tagja beszélt gyorsabb artikulációs tempóval, még ha közelítettek is egymáshoz a beszélőpárok artikulációs tempóértékei, a köztük levő reláció mindkét kondíció esetében megmaradt.



2. ábra. A beszélők artikulációs tempója

3.3 Alapfrekvencia

A 3. ábrán láthatóak a beszélők átlagos alapfrekvencia-értékei egyetértések és egyet nem értések esetén.

Elmondható, hogy a beszélők átlagos alapfrekvenciája 101,02 Hz és 130,53 Hz között mozgott az egyet nem értések során, és 111,45 Hz és 138,91 Hz között az egyetértések során. Tehát az egyet nem értő viták alkalmával alacsonyabb alapfrekvenciával nyilatkoztak meg a beszélők, s e megállapítás egyénenként is igaz a tízből kilenc adatközlő esetében. Csak B6 beszélő esetében volt magasabb az alapfrekvencia az egyet nem értések során.

A legmagasabb alapfrekvencia-értékek B3 és B4 beszélők esetében mérhetőek, akik beszédpartnerek voltak. Míg a legalacsonyabb értékek B9 és B10 beszélők esetében voltak találhatóak, akik szintén egy párt alkottak. Feltételezhető, hogy az alapfrekvencia tekintetében is lehetséges, hogy a diákok valamilyen módon alkalmazkodtak beszédpartnerükhöz, ugyanis B1 és B2 párok, valamint B5 és B6 vitapartnerek esetében is

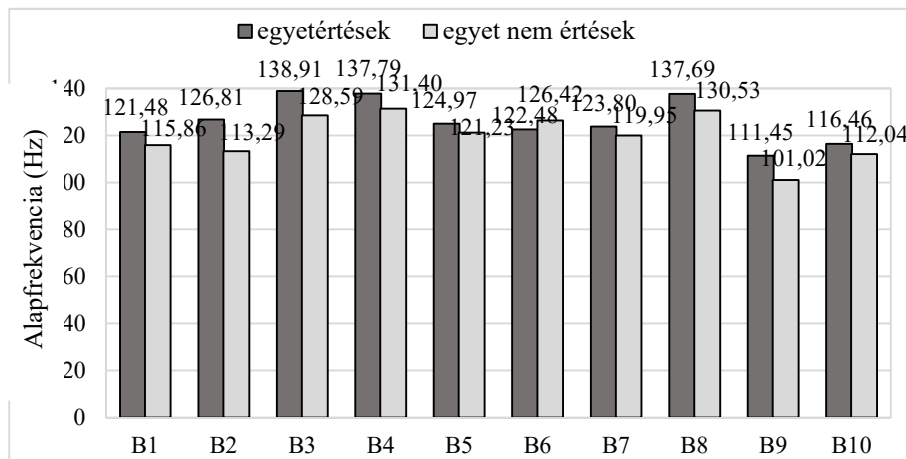
egymáshoz közeli alapfrekvencia-értékek voltak mérhetőek mind az egyetértéseknél, mind az egyet nem értéseknél.

3.4 Alapfrekvencia egyszerre beszéléseknél

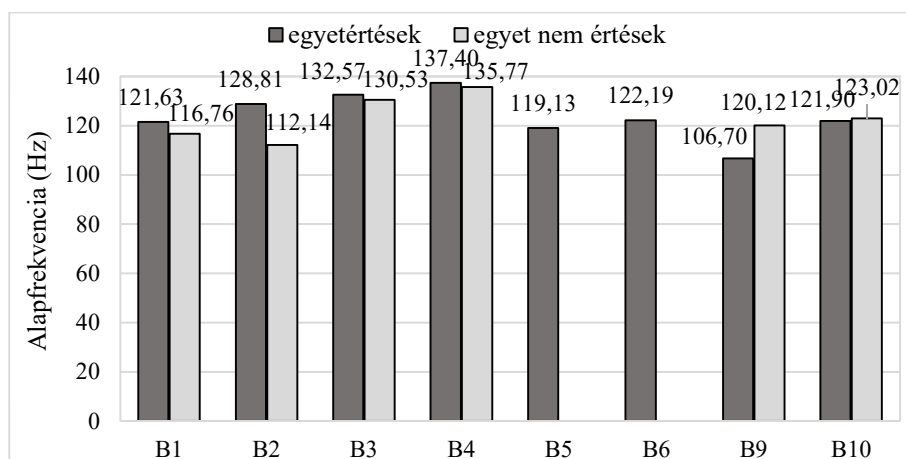
Az 4. ábrán látható a vitapartnerek átlagos alapfrekvenciája egyszerre beszélésekkor az egyetértések és az egyet nem értések esetében. Ezen értékek mérése azokon a szakaszokon történt, ahol nem jelenik meg a másik beszélő beszéde.

B7 és B8 beszédpartnereknél nem fordult elő egyszerre beszélés. Míg B5 és B6 beszélőknél, akik szintén párt alkottak, csupán az egyetértésekkor tűnt fel a jelenség, mindkettőjük esetében minimálisan alacsonyabb az alapfrekvencia értéke egyszerre beszéléskor.

B9 és B10 vitapartnereknél az egyszerre beszélések esetében magasabb alapfrekvencia volt mérhető egyet nem értéskor, mint egyetértéskor, holott az átlagos alapfrekvenciájuk esetében az arány épp fordított volt (vö. 3. ábra).



3. ábra. A beszélők átlagos alapfrekvenciája



4. ábra. A beszélők átlagos alapfrekvenciája egyszerre beszélések esetén

B1 és B2, valamint B3 és B4 beszélőpároknál minimálisan ugyan, de inkább emelkedett az alapfrekvencia egyszerre beszélésekkor az átlagos alapfrekvenciához képest, de arányuk az egyetértések és az egyet nem értések viszonylatában nem változott, ugyanúgy az egyet nem értések során volt magasabb az alapfrekvencia értéke egyszerre beszélésekkor is.

4 Következtetések

A diákok egyet nem értéseik során az egyetértésekben megjelenő paraméterektől eltérő alaphangmagassággal és artikulációs tempóval nyilatkoztak meg – az f_0 inkább alacsonyabb volt, az artikulációs tempó pedig általában magasabb, ugyanakkor viszonylag nagy az egyéni különbségek szerepe.

Ezen különbségek összefüggnek egyéb pragmatikai jellemzőkkel. Az egyszerre beszélés aránya nem minden beszélő esetében nagyobb az egyet nem értések során. Az egyetértésekkor előforduló egyszerre beszélések általában helyeslő megnyilatkozások, az egymás szavába vágás ritkább az egyet nem értéseknél.

Az eddigi eredmények alapján feltételezhető, hogy a beszédpartnerek valamilyen módon alkalmazkodnak egymáshoz, de ennek bizonyítása és okainak pontosítása még vizsgálatot igényel.

A kutatás a továbbiakban kiterjeszhető a nemek közti eltérések vizsgálatára is (lány-lány és lány-fiú párok vizsgálata is), valamint vizsgálhatóak még egyéb szuprasegmentumok (szünetek, intenzitás, beszéddallam) és pragmatikai jellemzők (diskurzuszjelölők).

Köszönetnyilvánítás

A kutatást az NKFI 128810 pályázat támogatta.

Irodalom

- Bándli J. 2011. *Az egyet nem értés pragmatikája*. PhD-értekezés. Pécsi Tudományegyetem.
- Boersma, P. – Weenink, D. 2012. *Praat: doing phonetics by computer (Version 5.1)*. Elérhető: http://www.fon.hum.uva.nl/praat/download_win.html (letöltés ideje: 2019. 01. 30.)
- Brown, P. – Levinson, S. 1987. *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.
- Edstrom, A. 2004. Expressions of disagreement by Venezuelans in conversation: reconsidering the influence of culture. *Journal of Pragmatics* 36: 1499–1518.
- García, C. 1989. Disagreeing and requesting by Americans and Venezuelans. *Linguistics and Education* 1: 299–322.
- Georgakopoulou, A. 2001. Arguing about the future: On indirect disagreements in conversations. *Journal of Pragmatics* 32: 1881–1900.
- Goffman, E. 1967. On face-work. In: *Interaction ritual: essays on face to face behavior*. New York: Doubleday Anchor. 5–46.
- Kakava, C. 2002. Opposition in modern Greek discourse: Cultural and contextual constraints. *Journal of Pragmatics* 34: 1537–1568.
- Koczogh, H. V. 2012. *The Effects of Gender and Social Distance on the Expression of Verbal Disagreement Employed by Hungarian Undergraduate Students*. PhD-értekezés. Debreceni Egyetem.
- Kotthoff, H. 1993. Disagreement and concession in disputes: On the context sensitivity of preference structures. *Language in Society* 22: 193–216.
- Likert, R. 1932. A Technique for the Measurement of Attitudes. *Archives of Psychology* 140: 1–55.
- Mérei F. 1971/2006. *Közösségek rejtett hálózata*. Budapest: Osiris Kiadó.
- Pearson, E. 1986. Agreement/disagreement. An example of results of discourse analysis applied to the oral English classroom. *ITL Review of Applied Linguistics* 74: 47–61.
- Pomerantz, A. 1984. Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In: Atkinson, J. – Heritage, J. (szerk.) *Structures of social interaction: Studies in conversation analysis*. Cambridge: Cambridge University Press. 57–101.
- Rees-Miller, J. 2000. Power, severity, and context in disagreement. *Journal of Pragmatics* 32: 1087–1111.
- Sacks, H. 1973. On the preferences for agreement and contiguity in sequences in conversation. Public lecture at the Linguistic Institute, University of Michigan. In: Button, G. – Lee, J. R. E. (szerk.) *Talk and Social Organisation*. Clevedon, UK: Multilingual Matters. 54–69.
- Schegloff, E. – Jefferson, G. – Sacks, H. 1977. The preference for self-correction in the organization of repair in conversation. *Language* 53: 361–382.

Alknyelvdok14

Az MTA Alkalmazott Nyelvészeti Munkabizottsága és az MTA Nyelvtudományi Intézete szeretettel meghívja Önt a 2020. február 7-én megrendezésre kerülő XIV. Alkalmazott Nyelvészeti Doktoranduszkonferenciára.

Az Alkalmazott Nyelvészeti Doktoranduszkonferenciák megrendezésére évente kerül sor. A tizennegyedik konferencia helyszíne az MTA Nyelvtudományi Intézete (1068 Budapest, Benczúr u. 33.). Előadók és résztvevők jelentkezését egyaránt várjuk.

A konferenciára előadással jelentkezni 2019. december 8-ig lehet.

Az absztrakt követelményei a honlapon olvashatók.

Részvételi díj: 4000 Ft az előadók, 1000 Ft a hallgatóság számára. A részvételi díj tartalmazza:

- az üléseken való részvétel lehetőségét,
- a kávészünetekben a kávé, üdítőt, ásványvizet, aprósüteményt,
 - a konferenciacsomagot (programfüzet, mappa, toll),
- az előadók számára a lektorálás utáni publikálási lehetőséget a konferenciakötetben.

Angol és magyar nyelvű jelentkezési lapok később a konferencia honlapjáról lesznek elérhetőek. A jelentkezések regisztrálásáról két munkanapon belül küldünk értesítést a jelentkező e-mail-címére.

A beküldendő kivonatok formai és tartalmi követelményeire vonatkozó útmutató a <http://www.nytud.hu/alknyelvdok20/absinfok.html> webhelyen található.

A konferenciával kapcsolatos további részleteket szintén a konferencia webhelyén (<http://www.nytud.hu/alknyelvdok20>) olvashatók.

Üdvözlettel,

Várad Tamás
Az Alkalmazott Nyelvészeti
Doktoranduszkonferenciák
szervezőbizottságának elnöke

Ludányi Zsófia
Az Alkalmazott Nyelvészeti
Doktoranduszkonferenciák
szervezőbizottságának titkára

Markó Alexandra
Az MTA Alkalmazott Nyelvészeti
Munkabizottságának elnöke

Gráci Tekla Etelka
Az MTA Alkalmazott Nyelvészeti
Munkabizottságának titkára