

Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2018

XII. Alkalmazott Nyelvészeti Doktoranduszkonferencia

Budapest, 2018. 02. 02.

Sorozatszerkesztő: Váradi Tamás

Szerkesztő: Ludányi Zsófia, Krepsz Valéria, Grácz Tekla Etelka

Lektorálta:

| | | |
|-----------------|--------------------|--------------|
| Balaskó Mária | Grácz Tekla Etelka | Robin Edina |
| Deme Andrea | Hunyadi László | Sass Bálint |
| Domonkosi Ágnes | Károly Krisztina | Simon Eszter |
| Farkas Judit | Ludányi Zsófia | Sólyom Réka |
| Fenyvesi Anna | Markó Alexandra | Schirm Anita |
| Fóris Ágota | Navracsics Judit | |
| Gósy Mária | Prószéky Gábor | |

Technikai szerk.: Ludányi Zsófia, Krepsz Valéria, Grácz Tekla Etelka

ISBN 978-963-9074-78-1

Kiadja: MTA Nyelvtudományi Intézet

Budapest, 2018.

Tartalom

| | |
|---|----|
| Dömötör Andrea: <i>Nem mind VP, ami állít. A névszói állítmány azonosítása számítógépes elemzőben</i> | 3 |
| Hantó Réka: <i>A spontán beszéd grammatikai komplexitása gyermekeknél. Jelzős és mellérendelő szintagmák, valamint összetett szavak előfordulása kisiskolás és kamaszkorban</i> | 11 |
| Huszár Anna: <i>A szóidőtartam és a beszédszakasz hosszának összefüggései spontán beszédben</i> | 23 |
| Kondacs Flóra: <i>A hát diskurzusjelölő prozódiai jellemzői a gyermeknyelvben</i> | 33 |
| Kresztyankó Annamária: <i>Magyarországi cigány közösségek interakciós stratégiái: az iskolai nyelvi gyakorlatok és az identifikáció</i> | 45 |
| Pomázi Bence: <i>Az N-nAk köszönhetően mint névutói szerkezet</i> | 58 |
| Sitkei Dóra: <i>Köztesnyelvi udvariassági választások a kérés beszédaktusában</i> | 70 |
| Rodrigo Tovar-Viera: <i>Linguistic realizations of research article abstracts written by Ecuadorian and North American academic authors</i> | 82 |

Nem mind VP, ami állít

A névszói állítmány azonosítása számítógépes elemzőben

Dömötör Andrea

PPKE BTK Nyelvtudományi Doktori Iskola
domotor.andrea@itk.ppke.hu

Kivonat: A kutatás annak lehetőségeit vizsgálja, hogyan azonosítható az esetrag nélküli névszók állítmányi szerepe egy pszicholingvisztikai indíttatású számítógépes mondatelemző számára. Ehhez egy konkrét szabályalapú modellt mutat be, amely 8 egyszerű, morfoszintaktikai címkékre alapozott szabály alapján, csak a kérdéses névszót megelőző mondatrészt figyelembe véve állapítja meg, hogy a névszó predikatív funkciót tölt-e be. A teszteredmények szerint amennyiben a kisebb, lokális elemzési feladatok (NP-sítés, a ragtalan névszók további eseteinek egyértelműsítése) jól vannak elvégezve, akkor a nagyobb egységekre alapozó második fázisú elemzésben a predikatív névszók azonosítása kevés erőforrással is nagy hatékonysággal megoldható.

1 Bevezetés

A tanulmány a nominális mondatok számítógépes elemzési lehetőségeit vizsgálja egy pszicholingvisztikai indíttatású elemzőrendszer keretében. Ezek a mondatok azért jelentenek a nyelvtechnológia számára különös kihívást, mert az elemzés középpontjában általában az ige és annak vonzatkerete áll, névszói állítmány esetén viszont legfeljebb a kopulával számolhatunk. A kopula viszont, mint ismeretes, nem mindig jelenik meg a mondatban hangzó formában, és a *van* szótó valamely alakjának jelenléte sem utal feltétlenül arra, hogy nominális mondatról van szó, hiszen a lokatív, egzisztenciális és birtokos mondatoknak is ez az igéje.

Éppen ezért a kutatás kiinduló elképzelése szerint a nominális mondatok elemzésének kulcsa nem a (zéró) kopula megtalálásában, hanem a predikatív névszók állítmányi szerepének felismerésében rejlik. Ezt erősítik meg a Universal Dependencies névszói predikátumokra vonatkozó irányelvei is. Eszerint, ha a kopula bármely számban vagy személyben elhagyható az adott nyelvben, akkor a függőségi elemzésben a predikatív névszó tekintendő fejnek, és ehhez kapcsolódik a kopula (ha van) egy 'cop' éllel (Nivre 2014). Azaz az elemzés során nem a kopula keresi a predikatív névszót, hanem fordítva.

Az elemzés szempontjából tehát kulcsfontosságú annak felismerése, hogy egy névszó predikatív szerepet tölt be. A magyarban azonban a morfológia alapján nem ismerhető fel az állítmányi szerep, hiszen a predikatív névszók a nominatívusziakhoz hasonlóan nem viselnek testes esetragot. A szintaktikai szerepre tehát csak a mondatbeli pozíció, azaz a kontextus mintázatai alapján lehet következtetni. Erre tesz kísérletet jelen kutatás.

Ludányi Zsófia, Krepsz Valéria, Gráczki Tekla Etelka (szerk.): *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2018*. Budapest: MTA Nyelvtudományi Intézet. 2018.

1.1 Az AnaGramma elemző

A célul tűzött elemzés az AnaGramma nyelvi elemzőrendszer (Prószéky–Indig 2015; Prószéky–Indig–Vadász 2016) működési elveit követi. Ennek fő tulajdonsága, hogy pszicholingvisztikai indíttatású, azaz az emberi szövegfeldolgozást próbálja modellálni. Ennek megfelelően az elemzés balról jobbra, szavanként halad. Az aktuális szó elemzése során így elsősorban az öt megelőző szavakra lehet támaszkodni, de rendelkezésre áll egy +2 token méretű ablak is, amely segítségével az éppen elemzett szót közvetlenül követő két szóra előretekintve valósul meg a kétfázisú elemzés első szakasza (Frazier–Fodor 1978). Az elemzés úgynevezett kereslet-kínálat elven működik, ami azt jelenti, hogy az éppen elemzett szó elindíthat keresleti szálakat (például: az ige keresi a vonzatait), vagy kielégítheti valamelyik aktív keresleti szál igényét (például: az alany kapcsolódik az igehez). Az elemző kimenete egy függőségi nyelvtanokra emlékeztető gráfrepresentáció.

Ezek alapján a névszók állítmányi szerepének e tanulmányban ismertetett vizsgálata sem a teljes mondat ismeretében, hanem a bal kontextus, illetve esetleg az ablak alapján történik. Ha a névszó felismerhetően predikatív, akkor a továbbiakban alanyt és esetlegesen kopolát fog maga mellé keresni.

1.2 Előzmény: nom-or-what

A testes esetrag nélküli névszók kezelésére az AnaGramma kereteiben már vannak jelentős eredmények. A nom-or-what elnevezésű algoritmus (Ligeti-Nagy et al. 2018) ezen névszók négy különböző funkcióját (esetét) tudja megkülönböztetni:

- (i) Nominativus (alany)
- (ii) Genitivus (jelöletlen birtokos)
- (iii) NP-t módosító elem (jelző)
- (iv) Névtós szerkezet része

Az algoritmus minden esetben csak az ablakot, tehát az éppen elemzett szót követő két tokent használja fel a döntéshez. Az elemzés ezen része tehát lokálisan, az első fázisban történik, az állítmányi szerep azonosítása azonban nem oldható meg így. Az ablak által nyújtott információ alapján csak nagyon ritkán lehet megállapítani, hogy az adott névszó predikatív-e, ehhez nagyobb egységekre (az NP-kre) és – a magyarra jellemző alany-állítmány szórend miatt – általában inkább a kérdéses szót megelőző mondatrészre van szükség.

A névszók predikatív tulajdonságának vizsgálata tehát a második fázisban, a bal kontextus, azaz az AnaGrammában tározónak nevezett adathalmaz alapján történik. A tározóban a megelőző kontextus elemei találhatóak, elemzett vagy részlegesen elemzett formában. Ebből tehát minden információ rendelkezésre áll, ami az addig elhangzott szavakból kiderülhetett; tudjuk, mely mondatrészek vannak már meg, és mi az, amit még keresünk. Ha a mondat több tagmondatból áll, ezek elhatárolása is menet közben történik, tehát a kérdéses névszóhoz érkeve már azt várjuk, hogy ismert a névszó tagmondatának eleje. Fontos megjegyzés még, hogy ekkor már csak azokról a névszokról kell dönteni, amelyek státuszát a nom-or-what algoritmus függőben hagyta.

Összefoglalva tehát a cél egy olyan szabályalapú algoritmus létrehozása, amely képes egy névszóról a bal kontextusa (és esetleg az öt közvetlenül követő szavak) alapján megállapítani, hogy állítmányi szerepben van-e.

2 Elméleti háttér

Az algoritmus részletes ismertetése előtt két elméleti kérdésre szeretnék kitérni. Egyrészt tisztázni kell, pontosan melyek azok a névszók, amelyeknek az esetleges predikatív „esetét” egyértelműsíteni akarjuk, másrészt az is némi magyarázatra szorul, hogy mit kell érteni az „eset-egyértelműsítés” alatt.

2.1 Predikatív névszók

A kutatás fókuszában a nominális mondatok állnak, azaz azok, amelyekben a predikátum szerepét nem egy (lexikális) ige, hanem egy névszó tölti be. A predikatív névszók lehetnek főnevek, melléknevek, ritkábban számnevek (l. [1]) és az ezeket helyettesítő névmások.

- [1] a. *Ez nekem sok.*
b. *A tegnapi eredmény 2:0 volt.*

Fontos még tisztázni, hogy a kopulás mondatok közül is csak a predikatívokkal (a tipológiát l. Higgins 1973) foglalkozunk, az azonosító mondatokkal (pl. *Én vagyok az apád.*) egyelőre nem, ezekben ugyanis nem mindig egyértelmű, hogy melyik névszót vagy névmást kell predikatívnek tekinteni, és melyiket alanynak, vagy hogy egyáltalán feltételezzünk-e alany-állítmányi viszonyt a két DP között. Ennek a mondattípusnak a számítógépes kezelése további kutatásokat igényel.

Jelen tanulmány nem foglalkozik a másodlagos predikátumokkal (pl. *Tibi okosnak látszik.*) sem, mivel ezek datívuszban állnak, a kutatás pedig csak az esetrag nélküli névszókat célozza.

Az algoritmusnak tehát akkor kell egy névszót predikatívnek ítélni, ha az az alany valamilyen attribútumát jelöli. Ha a névszói állítmány frázis, akkor a program a fejnek ítéli meg a predikatív jegyet, amit [2]-ben a vastag betű jelöl. Ha a további elemzés során az NP-összecsomagolás megtörténik, akkor a fej predikatív jegye természetesen az egész frázisé lesz, hasonlóan az esetragokhoz.

- [2] a. *Judit **ügyvéd.***
b. *Judit megfontolt, kezdő, ugyanakkor nagyon dörzsölt **ügyvéd.***

2.2 Milyen eset a névszói állítmány?

Az elérhető korpuszok általában minden esetrag nélküli névszót nominatívusznak jelölnek, a számítógépes elemzést viszont az segítené, ha csak az alany lenne NOM-ként címkézve, a többi funkció más jelölést kapna. A nom-or-what fentebb ismertetett kategóriái közül a (ii)–(iv) valóban kevésbé tekinthető nominatívusznak. A predikatív névszók esetében azonban nem ilyen egyértelmű a helyzet.

Szécsényi (2000) amellet érvel, hogy az állítmányi szerepű névszó az alannal való egyeztetés útján maga is nominatívuszi esetet kap. Ezzel szemben Kádár (2011) szerint egyáltalán nincs is értelme a névszói predikátumok esetéről beszélni, hiszen ezek nem argumentumok, így nincs szükségük absztrakt esetre. Az, hogy alakilag a nominatívuszi névszókkal egyeznek meg, azért van, mert a magyarban morfológiai szempontból ez az alapértelmezett (Schütze 2001).

Jelen tanulmány ebben a kérdésben nem kíván állást foglalni. A gépi elemzés szempontjából a predikatív tulajdonság mint egyfajta funkciójegy jelzi, hogy az adott elem egy alanyt és esetleg egy kopulát vár maga mellé. Az absztrakt eseteknek tehát itt valójában nincs jelentősége, a névszói állítmányt így egyszerűen esetrag nélküli, predikatív jeggyel bíró névszónak tekintjük.

3 A névszók eset egyértelműsítése

Bár a nom-or-what eset-egyértelműsítésről beszél (az általa egyébként részletesen nem tárgyalt predikatív névszók esetében is), valójában az „eset” megnevezés, amint láttuk, nem feltétlenül követi az elméleti nyelvészet esetfogalmait. Inkább a morfoszintaktikai címkében jelölendő tulajdonságokról van szó, amelyek szükségesek ahhoz, hogy az elemző egyrészt meg tudja határozni az NP-ket (nom-or-what), másrészt felismerje a nem igei állítmányt (predikatív névszók azonosítása).

A kutatás céljával tűzött predikatív névszókat azonosító algoritmus az „is-pred” nevet viseli, és a kétfázisú elemzési folyamat második fázisában valósul meg. Az elsőben, amint az 1.2 alfejezetben említettem, a nominatívusz, a genitívusz, az NP-módosító és a névutó előtti névszó azonosítása történik lokálisan, az ablak alapján. A második fázisban az is-prednek már csak azokról a névszókról kell döntenie, amelyeket a nom-or-what nem tudott egyértelműen azonosítani. Az első fázis jelenléte rendkívül fontos, enélkül kevés esély lenne arra, hogy a predikatív tulajdonságot a bal kontextus alapján fel tudjuk ismerni, amint azt [3] példái mutatják.

- [3] a. *Zsófi orvos.* – predikatív
 b. *Zsófi orvos barátja...* – NP-módosító
 c. *Zsófi orvos nélkül is kigyógyult az influenzából.* – névutós szerkezet része

Ez lényegében azt jelenti, hogy bár az is-pred algoritmus a pszicholingvisztikai indíttatás szerint alapvetően a már elhangzott elemeket veszi figyelembe, azért van valamennyi „rálátása” a kérdéses névszót követő szavakra is.

Az algoritmus bemenete tehát az eldöntendő szó, az azt megelőző mondatrész (tározó) és az azt közvetlenül követő két token (ablak). A kimenet pedig nagyjából a háromértékű logika értékeinek felel meg: *Pred*, ha a névszó egyértelműen predikatív, *Nonpred*, ha egyértelműen nem az, és *Undefined*, ha a meglévő információkból nem lehet biztos döntést hozni. Az algoritmus magas pontosságra törekszik, és kevesebb hangsúlyt fektet a fedésre: amennyiben a teljes mondat ismerete nélkül nem egyértelmű a predikativitás kérdése, inkább *Undefined* választ ad, minthogy hibázzon, hiszen később, a további szavak beolvasásával a bizonytalan ítélet pontosítható lesz.

A döntést segítő szabályok elsősorban a szófajcímkékre alapoznak, a lexikális információkat kevésbé veszik figyelembe. Az algoritmus tehát alapvetően a mondat szerkezet alapján azonosítja a predikatív névszókat. A szabályok kialakítása nyelvészeti megfontolások és intuíciók alapján történt, melyeket minden esetben korpuszmérésekkel ellenőriztem.

A felhasznált szabályok sorrendben a következők:

1. Ha van a tározóban vagy az ablakban ragozott ige, amely nem (potenciális) kopula → *Nonpred*
2. Ha a szó címkéje *ROMAN* (római szám) → *Nonpred*

3. Ha a következő szó címkéje EKSZ (-e kérdőszó) → Pred
4. Ha az ablakban van 1–2. számú létige → Pred
5. Ha a szó DP (része) → Nonpred
6. Ha a tározóban van egyértelműsített nominatívusz, és a szó magában áll vagy frázis vége → Pred
7. Ha a szó melléknév vagy számnév, és magában áll vagy frázis vége → Pred
8. Ha a szó többes számú melléknév → Pred
9. Egyébként → Undefined

A szabályok sorrendje fontos, a 7. és 8. például csak akkor működik, ha az 5. már kizárta azt a lehetőséget, hogy az adott melléknév alany legyen. Az 5. szabály egyébként DP-nek tekinti a határozott névelős főnévi csoportokat, a tulajdonneveket, a személyes és mutató főnévi névmásokat és bizonyos birtokos szerkezeteket. Ez utóbbiak közül az számít DP-nek, amelyek alany vagy azonosító, és nem tulajdonságot jelöl (l. [4]). A DP-k tehát (a birtokosokon kívül) a tározó információi alapján elméletileg egyértelműen azonosíthatók. Mivel a jelen algoritmusnak nem célja az NP-sítés, a tesztkorpuszban ezek kézzel vannak jelölve.

- [4] a. *Ádám tagja a szövetségnek.* – predikatív
 b. *Ádám a szövetség legifjabb tagja.* – azonosító

Azt, hogy a szó magában áll vagy frázis vége-e, az ablak alapján próbálja eldönteni az algoritmus. Ehhez egy reguláris kifejezéseket illesztő függvényt használ, amely olyan POS-tag-mintákat keres, amelyek egyértelműen jelzik a frázis végét: például hogyha a névszó után névelő vagy ragozott ige következik. Ennek kidolgozásában megint csak a nagy pontosság volt a cél: a függvény csak akkor jelez frázisvéget, ha teljes biztonsággal megállapítható, hogy a szót tartalmazó frázis már nem fog folytatódni.

[5]-ben a 6. és 7. szabályokra adok egy-egy példát. (A többi szabály egyértelmű.) Vastag betű jelöli a kérdéses névszót, az utána következő mondatrész ebben az elemzési fázisban nem látható.

- [5] a. *ez/DET_NM.NOM a/DET szombati/MN.0 buszjárat/FN.NOM **vice**/FN.pred* → predikatív a 6. szabály alapján
 b. *Eredeti/MN.0 formájában/FN.PSe3.INE nem/HA **azonos**/MN.pred a/a/DET kereszténység/FN.GEN istenfogalmával/FN.PSe3.INS ./SPUNCT* → predikatív a 7. szabály alapján

Az algoritmust, egyelőre teszt céllal, Python3-ban implementáltam.

4 Az algoritmus tesztelése

4.1 Tesztadat, módszer

Az is-pred algoritmus teszteléséhez az MNSz.²-ből (Oravecz–Váradi–Sass 2014) készítettem tesztkorpuszt. Ez 300 tagmondatot tartalmaz, melyek közül 200-ban van predikatív névszó. A tesztmondatok között nem szerepel azonosító mondat, hiszen mint 2.1-ben említettem, ezek elemzése nem feltétlenül követi a predikatív mondatot.

két. Egyéb tekintetben a korpusz random összetételű. Az MNSz.² morfológiai annotációjához képest annyi módosítást tettem, hogy a tulajdonnevek saját címkét kaptak. Ezen kívül speciális jelzéseket adtam a DP-knek (hogy az esetlegesen hibás NP-chunkolás ne zavarja az algoritmus értékelését), illetve szükség esetén kézzel javítottam a POS-tagek hibáit.

A tesztkorpuszon először lefuttattam a nom-or-what implementációját, majd az eredményt ismét kézzel javítottam, ahol szükséges volt. Ez után 896 egyértelműsítendő névszó maradt, amelyből 208 volt predikatív. Ezeket kézzel annotáltam a kérdéses névszók valódi funkciója szerint, a teljes mondatot figyelembe véve (Pred vagy Nonpred), így állt elő a kiértékeléshez szükséges gold standard. A kiértékelési módszer sajátossága, hogy az Undefined választ semmilyen esetben nem tekint jónak, akkor sem, ha egyébként az adott elemzési ponton valóban nem lehet még dönteni a névszó predikativitásáról, azaz az algoritmus tulajdonképpen jól működött, de az eredeti kérdéseinkre nem adott választ. A teszt tehát nem az *elvárt*, hanem a *hasznos* működést értékeli.

4.2 Eredmények és értékelés

Az is-pred script gold standardhoz viszonyított eredményei az 1. táblázatban láthatók. A sorok jelzik az algoritmus válaszait, az oszlopok a várt (kézi annotáció szerinti) értékeket. A tévesztési mátrix mellett a pontosság és fedés értékek is szerepelnek.

| | Pred | Nonpred | Pontosság (%) |
|-----------------------|-------------|----------------|----------------------|
| Pred (db) | 124 | 2 | 98,4 |
| Nonpred (db) | 0 | 546 | 100 |
| Undefined (db) | 84 | 139 | - |
| Fedés (%) | 59,6 | 79,5 | |

1. táblázat. Az is-pred algoritmus eredményei

Mint látható, az algoritmus magas pontossággal dolgozik, az Undefined válaszokat nem számítva csak kétszer tévedett. Ezek a hibák [6]-ban láthatók.

- [6] a. *Ebből egy ötödik, a férfi pólósoké rendkívüli családás.*
 b. *A három fiú közül az egyik nagydarab volt, mint egy súlyemelő.*

A [6a]-ban az algoritmus az *ötödik* számnevet predikatívnak ítélte a 7. szabály alapján, az ablakban található determináns ugyanis a frázis végét jelezte a program számára. Vagyis ezt a hibát nem az algoritmus szabályainak alkalmatlansága, hanem a frázis végét jelző függvény pontatlansága okozta, ez ugyanis nem számolt az ehhez hasonló magyarázó szerkezetekkel. A [6b]-ben pedig a 6. szabály alapján ítélte a program a *súlyemelő* szót predikatívnak, az ilyen jellegű összehasonlításokra ugyanis nem volt felkészítve.

Az 1. táblázatból az is látszik, hogy a predikatív névszókat gyakrabban specifikálja alul a program, mint a nem predikatívokat, azaz előbbieket esetében alacsonyabb a fedés. Úgy tűnik tehát, hogy az algoritmus a „nem” döntést könnyebben meg tudja hozni, mint az „igen”-t. Azt, hogy a névszó biztosan nem predikatív, közel 80% biztonsággal meg tudta mondani a program, míg az egyértelműen predikatív válasz esetén

ez a szám nem érte el a 60%-ot. Ennek az lehet oka, hogy a Nonpred választ adó szabályok feltételeinek igazsága könnyebben megállapítható a rendelkezésre álló információkból. Mindemellett a viszonylag alacsony fedés a teljes elemzés szempontjából nem olyan nagy probléma, a későbbiekben, a további szavak beolvasásával könnyen korrigálható.

A fedésértékek egyébként pszicholingvisztikai szempontból is érdekeseek, hiszen azt mutatják, hogy (legalább) ekkora arányban „valós időben”, azaz csak a szó elhangzásáig rendelkezésre álló információból eldönthető, hogy a névszó predikatív-e. De emberi értékeléssel ennél valószínűleg még jóval nagyobb számokat is kaphatnánk, hiszen az ember lexikális információkra is támaszkodik, amire a program egyáltalán nem.

A 2. táblázat azt mutatja, hogy melyik szabályt hányszor alkalmazta a program.

| | |
|---|-----|
| 1. ha van ige → Nonpred | 97 |
| 2. ha ROMAN → Nonpred | 3 |
| 3. ha EKSZ következik → Pred | 1 |
| 4. ha 1–2. személyű létige van az ablakban → Pred | 2 |
| 5. ha DP (része) → Nonpred | 446 |
| 6. ha a tározóban van nominatívusz → Pred | 65 |
| 7. ha melléknév vagy számnév → Pred | 44 |
| 8. ha többes számú melléknév → Pred | 14 |
| 9. egyébként → Undefined | 224 |

2. táblázat. A szabályok alkalmazásának gyakorisága

Úgy tűnik, a leggyakoribb nom-or-what által azonosítatlan névszó tulajdonnév vagy DP része (5. szabály). Ezek az elemzés második fázisában, balról jobbra elemezve a determinánsok segítségével már könnyebben egyértelműsíthetők. A predikativitás felismerésében legtöbbször a nominatívusz jelenléte segített (6. szabály), de a melléknévek és számnevek esetében sokszor enélkül is megállapítható az állítmányi szerep (7. és 8. szabály). Amint az előzőekből is kiderült, az algoritmus viszonylag nagy arányban, az esetek negyedében nem tud egyértelmű döntést hozni.

5 Konklúzió

Az is-pred algoritmus elérte a kitűzött célját: nagy pontossággal képes a pszicholingvisztikai indíttatású elemző elveit követve megállapítani, hogy egy névszó predikatív szerepet tölt-e be. A fedés viszonylag alacsonyabb mértéke részben a korlátozott információ (nem áll rendelkezésre a teljes mondat), részben az algoritmus „óvatossága” miatt van, úgy lett ugyanis felépítve, hogy csak biztos választ adjon, a bizonytalan eseteket hagyja függőben.

Az eredmények alapján elmondható, hogy a nagy pontosságra törekvés mellett is elég nagy számban lehet a névszók állítmányi szerepéről dönteni egy egyszerű, kevés szabályt tartalmazó algoritmussal. Ehhez viszont szükség van arra, hogy a bal kontextus ragtalan névszónak többi funkciója (különösen a nominatívusz) jól legyen azonosítva, illetve jól működjön az elemzőben az NP-sítés. Vagyis a predikatív névszók azonosítása akkor lehet sikeres, ha a kisebb, lokális elemzési feladatok jól vannak megoldva. Ez a már említett kétfázisú elemzési modell logikus következménye.

További kutatási feladat a pontosság-fedés arány optimalizálása esetleges további szabályokkal vagy a meglévők finomításával. Ezenkívül komoly elméleti megfontolások is igénylő kérdés az itt nem tárgyalt azonosító mondatok kezelése.

Köszönetnyilvánítás

Jelen kutatás az FK~125217 és a PD~125216 számú projekt keretében az FK17 és a PD17 pályázati program finanszírozásában a Nemzeti Kutatási Fejlesztési és Innovációs Alap által biztosított támogatással valósult meg.

Irodalom

- Frazier, L., Fodor, J. D. 1978. The sausage machine: A new two-stage parsing model. *Cognition* 6/4. 291 – 325.
- Higgins, R. 1973. *The Pseudo-Cleft Construction in English*. New York: Garland Press.
- Kádár, E. 2011. *A kopula és a nominális mondatok a magyarban*. Budapest: Akadémiai Kiadó.
- Ligeti-Nagy, N., Vadász, N., Dömötör, A., Indig, B. 2018. Nulla vagy semmi? Esetgyértelműsítés az ablakban. In: Vincze, V. (szerk.): *XIV. Magyar Számítógépes Nyelvészeti Konferencia* (Szeged, 2018. január 18–19.) Szeged: Szegedi Tudományegyetem. 25–37.
- Nivre, J. 2014. *Nonverbal Predication and Copulas in UD v2*. Elérhető: <http://universaldependencies.org/v2/copula.html> Letöltés dátuma: 2018. febr. 8.
- Oravecz, Cs., Váradi, T., Sass, B. 2014. The Hungarian Gigaword Corpus. In: Calzolari, N. et al. (szerk.): *Proceedings of the 9th International Conference on Language Resources and Evaluation* (Reykjavik, Izland, 2014. május 26–31.) European Language Resources Association. Reykjavik, Izland. 1719-1723.
- Prószéky, G., Indig, B. 2015. Magyar szövegek pszicholingvisztikai indíttatású elemzése számítógéppel. *Alkalmazott Nyelvtudomány*, 15(1-2): 29–44.
- Prószéky, G., Indig, B., Vadász, N. 2016. Performanciaalapú elemző magyar szövegek számítógépes megértéséhez. In: Kas, B. (szerk.): „Szavad ne feledd!": *Tanulmányok Bánréti Zoltán tiszteletére*. Budapest: MTA-NYTI. 223–232.
- Schütze, C. 2001. On the Nature of Default Syntax. *Syntax*, 4: 205–238.
- Szécsényi T. 2000. Esetgyeztetés a predikatív főnévi csoportban. In: Büky, M., Maleczki, M. (szerk.): *A mai magyar nyelv leírásának újabb módszerei IV*. Szeged: SZTE. 189–202.

A spontán beszéd grammatikai komplexitása gyermekeknél

Jelzős és mellérendelő szintagmák, valamint összetett szavak előfordulása kisiskolás és kamaszkorban

Hantó Réka

ELTE BTK Nyelvtudományi Doktori Iskola,
MTA Nyelvtudományi Intézet
hanto.reka@nytud.mta.hu

Kivonat: A jelen tanulmány hatéves kisiskolások és tizenhárom éves kamaszok spontán beszédét veti össze a grammatikai szerkesztettség tekintetében. Az elemzés az összetett szavak gyakoriságának meghatározására, valamint az egyes jelzőtípusok és a mellérendelő szintagmák arányának adatolására terjedt ki. Eredményeink szerint minden mutatóban a tizenhárom éves korosztály spontán beszédében volt adatolható több előfordulás a vizsgált nyelvi kategóriákra azonos beszédmennyiségben (200 szó). A jelzős szintagmák vonatkozásában az eltérés szignifikáns volt a két csoport között, melyet főként a minőség- és mennyiségjelzős szintagmák számának növekedése idézett elő. A mellérendelő szintagmáknál mindkét korcsoportban a kapcsolatos típusú szerkesztésmód fordult elő legtöbbször, ezt követte a választó, majd az ellentétes típusú szintagma. Statisztikai értelemben csak a jelzős szintagmák számadataiban találtunk eltérést a korcsoportok között, vagyis megállapítható, hogy a két életkori csoport nem áll egymástól távol az itt elemzett nyelvi kategóriák vonatkozásában.

1 Bevezetés

Az írás és a beszéd a nyelv két különböző, de egyenrangú megjelenési formája, melyek eltérő normarendszerrel rendelkeznek, amelyben a mondattani eltérések a legjelentősebb különbözőségek közé tartoznak (Lanstyák 2009), emellett az írott és a beszélt nyelv eltér egymástól a lexikai tömörség tekintetében is (Halliday 1979). Az írott nyelv tudatosabban szerkesztett, hiszen több idő van az átgondolásra, mint élőszóban, valamint kevesebb benne a tévesztés, mert lehetőség van a korrigálásra (Érsok 2006).

Az élőszóra jellemző a mellékmondatos szerkesztésmód, a lineáris jellegű, láncszerűen egymást követő tagmondatok használata, melyek az emlékezeti terhelés csökkentését szolgálják (Wacha 1988). A beszélt nyelv további sajátossága, hogy gyakoribbak benne az elliptikus, közbevetett mondatok, a mondattöredékek vagy az agrammatikus mondatok, valamint a mondatátszövődések (Keszler 1983), emellett jellemző a redundancia, a túlbiztosítás is (Gósy 2004).

Ludányi Zsófia, Krepsz Valéria, Grácsi Tekla Etelka (szerk.): *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2018*. Budapest: MTA Nyelvtudományi Intézet. 2018.

A valódi spontán beszéd jellegzetessége, hogy a gondolatok meghangosítását nem előzi meg előzetes felkészülés (Gósy 2005), szemben más beszéd típusokkal, mint például a felolvasás vagy a félspontán/félinterpretatív beszéd. Utóbbi során a beszélő szintén az adott pillanatban mondja el közlendőjét, azonban előzetesen megformálja azt. A beszédprodukciónak működésére több modellt is kidolgoztak, ezek közül a legismertebb Willem Levelt modellje, amely szerint makro- (pragmatikai, szemantikai, szintaktikai tervezés) és mikroszinten (nyelvi szerkezetek kiválasztása) tervezzük meg a kiejtendő szándékozott közlést, de a tervezés és a kivitelezés egyszerre zajlik le (Levelt 1989).

A spontán beszéd tanulmányozása gyermekek körében viszonylag új kutatási terület, ez idáig a legtöbb kísérlet a felnőttek spontán megnyilatkozásainak leírására irányult. A gyermeknyelvi kutatások zöme pedig az anyanyelv-elsajátítás kezdeti szakaszaira koncentrál. Ennek magyarázata, hogy a beszéd fejlődésében meghatározó az első hat év, ez alatt az időszak alatt zajlanak le az anyanyelv-elsajátítás legfontosabb nyelvi folyamatai, ezt követően látványosabb nyelvi változások már nem következnek be (Horváth 2013). Ugyanakkor fontos az első biológiai sorompó (6–7 éves kor) utáni időszak vizsgálata is, hiszen az anyanyelv elsajátítása ezekben az években is tovább folytatódik, tulajdonképpen életünk végéig (Horváth 2016).

Óvodások és kisiskolások (4–8 évesek) körében vizsgálták a KFM-módszer (közlélegységek fejlődési mutatója, Gerebenné et al. 1992) alapján irányított spontán beszédben a megnyilatkozások grammatikai komplexitását (Laczkó 2011). Az adatok szerint az életkor előrehaladtával fokozatosan növekednek a KFM-értékek, valamint ezzel párhuzamosan az MLU-értékek (az átlagos mondatösszehasonlító mérésszáma, Mabel 2010) is, vagyis egyre összetettebbé válik a gyermekek beszéde. A négyéves óvodások (kiscsoportosok) alanyi-állítmányi szerkesztésű mondataikat rendszerint tárgyi és minősítő jelzős bővítményekkel egészítik ki, összetett mondataikat a kapcsolatos mellérendelések uralják. A középső csoportosok (5 évesek) beszédében már nagyobb arányban jelennek meg a határozók, míg a nagycsoportosoknál (6 évesek) egyre többször és változatosabb formában fordulnak elő az alá- és mellérendelő szerkezetek. Mondataik száma kevesebb, mint a középső csoportosokéi, de összetettebbek. A hétévesek (első osztályosok) megnyilatkozásaiban új szófajok (pl. határozatlan névmások, névmási határozószók) és a nehezebb morfológiai szerkezetek (ható igék és főnévi igenév ragozása) is egyre gyakoribbá válnak. Második osztályban (8 évesek) változatos bővítmények, új szófajok (gyakorító igék, melléknévi igenevek, határozói igenevek) használata jellemző.

Horváth (2017) szintén a KFM-módszer alapján vizsgálta 6–9 éves gyermekek spontán beszédének a grammatikai összetettségét. Az eredmények szerint a 6, 7 és 8 évesek KFM-értékei között nem volt jelentős különbség, a 9 éveseké azonban már jóval magasabb volt. Az itt kapott eredmények szerint az idő előrehaladtával a grammatikai szerkesztettség nem lineárisan fejlődik ezekben az életkori sávokban. A lányok minden életkori csoportban magasabb értékeket értek el, de a két nem közötti eltérés nem volt szignifikáns. A gyermekek életkora azonban szignifikáns hatással volt a narratívában kiejtett szavak számára. 7–9 éveseknél vizsgálták a beszédszakaszok időtartamát is (Horváth 2016), a kísérleti adatok szerint ezekben az életkorokban a beszédszakaszok átlagos hossza jelentősen elmaradt a felnőtteknél tapasztalható értékektől, ugyanakkor az egyes életkori csoportok között nem volt kimutatható szignifikáns különbség. Vallent (2010) középiskolások körében vizsgálta a beszédidő és a szünettartási stratégiák realizációt. Kutatásában a lányok átlagosan több ideig beszéltek, mint a

fiúk, de a fiúk beszédében több néma szünet fordult elő, így a lányok narratívája folyamatosabbnak hatott (Vallent 2010).

Egy másik magyar vizsgálatban 6, 9 és 13 évesek spontán beszédében vizsgálták ugyancsak a KFM-eljárással a megnyilatkozások szerkesztettségét és folyamatosságát (Neuberger 2011). A 9 éveseknél gyakoribbak voltak az összetett mondatok, mint a fiatalabbaknál, és már nem csak kapcsolatos kötőszókat használtak a tagmondatok összekapcsolására. A vizsgált életkori csoportok közül a 13 évesek beszéde közelítette meg legjobban bizonyos nyelvi jellemzők tekintetében a felnőttek beszédét: egyre több volt az összetett mondat, és már többféle és mennyiségileg is több határozót és jelzőt alkalmaztak. Közepes korrelációt mutatott a gyermekek spontán beszédének szintaktikai komplexitása a beszédükben használt szavak számával. Megállapítható volt, hogy a bonyolultabb struktúrák létrehozásához nagyobb szókincs szükséges. Az egyéni eltérések jelentősek voltak mindegyik életkori csoportban.

Neuberger (2014) 6, 7, 9, 11 és 13 évesek spontán beszédében vizsgálta a tárgyias, határozós és jelzős szerkezetek előfordulását. Az adatok szerint a hatéves korosztályban a tárgyias szintagmák előfordulási aránya volt a legmagasabb (34,8%), a többi korcsoport esetében azonban a hely- és időhatározós szintagmák voltak többségben. A legritkábbak a kötött bővítményes és a birtokos szerkezetek voltak mindegyik korcsoportban. Az életkorral növekvő tendencia legmarkánsabban a hely- és időhatározós szintagmák számának megugrásával mutatkozott meg. A jelzős szerkezetek aránya is fokozatosan nőtt az életkor előrehaladtával, ezen belül a minősítő jelzős szintagmák aránya mutatkozott a legmagasabbnak.

A beszélő életkora nagymértékben befolyásolja a beszéd számos jellemzőjét, mint a grammatikai szerkesztettség, a temporális jellegzetességek, a megakadásjelenségek vagy például a szóhasználat (például Bóna 2012; Laczkó 2009, 2010; Markó–Dér 2011; Deme 2012; Neuberger 2012b). A jelen vizsgálatban a gyermekkor két meghatározó életkori szakaszát, a kisiskolás-, illetve a kamaszkort vettem össze. A kutatás célja az volt, hogy a grammatikai komplexitás vonatkozásában hasonlítsa össze a két korosztályt. Vizsgáltam az összetett szavak, a jelzős és a mellérendelő szintagmák gyakoriságát.

Kutatási kérdések voltak, hogy i) kamaszkorra milyen eltérések jelentkeznek a kisiskolás korúak spontán nyelvhasználatához képest, ii) az eltérés milyen mértékű (tendencia vagy szignifikáns eltérés), illetve iii) mely szerkesztési mód a leggyakoribb, illetve a legritkább a két korosztályban? Az előzetes hipotézisek a következők: 1. Az összetett szavak nagyobb arányban fordulnak elő a tizenhárom éveseknél. 2. A jelzős szintagmák nagyobb arányban jelennek meg az idősebbeknél. 3. A mellérendelő szintagmák ugyancsak gyakrabban fordulnak elő az idősebb korosztályban.

Ez idáig viszonylag kevés kísérlet irányult a kamaszkorúak nyelvhasználatának leírására magyar nyelven. Emellett a két korosztály, a kisiskolások és a kamaszok spontán beszédének jellemzőiről, lehetséges eltéréseiről is viszonylag keveset tudunk, így az itt kapott eredmények fontos adalékul szolgálhatnak ezen a területen.

2 Kísérleti személyek, anyag, módszer

A jelen kutatáshoz a GABI (Gyermeknyelvi Beszédatadtbázis és Információtár, Bóna et al. 2014) hangfelvételeit használtam fel. 14 felvételt választottam ki a hatéves kisiskolások, valamint ugyancsak 14 felvételt a tizenhárom éves kamaszok korosztályából. A két kísérleti csoportban a nemek fele-fele arányban szerepeltek. Az adatközlők

ép értelmű, ép hallású gyermekek, pszichés betegség nem állt fenn esetükben. A nemek közötti eltéréseket az adatközlők kis száma miatt fenntartásokkal lehetett elvégezni (az adatbázisban még nincs elegendő felvétel).

A felvételekből a spontánbeszéd-részeket vizsgáltam, melyek során az interjúkészítő arra kéri a gyermeket, hogy szabadon, lehetőleg folyamatosan beszéljen neki tetsző témákról. A spontán megnyilatkozások – az írással szemben – nem tartalmazzak központosítást, így a beszédfolyam szegmentálása, a mondatok elkülönítése problematikus (Gósy 2003; Neuberger 2012a). Emiatt a jelen vizsgálatban nem különítettem el mondatokat, hanem ehelyett ennél kisebb nyelvi egységeket, a szintagmatikus kapcsolókat vizsgáltam. Az elemzéshez a besoroló típusú grammatikát vettem alapul (Készler szerk. 2010). Ennek alapján mértem a jelzős, valamint a mellérendelő szintagmák egyes altípusainak gyakoriságát (5-5-féle altípus), valamint az összetett szavak előfordulási arányát is. A jelzős szintagmákon belül elkülönítettem minőség-, mennyiség-, birtokos, kijelölő és értelmező jelzős szintagmákat. A mellérendelő szintagmákon belül vizsgáltam a kapcsolatos, az ellentétes, a választó, a magyarázó és a következtető szerkezet gyakoriságát. Ezek adatolásakor a közbeékelte néma vagy kitöltött szüneteket nem vettem figyelembe, tehát ha egy szintagma szemantikailag összetartozó részeit vagy egy összetett szót szünet választott ketté, azokat az eseteket ugyanúgy besoroltam a megfelelő szintagmák, illetve az összetett szavak közé. (A grammatikai szerkezetek és a szünettartási stratégiák egymásra való hatása egy következő tanulmány anyagát képezheti a későbbiekben.)

A vizsgált nyelvi kategóriák gyakoriságát 200 szóra vetítve határoztam meg (db/200 szó). Az elemzés manuálisan történt. A 200 szó kb. 4–5 perces időtartamnak felel meg egy-egy felételből. A 200 szó összeszámlálásakor a nemzetközi gyakorlatnak megfelelően a megakadásjelenségeket nem vettem figyelembe. A beszédrészeket a felvételek azon szakaszából választottam ki, ahol a lehető legfolyamatosabb volt a beszéd. Ez különösen a hatéves korosztály esetében volt fontos, hiszen az ő közlésegyiségek átlagosan rövidebbek, mint a tizenhárom évesekéi. Így – ha a megadott időegység alatt adatolható előfordulások számát határoztuk volna meg – az eredmények torzítanának, hiszen a tizenhárom évesek hosszabb közlésegyiségekben értelem szerűen minden nyelvi kategóriára több előfordulást adatolhattunk volna. Az elemzéskor a töltelésként funkcionáló szavakat nem vettem figyelembe (például az *ilyen* diskurzusjelölő), mert ez szintén torzítaná az eredményeket.

A spontánbeszéd-szakaszokat manuálisan, beszédszakasz-szinten jegyeztem le a Praat programban (Boersma–Weenink 2018). Az eredményeken a statisztikai elemzést az SPSS szoftver felhasználásával végeztük el.

3 Eredmények

3.1 Az összetett szavak gyakorisága

Az összetett szavak vizsgálata során meghatároztam az adott összetétel előfordulását ismétlések nélkül, illetve ismétlésekkel együtt is (ugyanannak a lexémának a többszöri előfordulása különböző inflexiók, jeles és ragos alakjaiban).

A hatévesek csoportjában összesen 211 db előfordulást adatoltam ismétlés nélkül, és 253 db előfordulást ismétléssel. Az átlag csoportszinten 15,07 db előfordulás volt ismétlés nélkül és 18,07 db ismétléssel. A tizenhárom éveseknél csoportszinten

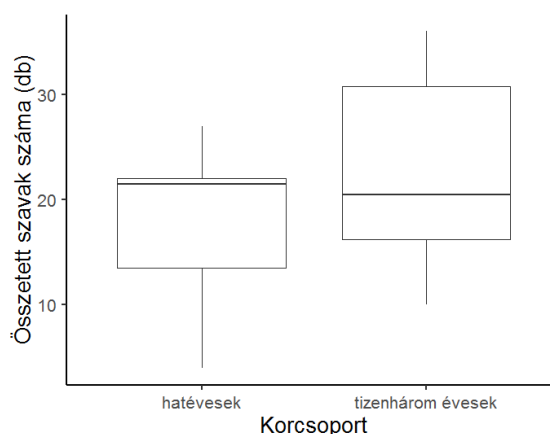
254 db, valamint 319 db előfordulás volt mérhető. Az átlag náluk 18,14 db ismétlés nélkül, valamint 22,78 db ismétléssel. A statisztikai számítás (UNIANOVA) szerint nincs különbség a két csoport között, a tendencia viszont kimutatható: a tizenhárom éveseknél fordul elő több összetett szó, vagyis ebben a tekintetben az ő spontán beszédük szerkezete kismértékben, de összetettebb.

A hatéves fiúknál 100 db előfordulás volt adatolható ismétlés nélkül, 124 db ismétléssel, a hatéves lányoknál összesen 111 db, valamint 129 db összetett szó jelent meg. A tizenhárom éves fiúknál az adatok 122 db, illetve 156 db előfordulás, a tizenhárom éves lányoknál 132, illetve 163 db előfordulás. A hatéves fiúknál átlagosan 14,28 db, illetve 17,71 db szóösszetétel volt adatolható 200 szóban, a hatéves lányoknál 15,85 db, illetve 18,42 db. A tizenhárom éves fiúknál átlagosan 17,42 db (ismétlés nélkül), illetve 22,28 db (ismétléssel) összetett szó fordult elő 200 szóban, a tizenhárom éves lányoknál ugyanezek az adatok 18,85 db, illetve 23,28 db. A nemi összevetésben sem volt mérhető statisztikai eltérés. Azonban mind a darabszámokból, mind az átlagokból látható, hogy a lányok spontán beszéde tartalmaz több szóösszetételt, vagyis – ugyan csak kismértékben – az ő beszédük szerkezetileg összetettebb.

Az adatok csoportokra vetített gyakoriságát az 1. ábra, a korpuszból vett nyelvi példákat az 1. táblázat szemlélteti.

| Kísérleti csoportok | Példák |
|------------------------|--|
| hatéves fiúk | <i>citromsárgák, rablótanyához, visszafordulunk, nyakkendőt</i> |
| hatéves lányok | <i>aranyszobát, királynő, gólyatoll, lovastáborba, delfinúszás, varázspálcát</i> |
| tizenhárom éves fiúk | <i>döntőképesség, kosárlabda-egyesületbe, gitártanárommal, mennybemenetelének,</i> |
| tizenhárom éves lányok | <i>történelemtanárral, iskolafelgyújtások, osztályközösségünk, diákszállószervezésbe</i> |

1. táblázat. Példák a korpuszból az összetett szavakra életkoronként és nemenként



1. ábra. Az összetett szavak gyakorisági eloszlása a két életkori csoportban ismétlésekkel

3.2 A jelzős szintagmák előfordulása

A jelzős szintagmák között adatoltam a besoroló típusú grammatika (Keszler szerk. 2010) alapján a minőség-, mennyiség-, birtokos, kijelölő és értelmező jelzős szintagmákat.

A hatévesek csoportjában összesen 67 db előfordulás volt adatolható mind az öt szintagmatípus figyelembe vételével, míg a tizenhárom éveseknél 152 db. A csoport-átlag a fiatalabbaknál 4,78 db / 200 szó, az idősebbeknél 9,71 db / 200 szó. Az UNIANOVA szerint szignifikáns különbség volt adatolható a két csoport között az életkor szerint: $F(1,27) = 15,424$; $p = 0,001$. A szélsőértékek figyelmen kívül hagyásával is szignifikáns a különbség a csoportok között.

A jelzőtípusok közül a minőségjelzős szintagma volt a leggyakoribb: a hatéveseknél összesen 35, az idősebbeknél 67 előfordulás volt adatolható csoportszinten. Gyakoriság szerint ezt követte sorrendben a mennyiségjelzős szintagma: a fiatalabbaknál 12 db, az idősebbeknél 39 db fordult elő. Birtokos jelzőből a hatévesek csoportjában összesen 7 db, a tizenhárom éveseknél 8 db, kijelölő jelzőből a fiatalabbaknál 9 db, az idősebbeknél 15 db, míg értelmező jelzőből az előbbi csoportban 4 db, az idősebbeknél 14 db fordult elő. Látható tehát, hogy elsősorban a minőség- és a mennyiségjelzők száma nő meg tizenhárom éves korra, a többi szintagmatípus aránya ezekhez képest elenyésző, valamint hasonló mennyiségű a két csoportban.

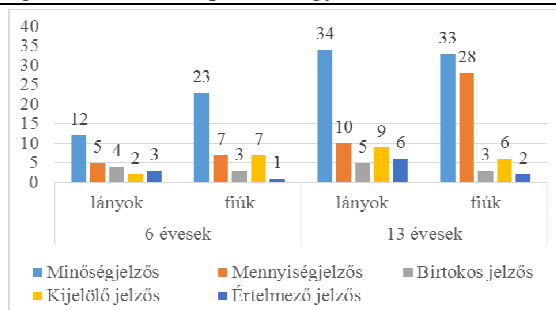
A hatéveseknél csoportszinten átlagosan 2,5 db minőségjelzős és 0,85 db mennyiségjelzős szintagma jelent meg 200 szóban. A tizenhárom éveseknél 4,78 db minőségjelzős, illetve 2,71 db mennyiségjelzős szerkezet fordult elő 200 szóra vetítve.

Ha 100%-nak tekintjük a jelzőtípusok összes előfordulását, akkor ezen belül a hatéveseknél a következőképpen oszlanak meg az egyes szintagmatípusok: 52,23% minőségjelzős, 17,91% mennyiségjelzős, 10,44% birtokos jelzős, 13,34% kijelölő jelzős, 5,97% értelmező jelzős szintagma. A tizenhárom éveseknél a minőségjelzős szintagmák aránya 44,07%, a mennyiségjelzős 25,65%, a birtokos jelzős 5,26%, a kijelölő jelzős 9,86%, az értelmező jelzős szintagmák aránya pedig 5,26%. Tehát az idősebbeknél is a minőségjelzős szintagmák aránya a legmagasabb, ugyanakkor náluk a mennyiségjelzős szintagmák aránya jobban megközelíti a minőségjelzős szintagmák arányát, mint a hatéveseknél.

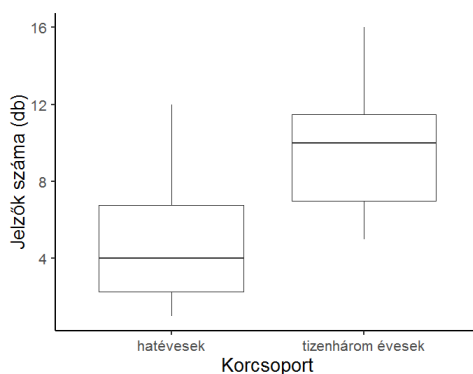
Az egyes szintagmatípusok eloszlása szerint nem tapasztalható lényegi eltérés a két életkori csoport között, vagyis az életkor nincs hatással arra, hogy a gyermekek spontán beszédükben mely jelzőtípusokat alkalmazzák gyakrabban vagy ritkábban: amely típus gyakori a fiatalabbaknál, az gyakori az idősebbeknél is, amely típusok viszont ritkább előfordulásúak a hatéveseknél, ritkán jelennek meg a tizenhárom éveseknél is.

Nemek szerinti eloszlásban a következőképpen alakultak az eredmények. A hatéves fiúknál a csoportban átlagosan 3,71 db, a tizenhárom éves fiúknál átlagosan 9,14 db, míg a hatéves lányoknál átlagosan 5,85 db, a tizenhárom éves lányoknál 10,28 db jelzős szintagma fordult elő. Bár nemenként kevés az adatközlők száma a két csoportban, de az eredményekből az látható, hogy mind a hatéveseknél, mind a tizenhárom éveseknél a lányok spontán beszédében jelent meg több jelzős szintagma.

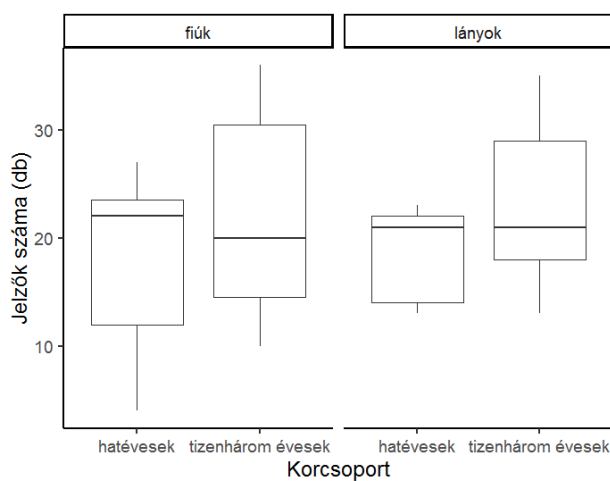
A kapott számadatok életkoronként és nemenként való eltéréseit a 2. ábra, az előfordulási gyakoriságot a 3–4. ábrák, a korpusz konkrét nyelvi példáit a 2. táblázat szemlélteti.



2. ábra. A jelzős szintagmák altípusainak gyakorisági eloszlása nemenként és életkoronként (db)



3. ábra. Az összes jelzőtípus gyakorisági eloszlása a két életkori csoportban



4. ábra. Az összes jelzőtípus gyakorisági eloszlása az életkor és nem szerint bontott csoportokban

| jelzőtípusok | hatévesek | | tizenhárom évesek | |
|--------------|--|--|--|--|
| | fiúk | lányok | fiúk | lányok |
| minőségj. | <i>meleg tej;</i> <i>porcukros nud-</i> <i>li</i> | <i>lila mókus;</i> <i>szép ruhába</i> | <i>tájfutó versenye-</i> <i>ket;</i> <i>könnyűségi skála</i> | <i>tengerparti család;</i> <i>hatalmas veszeke-</i> <i>dések</i> |
| menyiségj. | <i>sok játékunk;</i> <i>rengeteg</i> <i>dínókn</i> | <i>kettő copfja</i> | <i>sok unokatesó-</i> <i>immel;</i> <i>több helyre</i> | <i>két kishúgom;</i> <i>több edzés</i> |
| birtokos j. | <i>nagymamának</i> <i>a kávédaráló-</i> <i>ja;</i> <i>templom egere</i> | <i>a csoport</i> <i>nyuszija;</i> <i>Zoénak a ba-</i> <i>rátját</i> | <i>május elseje</i> | <i>iskola szelleme;</i> <i>a hárompontos vo-</i> <i>nala</i> |
| kijelölő j. | <i>ez a kettő;</i> <i>ebben az alak-</i> <i>ban</i> | <i>azon a va-</i> <i>sárnapon</i> | <i>ennek a dombnak</i> <i>a lábánál;</i> <i>itt, a lakótelepen</i> | <i>az egyik húgom;</i> <i>egyik nagyimhoz</i> |
| értelmező j. | <i>Torzonborz,</i> <i>rabló</i> | <i>ide, az oviba</i> | <i>itt, Halason</i> | <i>összevesztem a ta-</i> <i>nárral, a történe-</i> |

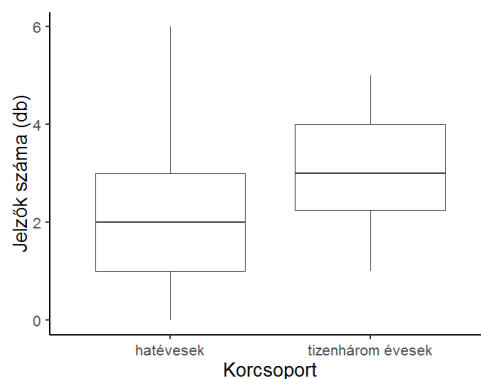
2. táblázat. Példák a korpuszból a jelzős szintagmák altípusaira életkoronként és nemenként

3.3 A mellérendelő szintagmák gyakorisága

A mellérendelő szerkezetek között adatoltam a kapcsolatos, a választó, az ellentétes, a magyarázó és a következtető típusú szintagmák előfordulását. A hatéveseknél összesen 23 db kapcsolatos, 5 db választó, 3 db ellentétes szintagma fordult elő, magyarázó és következtető szintagmára nem találtam példát. A tizenhárom éveseknél 35 db kapcsolatos, 6 db választó, 2 db ellentétes szintagma volt, magyarázó és következtető szintagmára ebben a csoportban sem találtam példát (4. ábra).

A csoportátlag a hatéveseknél 2,21 db előfordulás az összes típus figyelembe vételével, az idősebbeknél 3,07 db 200 szóra vetítve. Az eltérés a két csoport között nem szignifikáns, de a tendencia itt is kimutatható: a tizenhárom évesek beszédében jelent meg több mellérendelő szintagma. A mellérendelő szintagmákon belül a hatéveseknél az összes előfordulás 74,19%-a kapcsolatos, 16,12%-a választó, 9,67%-a ellentétes szintagma. A tizenhárom éveseknél a szintagmák 81,39%-a kapcsolatos, 13,95%-a választó, 4,65%-a ellentétes szerkezet. Megfigyelhető tehát, hogy mindkét csoportban a leggyakoribb a kapcsolatos szintagma, ezt követi a választó, majd az ellentétes típusú szerkezet.

A mellérendelő szintagmatípusok eloszlásában minőségi eltérés nem figyelhető meg a két csoport között, vagyis az életkor nincs hatással arra, hogy a mellérendelő szintagmák melyik típusát alkalmazzák a gyermekek spontán beszédükben kisebb vagy nagyobb. Nemek tekintetében a hatéves fiúknál 200 szóban átlagosan 2 db, a tizenhárom éves fiúknál átlagosan 3,2 db, míg a hatéves lányoknál átlagosan 2,4 db, a tizenhárom éves lányoknál átlagosan 2,8 db mellérendelő szintagma fordult elő. Az adatok előfordulásának gyakoriságát az 5. ábra, a korpuszból vett nyelvi példákat a 3. táblázat szemlélteti.



5. ábra. Az összes mellérendelő szintagmatípus gyakorisági eloszlása a két életkori csoportban

| kísérleti csoportok | kapcsolatos | választó | ellentétes |
|------------------------|---|---|---|
| hatéves fiúk | <i>pöttyösek és citromsárgák; valamikor anya, valamikor apa;</i> | <i>hét óra vagy hat óraker; a tizennégyes villamos vagy a tizenkettes</i> | - |
| hatéves lányok | <i>fürödtünk, napoztunk; játszunk felhőset meg földeset</i> | <i>fürödtem tengerbe vagy a medencébe; hétkor kel az ember vagy ötkor vagy hatkor</i> | <i>otthon nem, csak Helyén</i> |
| tizenhárom éves fiúk | <i>van külön női és férfi kategória; hétfőn és csütörtökön van az edzésem</i> | <i>zsírkrétát vagy lyukasztót kapnak; kosarazok vagy sportolok valamit</i> | <i>nem hajnali kettőig, hanem csak héttig</i> |
| tizenhárom éves lányok | <i>két nagymamám van és egy nagypapám; van reggeli, ebéd, vacsora</i> | <i>ő is megbetegedett vagy valami ilyesmi</i> | - |

3. táblázat. Példák a korpuszból a mellérendelő szintagmák altípusaira életkoronként és nemenként

4 Összegzés és következtetések

A jelen kutatás eredményei szerint a tizenhárom évesek spontán megnyilatkozásaikban tendenciaszerűen gyakrabban használnak összetett szavakat, jelzős és mellérendelő szintagmákat, mint a hatévesek. Ezek alapján első hipotézisünk, mely szerint az összetett szavak nagyobb arányban jelennek meg az idősebb korosztályban, teljesült, azonban a különbség a két életkori csoport között nem szignifikáns. Második hipotézisünk, mely szerint a jelzős szintagmák nagyobb arányban fordulnak elő a tizenhárom éveseknél, szintén teljesült: a különbség szignifikáns volt a két életkori csoport között. Harmadik hipotézisünk, mely szerint a mellérendelő szerkezetek ugyancsak

nagyobb arányban fognak megjelenni az idősebbeknél, igazolódott, de az eltérés ebben az esetben sem szignifikáns, viszont a tendencia kimutatható.

A jelzős és mellérendelő szintagmák egyes altípusainak gyakorisági eloszlása hasonlóan alakult a két csoportban, amely típus gyakori volt a hatéveseknél, az ugyancsak gyakori volt a tizenhárom éveseknél, amely típusok viszont ritkábbak voltak a fiatalabbaknál, azok ritkábban jelentkeztek az idősebbek szövegalkotásában is. Vagyis mindegyre az életkor nincs hatással. Megállapítható továbbá, hogy a vizsgált korcsoportok spontán beszédében a minőség- és mennyiségjelzős szerkezetek jelentkeznek a legnagyobb arányban a jelzőtípusok között, a minőségjelzők dominálnak mindkét korcsoportban. A mellérendelő szerkezetek között mindkét csoportban a kapcsolatos szerkesztés fordult elő a legtöbbször, ezt követte a választó, majd az ellentétes típusú szintagma, azonban utóbbi két altípus gyakorisági aránya jelentősen elmaradt a kapcsolatos szintagmák arányától.

A nemi összevetést csak fenntartásokkal lehet elvégezni az adatközlők kis száma miatt, ugyanakkor megfigyelhető volt, hogy a két életkori csoporton belül csaknem minden nyelvi kategória esetében a lányok beszédében volt adatolható több előfordulás, bár az eltérés kismértékű volt.

Az eredményekből azt a következtetést lehet levonni, hogy a hatéves kisiskolásokhoz képest tizenhárom éves korra a spontán beszéd kismértékben ugyan, de összetettebbé válik grammatikai szerkesztettségét tekintve, hiszen 200 szóra vetítve az idősebbek beszédében minden mutatóban több előfordulást lehetett adatolni. Ugyanakkor statisztikai értelemben – a jelzős szintagmák kivételével – nem adatoltunk eltérést, a két korcsoport a nyelvi fejlettség fokát tekintve nem áll igazán távol egymástól – ez adja a kutatás érdekességét. A magyarázat az lehet, hogy mindkét életkori szakaszban már részei a gyermekek az iskolai képzésnek, ahol nyelvi kompetenciájuk, pragmatikai és metanyelvi tudásuk nagyobb mértékben fejlődik, mint az óvodában eltöltött évek alatt. Mindkét korosztály ekkor még az általános iskolai képzésbe jár, melynek a szerkezete, a tanulnivaló mennyisége még nem éri el a középiskolai szintet, vagyis a terhelés hasonló. A jelzős szintagmák gyakoriságában tapasztalható nagyobb eltérés egyik oka valószínűleg az, hogy a tizenhárom évesek szókincse már fejlettebb és bővebb, emiatt valószínűsíthetően bátrabban használnak minősítő jelzőket a beszéd formálásakor.

A kutatást ki lehet terjeszteni a jelen vizsgálat középpontjában álló korosztályoknál fiatalabb vagy idősebb életkori csoportokra, mint a kisgyermekek és a kamaszok, a kamaszok és a fiatal felnőttek vagy a kamaszok és a középkorúak. Ha az említett életkori csoportokban is hasonló mennyiségi mutatókat adatolnánk, mint a jelen vizsgálati csoportokban, ez képet adna arról, hogy a magyar spontán beszédben az elemzett nyelvi kategóriák milyen arányban fordulnak elő. További adalékul szolgálhat a meglévő eredmények mellé többek között az elliptikus szerkezetek és az átlagosnál hosszabb beszédszakaszok szerkezetének elemzése, vagy a néma és kitöltött szünetek és a beszédjellel kitöltött szakaszok egymáshoz való viszonyulásának leírása is.

Köszönetnyilvánítás



A jelen vizsgálat az Emberi Erőforrások Minisztériuma ÚNKP-17-3 kód-számú Új Nemzeti Kiválóság Programjának támogatásával készült.

Irodalom

- Boersma, P., Weenink, D. 2018. *Praat: doing phonetics by computer* [Computer program]. Version 6.0.43, from <http://www.praat.org/>. Letöltés ideje: 2018. szeptember 8.
- Bóna, J. 2012. A spontán beszéd sajátosságai idősödő, idős és matuzsálemi korban. In: Markó, A. (szerk.) 2012. *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngékezdési időig*. Budapest: ELTE BTK–MTA NYTI. 100–115.
- Deme, A. 2012. Óvodások magánhangzóinak akusztikai jellemzői. In: Markó, A. (szerk.) 2012. *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngékezdési időig*. Budapest: ELTE BTK–MTA NYTI. 77–99.
- Érsok, N. Á. 2006. Szóbeliség és/vagy írásbeliség. *Magyar Nyelvőr*, 130(2): 165–176.
- Gerebenné, Várbíró K., Gósy M., Laczkó M. 1992. *Spontán beszéd megnyilvánulások szintaktikai elemzése DSS technika elemzésével*. Kézirat. Budapest.
- Gósy, M. 2003. Virtuális mondatok a spontán beszédben. *Beszédkutatás 2003*: 19–43.
- Gósy, M. 2004. *Fonetika, a beszéd tudománya*. Budapest: Osiris Kiadó.
- Gósy, M. 2005. *Pszicholingvisztika*. Budapest: Osiris Kiadó.
- Halliday, M. A. K. 1979. Differences between spoken and written language. In: *Communication through reading. Proceeding of the forth Australian Reading Conference*. Vol. 2. Adelaide, SA, 37–52.
- Horváth, V. 2013. Temporális szerveződés kilencéves gyermekek spontán beszédében. *Beszédkutatás 2013*. 144–159.
- Horváth, V. 2016. Kisiskolás gyermekek spontán beszédének jellemzői. *Alkalmazott Nyelvtudomány*, 2016(1): 1–14.
- Horváth V. 2017. Közlések grammatikai szerkesztettsége 6–9 éves gyermekek narratíváiban. *Anyanyelv-pedagógia*, 10(4): 5–18.
- Keszler, B. 1983. Kötetlen beszélgetések mondat- és szövegtani vizsgálata. In: Rácz, E., Szathmári I. (szerk.): *Tanulmányok a mai magyar nyelv szövegtana köréből*. Budapest: Tankönyvkiadó. 164–202.
- Keszler, B. (szerk.) 2010. *Magyar grammatika*. Budapest: Nemzeti Tankönyvkiadó.
- Laczkó, M. 2009. Középisikolás tanulók beszédének temporális jellemzői. *Magyar Nyelvőr*, 133(4): 447–467.
- Laczkó, M. 2010. Megakadásjelenségek a spontán és a szóközi beszédben. *Beszédkutatás 2010*: 184–198.
- Laczkó, M. 2011. Óvodások és kisiskolások spontán mondatalkotási folyamatai. *Magyar Nyelvőr*, 135(4): 440–459.
- Lanstyák, I. 2009. *A magyar beszélt nyelv sajátosságai*. Pozsony: Stimul.
- Levelt, W. 1989. *Speaking. From intention to articulation*. Cambridge: MIT Press, Cambridge MA.
- Markó, A., Dér Cs. 2011. Diskurzusjelölők használatának életkori sajátosságai. In: Navracsics, J, Lengyel Zs. (szerk.): *Lexikai folyamatok egy- és kétnyelvű közegben. Pszicholingvisztikai Tanulmányok II*. Budapest: Tinta Könyvkiadó. 49–61.
- Neuberger, T. 2011. Gyermekek spontán beszédének szerkesztettsége és folyamatossága. *Beszédkutatás 2011*: 83–95.
- Neuberger, T. 2012a. Virtuális mondatok gyermekek spontán beszédében. *Beszédkutatás 2012*: 217–233.

- Neuberger, T. 2012b. Szólóhívás gyermekek szóasszociációiban és spontán beszédében. In: Navracsics, J., Szabó D. (szerk.) *A mentális folyamatok a nyelvi feldolgozásban. Pszicholingvisztikai tanulmányok III.* Budapest: Tinta Könyvkiadó. 85–94.
- Neuberger, T. 2014. *A spontán beszéd sajátosságai gyermekkorban.* Budapest: ELTE Eötvös Kiadó.
- Rice, 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*, 5(2): 333–349.
- Vallent, B. 2010. Középiskolások spontán narratíváinak jellemzői. *Beszédkutató 2010*: 199–210.
- Wacha, I. 1988. Élő nyelvi (spontán) szövegek megnyilatkozásainak (szintaktikai) vizsgálati szempontjaihoz. (A gazdagréti kábeltelevízió élő nyelvi felvételei alapján). In Kontra, M. (szerk.): *Beszélt nyelvi tanulmányok.* Linguistica Series A. Studia et dissertationes 1. Budapest: MTA Nyelvtudományi Intézet. 102–158.

Források

- Bóna, J., Imre A., Markó A., Váradi V., Gósy, M. 2014. GABI – Gyermeknyelvi Beszédadatbázis és Információtár. *Beszédkutató 2014*: 246–251.

A szóidőtartam és a beszédszakasz hosszának összefüggései spontán beszédben

Huszár Anna

ELTE BTK Nyelvtudományi Doktori Iskola
zicsyhars@gmail.com

Kivonat: A beszéd időzítési viszonyait számos tényező befolyásolja. A beszédhangok időtartamának vizsgálatával több kutatás foglalkozott, mint a szavak időtartamának alakulásával. Mind a hazai, mind a nemzetközi szakirodalomban igazolták a szóhosszúság hatását a beszédhangok időtartamára: a több szótagból álló szavakat alkotó beszédhangok időtartama általában rövidebb. Ez a jelenség nagyobb prozódiai egységek (például beszédszakaszok, azaz szünettől szünetig tartó egységek) esetében is kimutatható lehet. Kutatásomban arra a kérdésre keresem a választ, hogy miként befolyásolja a beszédszakaszok hossza a szavak időtartamát a spontán beszédben. Az eredmények szerint a beszédszakaszok hosszának növekedésével csökken a szavak időtartama, ez a csökkenő tendencia a 3-4 szavas beszédszakaszokig mutatható ki.

1 Bevezetés

A beszéd időzítési viszonyait számos tényező befolyásolja: a beszédstílus, a beszédtema, a beszédhelyzet, a beszélő életkora, személyisége, érzelmi állapota, szociológiai háttere stb. (Gósy 2004). Az időviszonyok jellemezhetőek a tempóval (beszéd- és artikulációs tempó), de az egyes szegmentumok időtartamaival is.

A szavak időtartamára jelentős hatással van a gyakoriság, többen megállapították, hogy a gyakoribb szavak rövidebb időtartamban realizálódnak, mint a hasonló szerkezetű, de ritkábban előforduló szavak (Zipf 1929; Aylett–Turk 2004). A szavak időtartamának alakulását befolyásolja továbbá a megjósolhatóság, vagyis egy szó adott szövegkörnyezetben való előfordulásának a valószínűsége: az előre könnyebben jósolható szavak vagy szótagok rövidebb időtartamban valósulnak meg, mint az előre nehezen jósolhatók (Baker–Bradlow 2009; Bell et al. 2009). A szó típusa (tartalmas vagy funkciószó) szintén befolyásolhatja a szavak időtartamát. Bell és munkatársai (2009) azt találták, hogy a gyakoriság és a megjósolhatóság másképpen hat tartalmas és a funkciószavak időtartamára. A tartalmas szavak időtartama csökken, ha a szó gyakori, vagy ha az adott szövegben már többször előfordult. Ez a jelenség a funkciószavaknál is kimutatható, azonban kisebb mértékben. A szavak időtartamát a beszédstílusa is befolyásolja, Baker és Bradlow (2009) a szógyakoriság, az ismétlések (első vagy többedik említés az adott szövegben) és a beszédstílus (alul- és túlartikulált beszéd) együttes hatását vizsgálta a szavak időtartamára. Ők is kimutatták a gyakoribb szavak időtartambeli rövidülését és azt is, hogy a többször előforduló szavak első említése hosszabb időtartamban valósul meg. Az alulartikulált beszédben, a szavak idő-

Ludányi Zsófia, Krepsz Valéria, Gráczki Tekla Etelka (szerk.): *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2018*. Budapest: MTA Nyelvtudományi Intézet. 2018.

tartama általában rövidebb volt. Az ismételt szavak első említése mindkét beszédstílusban hosszabb időtartamban realizálódott, mint a második említésnél. A gyakoriság redukciós hatása az eredmények szerint csak az alulartikulált beszédben érvényesült.

Magyar nyelvre is több kutatás született már a szavak időtartamával kapcsolatban. Gósy (1997) spontán beszédben elemezte a szavak szótagszáma, a szavakat alkotó beszédhangok száma és időtartama közti összefüggéseket; és a szóidőtartamok nem lineáris növekedését tapasztalta mind a szótagszám, mind a beszédhangok számának emelkedésével. Ezen túl a gyakoriság időtartamra gyakorolt hatását is vizsgálta, és megállapította, hogy a gyakran ejtett szavak tendenciaszerűen rövidebb időtartamban valósulnak meg. Gósy és Horváth (2010) a *tehát* és *ilyen* szavak időzítési viszonyait elemezték, és azt találták, hogy a szavak funkciója is befolyásolhatja azok időtartamát: az adott szavak töltelékshöz ként rövidebb időtartammal realizálódnak, mint eredeti funkciójukban. Horváth (2014) a szavak időtartamát a beszélő életkora, a szó szótagszáma és morfológiai felépítése függvényében vizsgálta. Kisiskolások (9 évesek) és felnőttek spontán beszédét hasonlította össze. Eredményei azt mutatják, hogy a kisiskolások szóidőtartami hosszabbak, mint a felnőtt beszélők. A morfológiai struktúra szintén hatással van a szavak időtartamára, a bonyolultabb szerkezetű szavak időtartama hosszabb, azonban mindezt felülírhatja a gyakoriság redukciós hatása.

A beszédhangok időzítésére szintén több tényező hatással van: a közlésben elfoglalt pozíció, a hangsúlyosság, a szóhosszúság stb. (Lindblom 1968; Klatt 1975). A frázisvégi nyúlás jelenségét több nyelvben is kimutatták: többek között a svédben (Lindblom 1968), a spanyolban (Oller 1973), az angolban (Klatt 1975), a németben (Kohler 1983), a hollandban (Cambier-Langeveld 1997) a franciában (Tabain 2003), a finnben (Nakai et al. 2009) és az oroszban (Kachkovskaia (2014) is.

A magyarban is kimutatták a frázisvégi nyúlás jelenségét (Magdics 1966, Hockey–Fagyal 1999). White és Mány (2008) felolvasott mondatokban vizsgálta hosszú és rövid magánhangzók időtartamának alakulását frázisvégi pozícióban. Megállapították, hogy a rövid magánhangzók esetében nyúlás csak az utolsó szótagon, a hosszú magánhangzók esetében az utolsó előtti szótagon is tapasztalható. Markó és Kohári (2015) a frázisvégi nyúlás és a glottalizáció összefüggéseit elemezte felolvasásban és spontán beszédben. Eredményeik szerint a frázis végének jelzésében a tempó lassulása és a glottalizáció is fontos szerepet játszik. A frázisvégi lassulások nagyobb mértékben jelentkeztek a spontán beszédben, mint felolvasásban. Ezzel szemben a glottalizáció gyakoribb volt a felolvasásokban. Gósy és Krepsz (2016) a frázisvégi nyúlás és a szavak szótagszámának összefüggéseit vizsgálta felolvasott mondatokban. Kimutatták, hogy ugyanazon szavak utolsó szótagja akkor a leghosszabb, ha a szó a mondat végén helyezkedik el. A szavak szótagszáma szintén hatással van a magánhangzók időtartamára, a több szótagból álló szavakban a magánhangzók rövidebb időtartamban realizálódtak, mint a kevesebb szótagú szavakban. Krepsz (2017) spontán beszédben vizsgálta a frázisvégi nyúlás jelenségét. Megállapította, hogy nemcsak a magánhangzók, hanem a mássalhangzók is hosszabb időtartamban realizálódnak frázisvégen.

Menzerath (1928) azt a megállapítást tette, hogy egy nyelvi egység hossza és az azt felépítő kisebb egységek hossza között nem egyenes arányosság van, tehát nem egyenletesen növekszik a szótagszám emelkedésével a szóidőtartam, hanem a több szótagból álló szavakat relatíve rövidebben ejtjük, mint a kevesebb szótagosakat. Azt, hogy a szavak hossza befolyásolja a beszédhangok időtartamát, más nyelvekben is kimutatták (Lindblom 1968; Turk–Shattuck–Hufnagel 2000).

Magyar nyelvre is több kutatás készült, izolált szavakra Gombocz és Meyer (1909), majd később Tarnóczy (1974) végzett kísérletet. A *tát, tátog, tátogat, tátogatók* és *tátogatóknak* szavak magánhangzóinak időtartamait vizsgálták; és kimutatták, hogy ugyanazon hangkörnyezetben a szót felépítő magánhangzók időtartamai és a szóhosszúság között fordított arányosság van, tehát minél több szótagból áll egy szó, annál rövidebbek az azt alkotó magánhangzók időtartamai. Kassai (1993) ugyanezen szavakat vizsgálta mondatokban, és hasonló eredményekre jutott.

A jelenséget azonban nem minden kutatás támasztotta alá, White és Mády (2008) egy, két és három szótagos szavakban vizsgálta a magánhangzók időtartamainak alakulását, és nem talált összefüggést a szavak szótagszáma és az azokat alkotó magánhangzók időtartamai között. Más kutatás a hosszú magánhangzók rövidülését adatolta a szóhosszúság növekedésével hangsúlytalan helyzetben, a rövid magánhangzók esetében nem tapasztalt rövidülést (Kohári 2012). Krepsz (2015) a spontán beszédben mutatta ki a magánhangzók időtartam-csökkenését a szóhosszúság növekedésével.

Valószínűsíthető, hogy a szótagszám időtartamra gyakorolt hatása nemcsak a szavak szintjén, hanem nagyobb egységek (pl. mondatok, beszédszakaszok) szintjén is érvényesül. Többen igazolták felolvasott mondatokra, hogy ugyanazon szavak időtartama csökken, ha azok hosszabb mondatokban szerepelnek (Lehiste 1974, Rakerd et al. 1987). Spontán beszédre szintén azt találták, hogy rövidebb beszédszakaszokban általában lassabb az artikulációs tempó (Quené 2005, 2008), illetve kevesebb szót mondanak adott idő alatt (Yuan et al. 2006). Yuan és munkatársai (2006) kimutatták, hogy a beszédszakaszok hosszának az időtartamra gyakorolt hatása nem egyenletes, az angolban hétszavas beszédszakaszokig figyelhető meg a szavak időtartamának rövidülése.

Jelen kutatásomban arra a kérdésre keresem a választ, hogy a beszédszakaszok hossza miként befolyásolja a szavak időtartamát a spontán beszédben. Mennyiben egyeznek meg, illetve térnek el az eredmények az eddigi, más nyelvekre megállapított tendenciáktól. Hipotéziseim szerint (i) a beszédszakaszok hosszának növekedésével csökken az adott szótagszámú szavak időtartama, (ii) a feltételezett összefüggés a beszédszakasz hossza és a szóidőtartam között csak a rövidebb beszédszakaszok esetében mutatható ki.

2 Kísérleti személyek, anyag, módszer

Kutatásomhoz a Beszélt nyelvi adatbázisból (BEA) választottam az adatközlőket (Gósy et al. 2012). 10 adatközlő (5 nő és 5 férfi) spontán narratíváját elemeztem. A beszélők életkora 25 és 30 év közé esett. Az adatközlőknek a spontán narratíva részben a munkájukról és a hobbiukról kellett beszélniük. Beszélőnként 3 perces részeket választottam ki a narratívák közepéből.

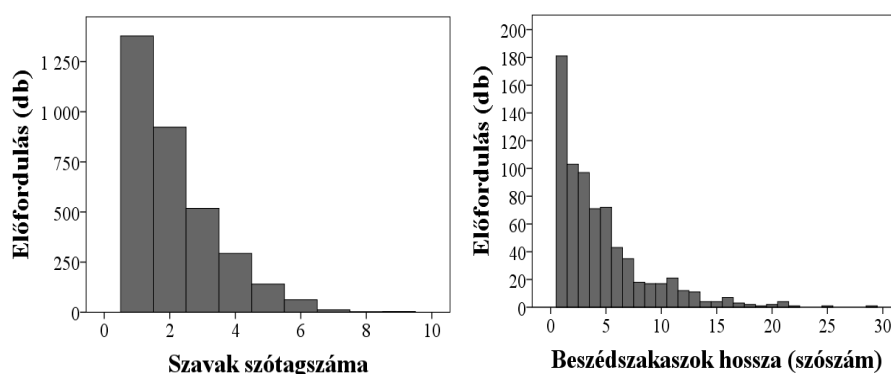
A hanganyagot a Praat szoftverben (Boersma–Weenink 2018) annotáltam beszédszakasz- és szószinten a hangszínek és a rezgéskép alapján, auditív ellenőrzés mellett. A beszédszakaszok elkülönítésénél a néma és kitöltött szüneteket vettem határjelölőnek. Az időtartamok automatikus kinyeréséhez egy erre a célra készített Praat szkriptet használtam.

Összesen 3331 szót és 726 beszédszakaszt adatoltam. Ez személyenként átlagosan 333 szót és 73 beszédszakaszt jelentett. A beszélőnkénti adatokat az 1. táblázat mutatja.

| Beszélő sorszáma | Szavak száma (db) | Beszédszakaszok száma (db) |
|------------------|-------------------|----------------------------|
| 1n | 374 | 54 |
| 2n | 382 | 65 |
| 3n | 339 | 77 |
| 4n | 337 | 54 |
| 5n | 389 | 52 |
| 6f | 280 | 76 |
| 7f | 292 | 63 |
| 8f | 291 | 101 |
| 9f | 279 | 93 |
| 10f | 368 | 93 |

1. táblázat. Szavak és beszédszakaszok száma beszélőnként (n = nő; f = férfi)

A szavak szótagszám szerinti és a beszédszakaszok szószám szerinti eloszlása nem volt egyenletes (1. ábra), mindkét esetben a rövidebb egységek előfordulása volt gyakoribb. Az elemzésekből – az alacsony előfordulásuk miatt – kizártam a hosszabb szavakat (a 6 szótagosnál hosszabbakat) és a hosszabb beszédszakaszokat (a 13 szótagosnál hosszabbakat), mivel előfordulásuk együttesen sem haladta meg az 5%-ot.



1. ábra. A szavak szótagszám szerinti és a beszédszakaszok szószám szerinti előfordulása

A statisztikai elemzéseket az SPSS 20.0 szoftverrel végeztem. A szavak időtartamára ható tényezők megállapításához egy többváltozós lineáris regressziós modellt építettem fel, melyben a függő változó a szóidőtartam (ms-ban kifejezve), a független változók pedig a beszédszakasz hossza (szószámban kifejezve) és a szavak szótagszáma voltak. A szavak szótagszámát azért láttam szükségesnek belevenni az elemzésbe, mivel a szótagszám jelentősen befolyásolja a szavak időtartamát (Menzerath 1928).

A nemek közti esetleges különbségek megállapításához a Mann–Whitney-próbát, az egyének közti eltérések megállapításához a Kruskal–Wallis-próbát alkalmaztam 5%-os szignifikanciaszinten.

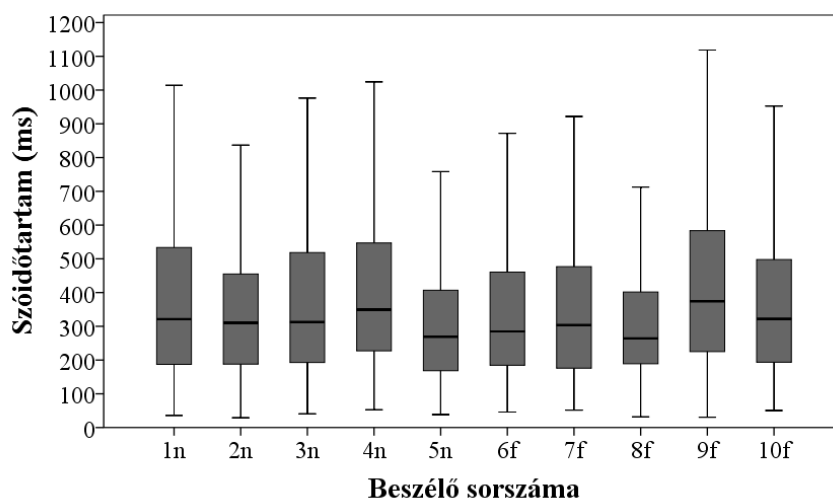
3 Eredmények

3.1 Nemek közti különbségek

A nemek között nem feltételeztem különbséget a szóidőtartamok tekintetében, azonban szükségesnek tartom ennek ellenőrzését. A szóidőtartamok átlaga a nők esetében 365,56 ms (szórás: 233,59 ms), a férfiak esetében 365,45 ms (szórás: 232,49 ms). A Mann–Whitney-próba nem mutatott szignifikáns különbséget a két nem szóidőtartamai között ($Z = -0,122$; $p = 0,903$), így a továbbiakban a nők és férfiak adatait nem vettem külön az elemzések során.

3.2 Egyéni különbségek

A szóidőtartamok beszélőnkénti alakulását a 2. ábra mutatja.



2. ábra. Egyéni különbségek a szavak időtartamában (n = nő; f = férfi)

A Kruskal–Wallis-próba szignifikáns különbséget mutatott a beszélők szóidőtartamai között ($\chi^2 = 68,68$; $p < 0,001$), azonban nem az összes adatközlőnél. Nem volt szignifikáns különbség a következő esetekben: az 1., 2., 3. és 10. adatközlő; az 5., 6., 7. és 8. adatközlő és a 4. és a 9. adatközlő adatai között. A beszélők három csoportra való bontása azonban lényegesen csökkentette volna az szóidőtartamok számát az egyes kategóriákban, így nem bontottam csoportokra az adatokat, hanem az egyéni különbségek csökkentése érdekében normalizáltam a szóidőtartamokat a következő képletet használva:

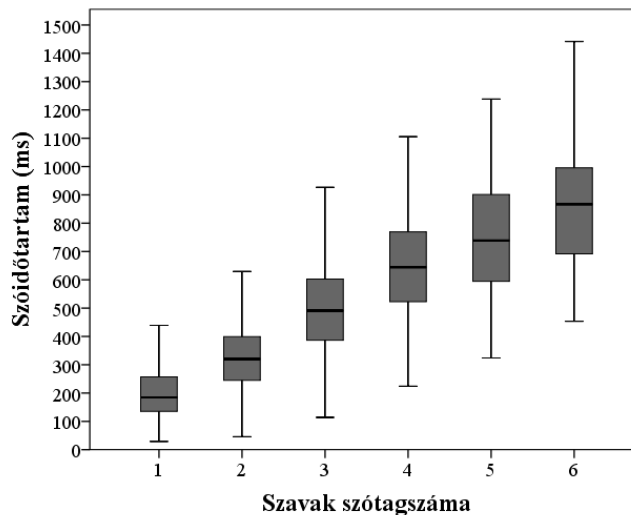
$$(1) \frac{E_{ij} - \bar{E}_i}{\sigma_i}$$

Ahol t_k a k -edik szó időtartama, \bar{t} a szóidőtartamok átlaga és \bar{t}_i az i -edik beszélő átlagos szóidőtartama. A további vizsgálatokban a normalizált szóidőtartamokkal dolgoztam.

3.3 A beszédszakasz hosszának és a szavak szótagszámának hatásai a szóidőtartamra

Egy többváltozós lineáris regressziós modellel vizsgáltam a szavak szótagszámának és a beszédszakasz hosszának a szóidőtartamra gyakorolt hatásait. Az eredmények szerint a felállított modell szignifikáns magyarázóerővel bír ($F = 2355,926$; $p < 0,001$), determinációs együtthatója: $R^2 = 0,629$. Ez az érték arra utal, hogy a szavak időtartamának alakulását 62,9%-ban a szavak szótagszáma és a beszédszakaszok hossza magyarázza. Az egyes független változók külön-külön is szignifikánsan magyarázzák a szóidőtartam alakulását ($B_{szótagszám} = 66,86$; $B_{szakaszhossz} = -11,756$; $p < 0,001$).

A regressziós együtthatók értékeiből megállapítható, hogy a szótagszám erősebb hatással van a szavak időtartamára, mint a beszédszakasz hossza. A szóidőtartamok szótagszám szerinti alakulását a 3. ábra szemlélteti.

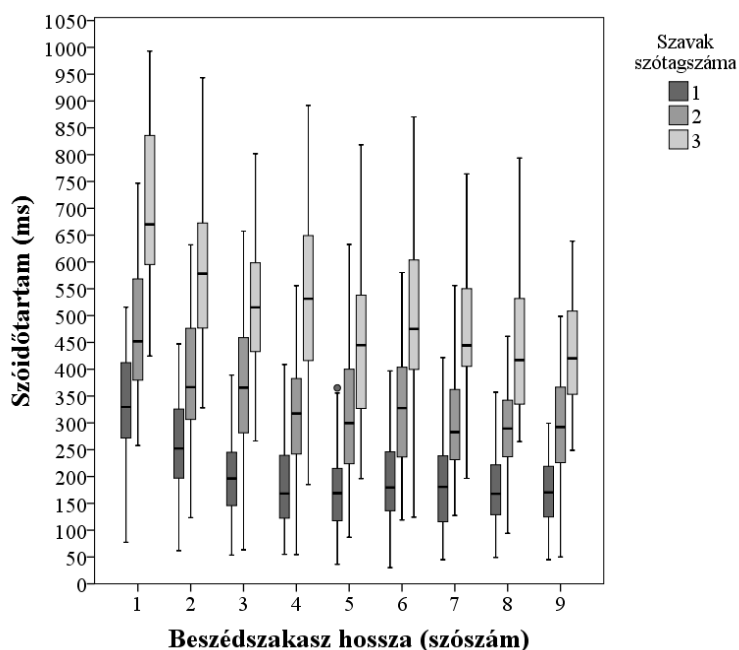


3. ábra. Szavak időtartamának alakulása a szótagszám szerint

Az ábráról leolvasható, hogy a szótagszám emelkedésével nő a szavak időtartama, a szóidőtartamok növekedésének mértéke azonban a szótagszám emelkedésével csökken, megfigyelhető tehát a kiegyenlítődési tendencia (Menzerath 1928).

A szavak szótagszámán kívül a beszédszakasz hossza is szignifikánsan befolyásolja a szóidőtartamokat. A regressziós együttható negatív előjele ($B = -11,756$) a két mennyiség közötti fordított arányosságra utal; tehát a beszédszakaszok hosszának növekedésével csökken az azokat alkotó szavak időtartama (4. ábra).

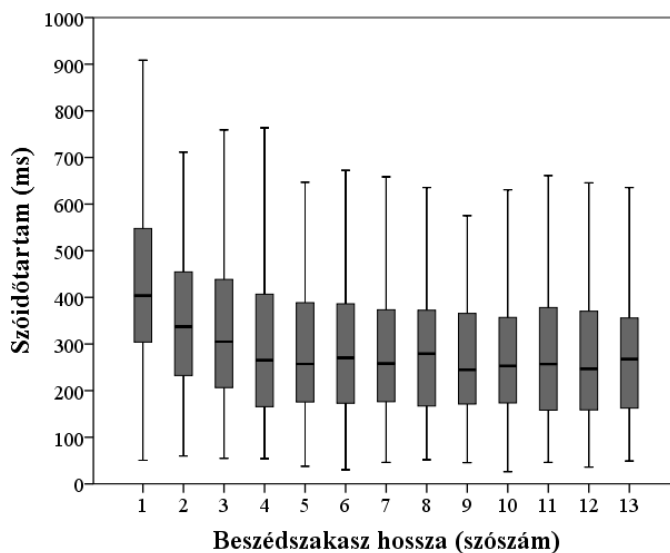
Az ábráról leolvasható, hogy a rövidebb beszédszakaszok esetében valóban csökken a szavak időtartama a beszédszakasz hosszának növekedésével. A hosszabb beszédszakaszoknál azonban nem olvasható le jelentősebb eltérés az adott szótagszámú szavak időtartamai között. Ez a tendencia az egy szótagos szavaknál rajzolódik ki a leginkább, a szavak szótagszámának emelkedésével kevésbé egyenletes a szóidőtartamok alakulása (pl. a három szótagú szavak időtartamainak mediánja nagyobb a négyzavas beszédszakaszokban, mint a háromszavasakban; a két szótagú szavak időtartamainak mediánja nagyjából megegyezik a két- és háromszavas beszédszakaszokban).



4. ábra. A szavak időtartamának alakulása a beszédszakasz és a szótagszám függvényében

A szavak időtartamát az eddig említett paramétereken túl a szavaknak a beszédszakaszokban elfoglalt helye is befolyásolhatja, ismeretes a frázisvégi nyúlás jelensége (Fletcher 2010). A szavak pozícióját nem tettem bele a lineáris regressziós modellbe, ahogy az ábrákon sem jelöltem külön az eltérő pozícióban lévő szavak időtartamát; hanem feltételeztem, hogy elegendő adatmennyiség esetén a pozícióhatás kiegyenlítődik, tehát az egyes csoportokban hasonló arányban jelennek meg a különböző pozícióban lévő szavak. Ebből is következhet az, hogy az egyre több szótagból álló szavaknál nehezebben fedezhető fel tendencia a szóidőtartamban, hiszen a mért adatok száma is jelentősen kevesebb (egy szótagos szavakból 1378, két szótagosakból 923, három szótagosakból 518 szóidőtartamot elemeztem, az ennél hosszabb szavak száma összesen sem érte el az 500-at). A továbbiakban érdemes lehet bővíteni a regressziós modellt egyéb független változók bevonásával (pl. frázispozíció), növelve ezzel a modell magyarázóerejét.

Az 5. ábra egyben mutatja az egy, két és három szótagos szavak időtartamát a beszédszakaszok hosszának függvényében.



5. ábra. A szóidőtartamok alakulása a beszédszakaszok hosszának függvényében

Az ábráról leolvasható, hogy az említett csökkenő tendencia a 3-4 szavas beszédszakaszokig figyelhető meg.

Az adatokat két részre osztva (1. ahol a beszédszakaszok hossza ≤ 4 szó; 2. ahol a beszédszakaszok hossza ≥ 4 szó) elvégezhető a lineáris regressziószámítás külön-külön a két csoportra. A rövidebb beszédszakaszok esetében a lineáris regressziós modell szignifikáns magyarázóerővel bír ($F = 58,262$; $p < 0,001$), míg a hosszabb beszédszakaszok esetében nem ($F = 0,122$; $p = 0,727$), tehát megállapítható, hogy csak a rövidebb (maximum 4 szavas) beszédszakaszok hossza van hatással a szavak időtartamára.

4 Következtetések

A jelen kutatásban a beszédszakaszok hosszának hatását vizsgáltam a szavak időtartamára. Az első hipotézisem az volt, hogy a beszédszakaszok hosszának növekedésével csökken a különböző szótagszámú szavak időtartama. Ezt a hipotézist az eredmények megerősítették, hiszen valóban megfigyelhető volt, hogy a rövidebb beszédszakaszokban hosszabb a szavak átlagos időtartama, mint a hosszabb beszédszakaszokban.

A második hipotézisem az volt, hogy a feltételezett összefüggés a beszédszakasz hossza és a szóidőtartam között csak a rövidebb beszédszakaszok esetében mutatható ki. Ezt a hipotézist szintén megerősítették az eredmények, a szóidőtartamban megfigyelt csökkenő tendencia a 3-4 szavas beszédszakaszokig volt kimutatható.

Megállapítható tehát, hogy egy nagyobb egység (beszédszakasz) mérete befolyásolja az azt felépítő kisebb egység (szó) időtartamát (Menzerath 1928). Ennek oka lehet, hogy létrejön egy kiegyenlítődési tendencia, amely csökkenti a különböző hosszúságú egységek időtartamai közötti különbséget. Ennek eredményeképpen nincs egyenes arányosság a beszédszakasz mérete és időtartama között, tehát a kevesebb szóból álló beszédszakaszokat általában lassabban, a több szóból álló beszédszakaszokat pedig gyorsabban ejtjük. Azt, hogy beszédünk során a beszédszakaszok időtartamainak kiegyenlítésére törekszünk, több okból is tehetjük: a levegővétele szempontjából is így kényelmesebb. A spontán beszédre azonban jellemző lehet, hogy eltérő az az információmennyiség, amelyet éppen közölni szeretnénk. A sok gondolatból álló közlést így relatíve gyorsan szavakba öntjük (kevesebb szünetet tartunk, ezáltal hosszabb beszédszakaszokat hozunk létre, melyekben az egyes szavak időtartamai rövidebbek, mint általában). Ha azonban gondolkodási időt szeretnénk nyerni, közlésünkben megnő a néma és kitöltött szünetek száma, és lassul a tempó is (ilyenkor gyakoribbak az egy-két szavas beszédszakaszok, melyekben a szavakat szokásosnál hosszabban ejtjük). Az eredmények által többet tudhatunk meg a szegmentális szerkezet és a prosódia összefüggéseiről.

Irodalom

- Aylett, M., Turk A. 2004. The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, 47: 31–56.
- Baker, R. E., Bradlow A. R. 2009. Variability in Word Duration as a Function of Probability, Speech Style, and Prosody. *Language and Speech*, 52: 391–413.
- Bell, A., Brenier J. M., Gregory M. L., Girand C., Jurafsky D. 2009. Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, 60: 92–111.
- Boersma, P., Weenink D. 2018. *Praat: doing phonetics by computer*. Elérhető: http://www.fon.hum.uva.nl/praat/download_win.html
- Cambier-Langeveld, T. 1997. The domain of final lengthening in the production of Dutch. In: de Hoop, H., Coerts J. A. (szerk.) *Linguistics in the Netherlands*. Amsterdam: John Benjamins. 13–24.
- Gombocz, Z., Meyer E. A. 1909. *Zur Phonetik der ungarischen Sprache*. Uppsala: Berlings Buchdruckerei.
- Gósy M. 1997. A szavak időzítési sajátosságai a spontán beszédben. *Beszédkutatás 1997*: 39–49.
- Gósy M. 2004. *Fonetika, a beszéd tudománya*. Budapest: Osiris Kiadó.
- Gósy M., Gyarmathy D., Horváth V., Grácsi T. E., Beke A., Neuberger T., Nikléczy P. 2012. BEA: Beszélt nyelvi adatbázis. In: Gósy M. (szerk.) *Beszéd, adatbázis, kutatások*. Budapest: Akadémiai Kiadó. 9–24.
- Gósy M., Krepesz V. 2016. A magánhangzó nyúlása a szótagszám és a mondatpozíció függvényében. *Beszédkutatás 2016*: 59–88.
- Gósy, M., Horváth V. 2010. Changes in articulation accompanying functional changes in word usage. *Journal of the International Phonetic Association*, 40(2): 135–161.
- Fletcher, J. 2010. The prosody of speech: Timing and rhythm. In: Hardcastle, William J., Laver, J., Gibbon, F. E. (szerk.): *The handbook of phonetic sciences*. 2nd edition. Wiley-Blackwell: Oxford. 521–602.
- Hockey, B. A., Fagyal Zs. 1999. *Phonemic length and pre-boundary lengthening: an experimental investigation on the use of durational cues in Hungarian*. In: Ohala, J. J. (szerk.) Proceedings of 14th International Congress of Phonetic Sciences. San Francisco, 1–7. augusztus. 313–316.
- Horváth, V. 2014. Szóidőtartamok gyermekek és felnőttek spontán narratíváiban. *Beszédkutatás 2014*: 87–97.

- Kachkovskaia, T. 2014. Phrase-final lengthening in Russian: Pre-boundary or pre-pausal? In: Ronzhin, Andrey – Potapova, Rodmonga – Delic, Vlado (szerk.) *Speech and computer*. Novi Sad: Springer International Publishing, 353–359.
- Kassai I. 1993. Gyorsult-e a magyar beszéd tempója az elmúlt 100–120 évben? *Beszédkutató 1993*: 62–69.
- Klatt, D. H. 1975. Vowel lengthening in syntactically determined in a connected discourse. *Journal of Phonetics*, 3: 129–140.
- Kohári A. 2012. Rövid és hosszú magánhangzók a szóhosszúság függvényében. In: Váradí T. (szerk.) VI. Alkalmazott Nyelvészeti Doktoranduszkonferencia. Budapest: MTA Nyelvtudományi Intézet, 80–88.
- Kohler, K. J. 1983. Prosodic boundary signals in German. *Phonetica*, 40: 89–134.
- Krepsz V. 2015. Magánhangzó-időtartamok alakulása a hangsor hossza és az életkor függvényében. In: Váradí Tamás (szerk.) 2015. XI. Alkalmazott Nyelvészeti Doktoranduszkonferencia. Budapest: MTA Nyelvtudományi Intézet, 17–29.
- Krepsz V. 2017. Szótag nyúlása a frázisvégen. In: Gósy M., Krepsz V. (szerk.) *Morfémák időzítési mintázatai a beszédben*. Budapest: MTA Nyelvtudományi Intézet, 156–174.
- Lehiste, I. 1974. Interaction between test word duration and length of utterance. *The Journal of the Acoustical Society of America*, 55: 160–169.
- Lindblom, B. 1968. *Temporal organisation of syllable production. Quarterly progress and status report 2/3*. Stockholm: Speech Transmission Laboratory.
- Magdics Klára 1966. A magyar beszédhangok időtartama. *Nyelvtudományi Közlemények*, 68. 125–139.
- Markó, A., Kohári A. 2015. Glottalization and timing at utterance final position in Hungarian: Reading aloud vs. spontaneous speech. In Wolters M. et al. (szerk.): *Proceeding of 18th International Congress of Phonetic Sciences*. Glasgow, 2015. augusztus 10–14. 722–726.
- Menzerath, P. 1928. Über einige phonetische Probleme. *Actes du premier congrès international de linguistes*. Sijthho: Leiden. 104–105.
- Nakai, S., Kunnari S., Turk A., Suomi K., Ylitalo R. 2009. Utterance-final lengthening and quantity in Northern Finnish. *Journal of Phonetics*, 39: 29–45.
- Oller, K. D. 1973. The effect of position in utterance on speech segment duration in English. *Journal of the Acoustical Society of America*, 54: 1235–1247.
- Quené H. 2005. Modeling of Between-Speaker and Within-Speaker Variation in Spontaneous Speech Tempo. *Proceedings of the 9th European Conference on Speech Communication and Technology*, Interspeech. Lisbon, 4–8. szeptember. 2457–2460.
- Quené, H. 2008. Multilevel modeling of between-speaker and within-speaker variation in spontaneous speech tempo. *Journal of the Acoustical Society of America*, 123: 1104–1113.
- Rakerd, B., Sennet W., Fowler C. A. 1987. Domain-final lengthening and foot-level shortening in spoken English. *Phonetica*, 44: 147–155.
- Tabain, M. 2003. Effects of prosodic boundary on /aC/ sequences: Articulatory results. *Journal of the Acoustical Society of America*, 113: 2834–2849.
- Tarnóczy, T. 1965. Can the problem of automatic speech recognition be solved by analysis alone? In: Commins, D. E. (szerk.) *Reports of the Fifth International Congress of Acoustics II*. Liege, 7–14. szeptember. 371–387.
- Turk, A. E., Shattuck-Hufnagel S. 2000. Word-boundary-related duration patterns in English. *Journal of Phonetics*, 28: 397–440.
- White, L., Mády K. 2008. The long and the short and the final: phonological vowel length and prosodic timing in Hungarian. In: Barbosa, P. A., Madureira S., Reis C. (szerk.) *Proceedings 4th Speech Prosody Conference*. Campinas, 6–9. május. 363–366.
- Yuan, J., Liberman M., Cieri C. 2006. *Towards an integrated understanding of speaking rate in conversation*. Elérhető: <https://pdfs.semanticscholar.org/f7b1/8ae6558b10305e504a24fcd79749aad93d3a.pdf> Letöltve: 2018. július 15.
- Zipf G. K. 1929. Relative frequency as a determinant of phonetic change. *Harvard Studies in Classical Philology*, 15: 1–95.

A *hát* diskurzusjelölő prozódiai jellemzői a gyermeknyelvben

Kondacs Flóra

SZTE BTK Nyelvtudományi Doktori Iskola
kondacs.flora@gmail.com

Kivonat: Kutatásomban a *hát* diskurzusjelölő prozódiai jellemzőivel foglalkozom az óvodások megnyilatkozásaiban. Azt vizsgáltam, hogy a 2,9–6,11 éves korosztály spontán beszédében fellelhetőek-e a *hát* két funkciója, a hezitáló és általános válaszjelölő között a prozódia tekintetében olyan különbségek, amelyek segíthetnek a két funkció elkülönítésében. A *hát* diskurzusjelölő prozódiai jellemzőit, vagyis az alapfrekvencia minimumát és maximumát, az időtartamot és a szünet jelenlétét és hiányát a Praat 6.0.31 szoftver segítségével előfordulásonként nyertem ki a hanganyagokból. Nemcsak a diskurzusjelölők, hanem a diskurzusjelölőket tartalmazó fordulók jellemzőit is megvizsgáltam. Az eredmények azt mutatják, hogy a nominális változókat (nemet, korcsoportot, helyszínt) is figyelembe véve a prozódiai jegyek közül a *hát* időtartama az, amelyik minden változó esetében szignifikáns eltérést mutat.

1 Bevezetés

A tanulmány első fejezetében a *hát* diskurzusjelölő jellemzőit ismertetem röviden az eddigi kutatások tükrében, majd a jelölők eddig leírt prozódiai jellemzőit fogalmazom meg részletesebben. A második részben a kutatásban felhasznált korpusz anyagáról és kísérleti személyeiről (2.1.), a felhasznált módszerről (2.2.) írok részletesebben. A harmadik fejezetben az elért eredményeket ismertetem, míg a negyedik rész az eredményekből levont következtetéseket tartalmazza.

A diskurzusjelölők (Fraser 1999, Dér 2005, Schirm 2007–8) olyan társalgásszervező elemek, melyek pragmatikai viszonyokat jelölnek, és diskurzusszegmenseket kötnék össze. Mind a felnőttek, mind a gyermekek nyelvhasználatában az egyik leggyakoribb diskurzusjelölő a *hát*, amely főleg élőszóban használatos (Dér–Markó 2007: 63). Bell (1998) a diskurzusjelölők jelentésével foglalkozott, és megalkotta a magperiféria modellt. Egy diskurzusszervező elem magjelentése, pragmatikai instrukciója alatt azt érti, hogy a diskurzusjelölő utasítást ad a hallgatónak a diskurzusszegmensek kapcsolatát és értelmezését illetően, a hallgató pedig az instrukció által vonja le a megfelelő következtetéseket. A társalgást szervező elemek interpretációja a jelölő szemantikájából, fonológiájából és szintaktikai tulajdonságaiból, illetve a központi instrukcióból és a kontextus kölcsönhatásából származik. A magjelentés mellett előforduló további jelentések lesznek a periférikus jelentések (Schirm 2011b: 186). Megfigyelhető, hogy a gyermekek elsőként a magjelentéseket sajátítják el az előfordulási

Ludányi Zsófia, Krepsz Valéria, Grácsi Tekla Etelka (szerk.): *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2018*. Budapest: MTA Nyelvtudományi Intézet. 2018.

gyakoriságuk miatt (Schirm 2011a: 24). A *hát* magjelentésének a bizonytalanság kifejezését tartják (Schirm 2011b), amely számos funkcióban megfigyelhető.

A *hát* diskurzusjelölő leggyakoribb gyermeknyelvi funkcióinak Markó és Dér (2011) az általános válaszelölő (1), a beszédtervezői, hezitáló (2) és a konklúziószerű használatát (3) emelték ki. Kutatásomban a saját korpuszomban a legtöbb adatot szolgáltatató két funkciót, a hezitálót és az általános válaszelölőt hasonlítottam össze a *hát* elemet tartalmazó forduló és az önállóan álló *hát* prozódiai jellemzőinek figyelembevételével.

A felnőttek *hát*-használatát megvizsgálva látható (Markó–Dér 2011a), hogy a *hát* általános válaszelölő szerepben (4) és beszédtervezői (hezitáló) funkcióban (5) áll leggyakrabban, akárcsak az óvodások nyelvhasználatában. A *hát* harmadik leggyakoribb előfordulása a felnőttek spontán beszédében a magyarázó, pontosító funkció volt (6). Mindezek mellett a felnőtt beszélők (vö. Markó–Dér 2011b) használták még a *hát* elemet mondanivaló továbbfűzésére (7) és konklúziószerű szerepben is (8). A *hát* diskurzusjelölő az érzelmi többlettartalom kifejezésére is szolgálhat, ahogy ismétlése által a retorikusság is erősödhet (9) (Schirm 2009: 308–309). Előfordul még kérdésbevezető szerepben is a *hát* diskurzusjelölő (10) (Kondacs 2017).

- [1] F (adatrögzítő, felnőtt): *és szerinted mit csinálhatott a nyúl meg a róka?*
Gy (adatközlő, gyermek): *Hát elmentek.* (Kondacs 2017ⁱ)
- [2] F: *és mit csinált utána a tündér?*
Gy: *háát (0.2) táncolt* (Kondacs 2017)
- [3] *az is jó meg így tornázunk így táncolok úgyhogy futkározok edzek [szünet] hát ennyi* (Markó–Dér 2011b: 8)
- [4] Interjúkészítő: *és mi az ami különös vagy mi tetszik benne?*
Adatközlő: *hát szerintem a gyerekekkel való fog [foglalkozás] tehát hogy gyerekekkel foglalkozhatom* (Markó–Dér 2011a: 58).
- [5] *pont azt kell használni ami hat [szünet] ja és hát fee hát ez ez ez hát ma nincs más ni ne nem nem lehet bocsánat hogy bele belezörejedsz öö hatni kell* (Markó–Dér 2011a: 59).
- [6] *de hogy ezeket mind meg kellett élni ahhoz hogy most [szünet] a mostani [szünet] gondolkodásom tát [tehát] jó hát ez nem egy olyan nagy út* (Markó–Dér 2011a: 9).
- [7] *és akkor megkapta ez a kislány hát ő volt a leg [szünet] rendesebben aki hordta a készülékeket* (Markó–Dér 2011b: 9).
- [8] *és azt gondolom hogy a diákok egy jó része szeretheti mert mert hogy visszajönnek öö szakszemináriumokra hát ez az én egyetemre kerülésnek a története* (Markó–Dér 2011b: 9).
- [9] *Mi éppen azt mondjuk – most ebben a felszólalásban is, a három ellenzéki párt álláspontját –, hogy mi szükségesnek tartjuk a probléma megoldását, de ez nem a probléma megoldása. Hát hol van az alapító okirat tartalma? Hát azt kell megmondani! Hát ahhoz milyen szándékok? És tessék monda-*

ⁱ A példákban a konverzációelemzési átírási konvenciókat használtam fel. (Jefferson 1984).

ni, 28 százalék meg tudja akadályozni egy szándék megvalósítását? (Schirm 2009: 308–309)

- [10] F: *és a kutya is HÁT szerinted ki mehetett még el?*
Gy: *Nem tudom.*”, illetve új téma bevezetőjeként is használhatjuk
F: *NAGYON ügyes voltál (.) hát már csak egy mese van hátra. Arra is figyelj ilyen ügyesen*

A diskurzusjelölők prozódiai jellemzőit eddig a hazai szakirodalomban csak néhányan vizsgálták (pl. Abuczki 2014, Dér–Markó 2010, Dér–Markó 2017). A *prozódia* a „beszédprodukciós folyamat által létrehozott komplex beszédjelnek az a vetülete, amely az idő, a frekvencia és az intenzitás folyamatváltozásaiként írható le, és amelynek az észlelése kizárólag állandó viszonyításban, nagyobb egységeken (minimálisan szótagok viszonylatában) lehetséges” (Markó 2015: 18). A kezdeti prozódiai vizsgálatokban a határozók prozódiai jellemzőit vizsgálták meg (Meyer 1986, Altenber 1987), majd ezután kezdtek el az angol diskurzusjelölők jellemzőinek leírásával foglalkozni. A korábbi szakirodalom megállapította, hogy a vizsgált elemek jellemzően elkülönülnek a mondat többi részétől, prozódiaileg függetlenek, amit az egységet megelőző és követő szünet jelez (Zwicky 1985: 303–304). Ezt az állítást azonban mind a hazai (Dér–Markó 2010), mind a nemzetközi kutatások (Hirschberg–Litman 1993: 516) cáfolták. A hazai és a nemzetközi prozódiai vizsgálatok általánosabb és specifikusabb jellemzőket fogalmaztak meg. Kawamori–Shimazu–Kawabata (1996) a diskurzusjelölők két funkciójának, a hezitálónak és a válaszadó szerepkörnek figyelték meg a prozódiai jellemzőit. Megállapították, hogy a diskurzusjelölők intonációs vonásai kategóriafüggek. Tseng és mtársai (2006) a diskurzusjelölők és a beszéd-töltelékek elkülönítésével foglalkoztak a Prosodic Phrase Group modelljének segítségével, és megállapítják, hogy a két fogalom prozódiaileg elkülöníthető. Wichmann és mtársai (2010) az *of course* elemmel kapcsolatban megfigyelték, hogy nincsen egyetlen olyan prozódiai vonás, amely a diskurzusjelölőkre volna jellemző. Beňuš (2012) a szlovák *no* jelölővel foglalkozott, és megállapította, hogy csak a visszajelző és a beszélőt folytatásra buzdító funkció mutat a többi szerepkörrel összevetve kiemelkedő prozódiai jegyeket. Azonban a felnőtt nyelvhasználatban megjelenő *hát* diskurzusjelölő pragmatikai szempontból teljesen elkülönülő funkcióinak azonosítását Dér–Markó (2017) végezte el, amelyhez a prozódiai konkomitanciákat keresték meg. A megvizsgált prozódiai jegyek: a szünet, az időtartam és az alaphfrekvencia. Az alaphfrekvenciát Praat programmal mértem ki. A szünet (Gósy 2000: 2) „beszéd folyamatban jelentkező olyan kismértékben akaratlagos beszédkimaradás, amely néma vagy jellel kitöltött, de független a beszédhang képzésétől”. Kowal és mtársai (1983) a szünetek minimum időtartamát 200 és 300 ms közé teszi, melyek a beszéd közben megjelenő kognitív hiányt mutatják meg.

Kutatásomban nemcsak a diskurzusjelölők, hanem a diskurzusjelölőt tartalmazó forduló prozódiai jellemzőivel is foglalkozom. A konverzációelemzői (Iványi 2001, Németh 2014) megközelítése szerint a beszélgetés egy interakcióba ágyazott, strukturálisan szervezett folyamat. A beszélgetés legkisebb egysége a forduló, amelyet az alkot, amit „egy társalgásban részt vevő beszélő mond addig, amíg vagy átadja a szót a következő beszélőnek, vagy befejezi mondanivalóját szóátadás nélkül” (Sacks et al. 1974, Németh 2014: 30). A társalgáselemzési nézőpont alapján a cselekvés egy olyan „fő munka, amelyet a forduló végez” (Levinson 2013: 107, Németh 2014: 30). Minden forduló legalább egy cselekvést kell, hogy tartalmazzon (Schegloff 2007).

A jelen kutatásban arra a kérdésre keresem a választ, hogy a *hát* prozódiai megvalósítása eltér-e attól függően, hogy az óvodáskorúak a félintézményes (Schirm 2017) szövegekben milyen funkciókban használják őket. Feltételeztem, hogy a *hát* időtartama, a *hát*-ot követő vagy nem követő szünet és az alapfrekvencia minimuma és maximuma támogatják a két leggyakoribb funkciónak, a hezitálónak és az általános válaszelőjelnek az elkülönítését. Továbbá feltételeztem azt is, hogy bizonyos nominális változók (nem, korcsoport, helyszín) szintén befolyásolják a prozódiai jegyek súlyát a két funkció kapcsán.

2 Anyag, kísérleti személyek, módszer

2.1 Anyag és kísérleti személyek

A kutatáshoz a saját gyűjtésű korpuszomat használtam fel, amely 167 felvételt tartalmaz, 167 különböző beszélővel. Az összes felvétel teljes időtartama 18 óra 26 perc, a felvételek átlagos hossza: 6 perc 58 másodperc. Az adatközlők egy magyarországi falusi intézményből, a szegvári Kurca-parti Óvodából; egy magyarországi városi óvoda két intézményéből, a Mindszenti Károly és Móricz Óvodából; illetve egy határon túli városi óvoda, a palicsi Kalimero Óvoda három kiscsoportjából, kettő középső és négy nagycsoportjából valók. A magyarországi falusi óvodából 80 fővel, a városiból 55-tel, míg a határon túli városiból 34 fővel rögzítettem anyagot. A felvételek rögzítése előtt a szülőkkel hozzájárulást íratattam alá. Az anyaggyűjtés során célom volt, hogy adataim diverzek legyenek, vagyis különféle településtípusról valók legyenek, és ne ugyanazt a nyelvjárást beszéljék mindannyian. A gyerekek átlagéletkora 5 év 8 hónap. A nemi megoszlásuk 85 lány, 82 fiú. A *hát* diskurzusjelölőt produkáló nagycsoportos korosztály életkora 5,6–6,11 év közötti volt, a középsősöké 4,3 és 5,5 év közötti, míg a kiscsoportosok 2,9 és 4,2 év közöttiek voltak. A hangfelvételeket egyesével vettem föl egy külön helyiségben, mely nem volt zajszigetelt.

A felvételeket diktafonnal és lappal rögzítettem, így hang- és videófelvétel készült. A hanganyagok diktafonnal kerültek rögzítésre .wav formátumban. A felvevő típusa: Olympus VN-732PC. A mérhető frekvenciatartomány: 70Hz–19kHz. A videófelvételt Samsung lappal telepített Debut Video Capture Software segítségével rögzítettem. A gyerekeket délelőttönként, mindig ugyanabból az óvoda csoportból vittem ki az anyaggyűjtésre. A gyermekek beszédindulását az óvónők és az óvónők által megkérdezett szülők átlagban 2–2,5 éves korra tették. Az anyanyelv-elsajátítására jellemző volt, hogy a lányok hamarabb kezdtek el szavakat használni minden csoport esetén (16 hónapos kor után). A gyermekek általánosságban ép hallásúak, hallásérült 4 fő, akik nem produkálták a vizsgált *hát* elemet. Többségük egynyelvű, viszont a szerbiai magyar gyerekeknél erős kétnyelvű környezet van. Az óvónők szerint 12 gyerek lehet, aki szigorúan véve kétnyelvű a szülők nemzetiségét is figyelembe véve, de minden kisgyerekre jellemző, hogy több szót is értenek mindkét nyelven. A magyarországi városi intézményekben a roma gyerekek esetén is szintén fennáll a kétnyelvűség lehetősége, de az óvónők tisztán kétnyelvűnek nem írtak be senkit. A szegvári óvodában egy kétnyelvű kisgyerek van. A szociális háttérük kapcsán elmondható, hogy a gyerekek többsége megfelelő anyagi körülmények között él. A 167 adatközlőből 20 fő nehéz szociális helyzetben él.

Az alkalmazott módszer történetmesélés volt képek és rövidebb-hosszabb (10–60 s) bábmések segítségével, melyek rögzítésében egy Mészáros Vincéné díjas bábművész segített. A gyerekek a mese megnézése előtt azt az instrukciót kapták, hogy figyeljenek arra, amit mutatni fogok nekik, és a mese megnézése után a következő kérdéseket hallották: *Mit láttál? Mi történt a mesében? A mesék végén, azokból kiindulva további beszédtemák is felmerültek: Mi a kedvenc állatod? Mit szoktál otthon játszani? Milyen évszak van most?*

A meséket különféle bábozási technikákkal rögzítettük (kesztyűbáb, óriásbáb, síkbáb). A gyerekek kevés szereplős, állatos meséket láttak, melyben az auditív és a vizuális ingerek együttesen jelentek meg, hiszen Zsák (2016) eredményei szerint a gyerekek mesemegértése a két csatorna megjelenésével javul. A kiscsoportosok 3–4 mesét, a középsősök 5–6 mesét, míg a nagycsoportosok 9-et láttak. A hát összesen 69 adatközlőnél jelent meg. Ahhoz azonban, hogy statisztikailag összevethető legyenek az adatok, egy adatközlőtől nemcsak egyet használtam fel, így a két funkció megvizsgálására összesen 126 adatot vettem alapul: 63 általános válaszjelölői és 63 hezitáló szerepkörben álló hát-ot. Törekedtem arra is, hogy minden településtípus minden korcsoportjából mindkét funkció esetén meglegyen a 9 adat, mivel a lehető legtöbb mintát szerettem volna felhasználni a csoportokból, azonban a Mindszenti középső csoport gyermekei a hát jelölő általános válaszjelölői szerepet betöltő, míg a palicsi óvoda kiscsoportjának tagjai a hezitáló funkciójú jelölőket nem produkáltak vizsgálható mennyiségű számban. A funkciók megállapításakor a kontextuális jellemzőket vettem alapul, vagyis megfigyeltem, hogy az általános válaszjelölő szerepkör során a feltett kérdés után a beszélő választ ezzel az elemmel vezette be (11). A hezitáló funkció megállapításakor pedig főként szókeresés céljából használta az elemet, vagyis akkor, amikor a beszélőnek még időre volt szüksége a beszéde rendezéséhez (12).

[11] F: *És ő veled mit szoktatok játszani? Vagy ő még nagyon kicsi?*

Gy: *Hát olyan 4 hónapos* (Kondacs 2017).

[12] F: *jól van. és TI mit szoktatok itt az oviban játszani?*

Gy: *az oviba? háát (.) játékokkal (.) bábjátékokkal* (Kondacs 2017).

Emellett más funkciókra is találtam példát, de azok a kis darabszám miatt összevethetetlenek. Ilyen szerepek voltak például az evidencialitást jelölő funkció (12) és a mondanivaló továbbvitele (13) (Kondacs 2017).

[13] F: *ÉS szerinted meg is ehette az egeret?*

Gy: *ah, HÁT persze* (Kondacs 2017).

[14] Gy: *Apa anyja meg én illetve ők itt voltak (.) de én nem itt voltam (.) de ÉN is mentem már mentőautóba. De ő unokatesóm nagyon szeret (.) Zolcsi biztos már voltál nála (.) UGYE még közép nagycsoportba (.) hogy nagyon szereti a mentőket de csak akkor (.) ha nem őt viszi el.*

F: *<Hát az>.*

Gy: *<Hát mondom> Zolcsikám NEKED az nagyon jó (.) de mondom a mentőbe nem annyira jó feküdni* (Kondacs 2017).

2.2 Módszer

A hát diskurzusjelölő prozódiai jellemzőit a Praat 6.0.31 szoftver segítségével előfordulásonként nyertem ki a hanganyagokból. Mivel a felvételek nem stúdiókörülmé-

nyek között készültek, így a *hát* nem minden előfordulását vettem bele az adatokba, a zajos hanganyagokat még a kutatás elején kiszűrtem. Zajos felvételnek azok számítottak, ahol a spektrogram segítségével nem tudtam kijelölni a diskurzuszjelölő /h/-tól /a/-ig, illetve /t/-ig tartó szakaszát. Azokat a hanganyagokat, amelyekben csak minimális alapzaj volt érzékelhető, feldolgoztam, hiszen ez az alapzaj egységesen megtalálható volt minden felvételen. Ennek ellenére az eredményeim értékelése során figyelembe kell venni ezt a tényt is.

A kutatáshoz a következő prozódiai információkat használtam fel: a diskurzuszjelölő időtartama (másodpercben), a diskurzuszjelölőt követő szünet léte (nincs, van) és a szünet időtartama, a diskurzuszjelölő alaphangfrekvenciájának (f_0) minimum- és maximum-értéke (Hz), a *hát*-ot is tartalmazó forduló alaphangfrekvenciájának minimuma és maximuma (Hz).

A diskurzuszjelölőket és a fordulókat az Audacity programmal vágtam ki a felvételekből. A diskurzuszjelölőket minden esetben a Praat 6.0.31 szoftver segítségével mértem ki, a szegmentális határokat figyelembe véve, a spektrogram segítségével kijelöltem a /h/ elejétől a /t/ végéig. Két példában csak a /ha/ rövidített alak jelent meg. Mindkét eset általános válaszzjelölő funkcióban jelent meg. A forduló (Sacks et al. 1974) megállapítása során mindig legalább egy olyan cselekvést kerestem ki, mely tartalmazza a *hát* diskurzuszjelölőt. Például: *Hát elemek* (Kondacs 2017). Spektrogram segítségével mértem ki a fordulókat is.

Ahhoz, hogy statisztikai vizsgálatokat is végezhessenek az adataimon, tehát megfelelő számú mintám legyen, a korpuszomban megjelenő összes *hát* előfordulást megvizsgáltam. Elsőként megfigyeltem, hogy az általam gyűjtött adatok átlagai és szórásai mutatnak-e különbséget nemek (lány-fiú), életkor (kiscsoport-középső csoport-nagycsoport) és elhelyezkedés (belföld-külföld) tekintetében. A helyszín tekintetében a magyarországi és a Magyarországon kívüli csoportra tudtam az adataimat bontani, részletesebb csoportosítás esetén nem állna rendelkezésemre elegendő adat a statisztikai elemzéshez.

Ezután a statisztikai elemzést az SPSS Statistics 24.0 szoftver segítségével végeztem el, az adatok közzétevése során pedig a University of Washington által kiadott elveket követtem (Writing an APA Empirical (lab) Report, 2010). A statisztikai vizsgálat előtt normalitásvizsgálatot végeztem a Shapiro–Wilk-próbával, melynek eredménye alapján az adataim nem mutatnak normális eloszlást. A hipotézisem ellenőrzése érdekében az adataimat Mann–Whitney-próbával hasonlítottam össze. 63 általános válaszzjelölő és 63 hezitáló funkcióban előforduló *hát* adatait vizsgáltam meg.

3 Eredmények

Az általam feldolgozott adatok általános információit az 1–4. táblázatok mutatják.

A nemek megoszlása alapján a *hát* diskurzuszjelölőt hezitáló funkcióban 22 alkalommal használták a fiúk és 41 alkalommal a lányok, általános válaszzjelölő funkcióban pedig 23 alkalommal fordult elő a fiúknál és 40 alkalommal a lányoknál. Az 1. táblázat alapján elmondható, hogy a nemek tekintetében az egyes prozódiai jellemzők átlagában nem mutatkozik számottevő eltérés az óvodáskorúak spontán beszédében előforduló *hát* két funkciója kapcsán. A két nem tekintetében az általános válaszzjelölői *hát*-ot tartalmazó forduló esetében a lányok eredménye 4,8%-kal magasabb értéket mutatott a fiúk eredményéhez képest, illetve a hezitáló *hát* esetében 5,4% volt a különbség a lányok javára a két nem között, illetve a hezitáló *hát*-ot tartalmazó forduló

alapfrekvenciájának értéke a lányoknál 4,6%-kal magasabb értéket mutatott a fiúk eredményéhez viszonyítva.

| | Válaszjelölő | | Hezitáló | |
|--|----------------|----------------|---------------|----------------|
| | Fiú | Lány | Fiú | Lány |
| <i>hát</i> időtartama (s) | 0,291±0,127 | 0,335±0,118 | 0,127±0,156 | 0,118±0,265 |
| forduló időtartama (s) | 2,991±1,719 | 3,109±1,490 | 1,719±1,903 | 1,490±2,402 |
| <i>hát</i> f ₀ -min. (Hz) | 269,103±55,214 | 263,557±80,190 | 55,214±97,934 | 80,190±109,369 |
| <i>hát</i> f ₀ -max. (Hz) | 331,334±64,545 | 321,013±61,545 | 64,545±61,678 | 61,545±72,059 |
| forduló f ₀ -min(Hz) | 127,339±61,050 | 143,374±58,142 | 61,050±70,945 | 58,142±52,110 |
| forduló f ₀ -max (Hz) | 41,401±71,202 | 373,797±69,923 | 71,202±59,875 | 69,923±86,311 |
| <i>hát</i> f ₀ -átlaga (Hz) | 294,616±36,249 | 294,456±58,702 | 36,249±56,037 | 58,702±71,969 |
| forduló f ₀ -átlaga (Hz) | 261,461±25,513 | 274,859±40,336 | 25,513±39,763 | 40,336±44,561 |

1. táblázat. A hezitáló és az általános válaszjelölő szerepek időtartamának és alapfrekvenciájának átlag- és szórásértékei a nemek tekintetében.

| | Korcsoport | | |
|--|----------------|----------------|----------------|
| | Kicsi | Középső | Nagy |
| <i>hát</i> időtartama (s) | 0,354±0,100 | 0,303±0,121 | 0,306±0,135 |
| forduló időtartama (s) | 3,076±1,273 | 3,363±1,411 | 2,861±1,837 |
| <i>hát</i> f ₀ -min. (Hz) | 275,913±98,764 | 261,147±62,675 | 261,651±56,505 |
| <i>hát</i> f ₀ -max. (Hz) | 326,010±79,004 | 339,839±65,511 | 313,923±45,816 |
| forduló f ₀ -min(Hz) | 145,723±57,905 | 114,391±46,683 | 147,471±65,052 |
| forduló f ₀ -max (Hz) | 395,389±72,828 | 405,806±72,036 | 369,244±69,960 |
| <i>hát</i> f ₀ -átlaga (Hz) | 299,203±82,494 | 296,257±30,117 | 290,226±34,665 |
| forduló f ₀ -átlaga (Hz) | 289,675±43,205 | 260,303±27,013 | 263,273±31,848 |

2. táblázat. A vizsgált adatok átlag- és szórásértékei az általános válaszjelölő funkció kapcsán az egyes korcsoportok tekintetében.

| | Korcsoport | | |
|--|-----------------|----------------|-----------------|
| | Kicsi | Középső | Nagy |
| <i>hát</i> időtartama (s) | 0,693±0,222 | 0,663±0,231 | 0,686±0,282 |
| forduló időtartama (s) | 3,759±2,452 | 5,016±2,267 | 4,070±2,043 |
| <i>hát</i> f ₀ -min. (Hz) | 253,972±120,833 | 235,231±94,213 | 226,161±102,353 |
| <i>hát</i> f ₀ -max. (Hz) | 396,455±86,349 | 339,461±48,821 | 357,741±62,424 |
| forduló f ₀ -min(Hz) | 138,629±67,497 | 146,782±69,306 | 118,986±52,599 |
| forduló f ₀ -max (Hz) | 417,839±109,145 | 412,997±61,277 | 406,076±63,944 |
| <i>hát</i> f ₀ -átlaga (Hz) | 315,579±90,115 | 305,464±50,176 | 292,398±58,890 |
| forduló f ₀ -átlaga (Hz) | 295,428±54,593 | 283,378±37,304 | 265,975±34,478 |

3. táblázat. A vizsgált adatok átlag- és szórásértékei a hezitáló funkció kapcsán az egyes korcsoportok tekintetében.

Az általános válaszjelölő funkcióban előforduló *hát* 18-18 alkalommal fordult elő a kiscsoportos (2. és 3. táblázatban *kis*) és középsőcsoportos (2. és 3. táblázatban *középső*) óvodásoknál, 27 alkalommal pedig a nagycsoportosok (2. és 3. táblázatban *nagy*) nyelvhasználatában. A *hát* hezitáló funkcióban szintén 18-18 alkalommal jelent meg a

kiscsoportos és középső csoportos, míg 27 alkalommal a nagycsoportosok hanganyagában. Az általános válaszelölői funkciójú *hát* esetén az alapfrekvencia átlagértéke és a *hát* időtartamának átlaga is csökken az életkor előrehaladtával. A további adatok kapcsán nem találtam az egyes korcsoportok tekintetében nagy eltérést az átlagokban.

| | Belföld | Külföld |
|--|--------------------|-----------------|
| | Válaszelölő | |
| <i>hát</i> időtartama (s) | 0,330±0,116 | 0,293±0,136 |
| forduló időtartama (s) | 2,896±1,299 | 3,491±2,072 |
| <i>hát</i> f ₀ -min. (Hz) | 269,013±77,753 | 257,003±54,477 |
| <i>hát</i> f ₀ -max. (Hz) | 322,803±58,037 | 329,726±73,611 |
| forduló f ₀ -min (Hz) | 149,727±61,803 | 107,003±39,152 |
| forduló f ₀ -max (Hz) | 372,337±65,015 | 424,218±77,150 |
| <i>hát</i> f ₀ -átlaga (Hz) | 296,979±56,735 | 288,351±34,905 |
| forduló f ₀ -átlaga (Hz) | 274,537±38,798 | 258,545±25,290 |
| | Hezitáló | |
| <i>hát</i> időtartama (s) | 0,711±0,253 | 0,503±0,110 |
| forduló időtartama (s) | 4,254±2,255 | 4,236±2,352 |
| <i>hát</i> f ₀ -min. (Hz) | 239,399±104,128 | 220,497±113,177 |
| <i>hát</i> f ₀ -max. (Hz) | 364,496±71,828 | 358,083±56,657 |
| forduló f ₀ -min (Hz) | 136,302±63,598 | 109,966±50,745 |
| forduló f ₀ -max (Hz) | 412,537±79,879 | 404,680±66,686 |
| <i>hát</i> f ₀ -átlaga (Hz) | 305,713±68,045 | 285,003±59,475 |
| forduló f ₀ -átlaga (Hz) | 281,129±45,409 | 268,763±423,823 |

4. táblázat. A hezitáló és az általános válaszelölő szerepek átlag- és szórásértékei a helyszín tekintetében.

A *hát* általános válaszelölő funkcióban 45 alkalommal volt fellelhető a belföldi óvodások beszédében és 18 alkalommal a határon túli óvodásokéban. A hezitáló funkcióban 54 alkalommal jelent meg, a magyarországi óvodásoknál és mindössze 9 alkalommal volt adathozható a határon túli óvodában. Az adatok alapján elmondható, hogy a külföldön élő óvodások mind az általános válaszelölői, mind a hezitáló funkcióban használt *hát*-ot rövidebb időtartammal ejtik ki. Emellett a *hát* alapfrekvencia-minimuma és -maximuma, illetve a forduló alapfrekvencia-maximuma esetén is alacsonyabb a külföldön élő óvodások frekvenciájának átlaga.

A *hát*-ot követő szünetekkel kapcsolatban elmondható, hogy összesen 38 esetben jelent meg néma szünet, míg 13-szor kitöltött szünet (*öö*, *mm*) követte a *hát* elemet. A néma szünetek átlagos időtartama 1,33 másodperc, míg a kitöltött szünet időtartama átlagosan 0,32 másodperc volt.

Az adatok átlagainak és szórásainak vizsgálata során minden esetben 95%-os konfidenciaszinttel dolgoztam. A statisztikai elemzés kapcsán a Mann–Whitney-próba elvégzése után elmondható, hogy hezitáló és általános válaszelölő funkciók között szignifikáns a különbség a *hát* időtartama ($U = 250$, $Z = -8,465$, $p < 0,01$), a forduló időtartama ($U = 1369,5$, $Z = -3,001$, $p < 0,01$), a *hát* alapfrekvenciájának maximuma ($U = 1350$, $Z = -0,671$, $p < 0,01$), és a forduló alapfrekvenciájának maximuma ($U = 1559$, $Z = -2,076$, $p = 0,038$) alapján.

A statisztikai vizsgálatot elvégeztem a nemekre lebontva is. A fiúk esetében 45 alkalommal fordult elő a *hát*, ebből 23 alkalommal általános válaszelölő 22 alkalom-

mal pedig hezitáló funkcióban. A teszt alapján szignifikáns eltérést a fiúk esetében egyedül a *hát* időtartama mutatott a két funkció között ($U = 48,5$, $Z = -4,646$, $p < 0,01$). A lányok esetében 81 előfordulása volt a *hát* diskurzusjelölőnek, ebből 40 alkalommal általános válaszelölő, 41 alkalommal hezitáló funkcióban.

A Mann–Whitney-próba alapján a lányok esetében szignifikáns a különbség a két funkció között a *hát* időtartama ($U = 68,5$, $Z = -7,101$, $p < 0,01$), a forduló időtartama ($U = 540,5$, $Z = 1,360$, $p < 0,01$), a *hát* frekvenciájának maximuma ($U = 498$, $Z = -3,042$, $p < 0,01$), a *hát*-ot tartalmazó forduló frekvenciájának minimuma ($U = 598$, $Z = -2,097$, $p = 0,036$) és a *hát*-ot tartalmazó forduló frekvenciájának maximuma ($U = 531$, $Z = -2,73$, $p < 0,01$) tekintetében.

Szintén megvizsgáltam az adataimat korcsoportokra lebontva is. A kiscsoportos óvodások összesen 18-18 alkalommal használták az általános válaszelölő és a hezitáló funkciójában a *hát* diskurzusjelölőt. A teszt alapján szignifikáns különbség van a két funkció között a *hát* időtartamának tekintetében ($U = 14$, $Z = -4,684$, $p < 0,01$). A középső csoportosok esetében 18 alkalommal az általános válaszelölő, míg 28 esetben a hezitáló szerepkör fordult elő. Szignifikáns eltérést a *hát* időtartama ($U = 14$, $Z = -4,685$, $p < 0,01$) és a forduló időtartama ($U = 87$, $Z = -2,373$, $p = 0,18$) mutatott. A nagycsoportosok 27-27 esetben használták a *hát* két funkcióját. Szignifikáns eltérést a *hát* időtartama ($U = 60,5$, $Z = -438,5$, $p < 0,01$) a forduló időtartama ($U = 220,5$, $Z = -2,491$, $p = 0,013$), a *hát* frekvenciájának maximuma ($U = 216$, $Z = -2,569$, $p = 0,01$) és a forduló frekvenciájának maximuma ($U = 245$, $Z = -1,825$, $p = 0,039$) mutatott.

A magyarországi óvodások és a határon túli óvodások által produkált adatok tekintetében is megvizsgáltam a két funkciót. A magyarországi óvodások összesen 99 alkalommal produkálták a *hát* diskurzusjelölőt, 45 esetben használták általános válaszelölő funkcióban, 54-szer pedig hezitáló szerepben. A statisztikai vizsgálat után elmondható, hogy a *hát* időtartama ($U = 122$, $Z = -7,683$, $p < 0,01$), a forduló időtartama ($U = 802$, $Z = -2,899$, $p < 0,01$), a *hát* frekvenciájának maximuma ($U = 836$, $Z = -2,663$, $p < 0,01$) és a forduló frekvenciájának maximuma ($U = 788$, $Z = -3,001$, $p < 0,01$) esetében van szignifikáns különbség a két funkció között. A határon túli óvodások 27 alkalommal használták a *hát* diskurzusjelölőt, ebből 18 általános válaszelölő funkcióban, 9-szer pedig hezitáló szerepkörben. A teszt alapján egyedül a *hát* időtartama ($U = 21$, $Z = -3,089$, $p < 0,01$) mutat szignifikáns különbséget.

Annak tekintetében is elvégeztem a statisztikai vizsgálatot, hogy a *hát* után tartott-e szünetet az adatközlő. A 126 adat közül összesen 75 esetben nem követte szünet a diskurzusjelölőt. Ebből 58 alkalommal az általános válaszelölő funkció szerepelt, míg 17-szer a hezitáló funkció. A Mann–Whitney-próba alapján szignifikáns különbséget mutatott a *hát* időtartama ($U = 82$, $Z = -5,203$, $p < 0,01$), és a *hát* frekvenciájának maximuma ($U = 322$, $Z = -2,164$, $p = 0,03$). Összesen 51 esetben követte szünet a *hát* diskurzusjelölőt, ebből 5 általános válaszelölő funkció után fordult elő, 46 pedig hezitáló funkció után jelent meg. A teszt alapján szignifikáns különbség egyedül a *hát* időtartama ($U = 19,5$, $Z = -3,026$, $p < 0,01$) tekintetében van a két csoport között.

4 Következtetések

Dolgozatomban a *hát* diskurzusjelölőnek a korpuszomban megjelenő két leggyakoribb, az általános válaszelölő és a hezitáló funkcióját alapul véve megállapítottam,

hogy a prozódia tekintetében vannak olyan különbségek, amelyek segíthetnek a két funkció elkülönítésében.

A nominális változókat (nemet, korcsoportot, helyszínt) is figyelembe véve a prozódiai jegyek közül a *hát* időtartama az, amelyik minden változó esetében szignifikáns eltérést mutat. Kiemelném még a forduló időtartamát, a *hát* frekvencia maximumának és a forduló frekvencia maximumának értékeit, melyek szintén szignifikáns eltéréseket mutatnak bizonyos nominális változó vizsgálatokor. Azt mutatják az eredmények, hogy a diskurzusjelölők funkcióinak elkülönítéséhez nemcsak a szemantikai és pragmatikai jellemzőket, hanem a prozódiai információkat is érdemes megvizsgálni, hiszen abban is vannak különbségek.

Megfigyelhető volt, hogy a vizsgált *hát*-okat az életkor előrehaladtával egyre több óvodás használja. Az látszik, hogy a kiscsoportosok közül 20 beszélő használta a 48-ból (42%), a középsősöktől 17 beszélő az 52-ből (33%), míg a nagycsoportos korosztályból összesen 32 óvodás nyelvhasználatában jelent meg a *hát* elem a 69 beszélőből (46%).

A felnőtt nyelvhasználókkal végzett vizsgálathoz (Dér–Markó 2017) hasonló eredmények születtek az óvodás korosztály tekintetében is, vagyis a 67 db hezitáló funkcióban megjelenő *hát* után sokkal jellemzőbbek voltak a kitöltött szünetek (13 eset) és néma szünetek (33 db), mint az általános válaszelölő funkció esetén, ahol mindösszesen 5 db néma szünet volt.

További kutatási célom az, hogy a funkciók adatszámainak bővítésével a gyermeknyelvi diskurzusjelölők jellemzőiről olyan általánosítást mutassak meg, amellyel segítem az anyanyelv-elsajátítás folyamatának pontosabb megismerését.

Köszönetnyilvánítás



A jelen vizsgálat az Emberi Erőforrások Minisztériuma ÚNKP-17-3 kódszámú új Nemzeti Kiválóság Programjának támogatásával készült.

Köszönettel tartozom doktoranduszársamnak, Kovács Viktóriának, amiért segítséget nyújtott a statisztikai elemzésben.

Irodalom

- Abuczki, Á. 2014. *A core/periphery approach to the functional spectrum of discourse markers in multimodal context – A corpus-based analysis of mondjuk (~'say'), ugye (~'is that so?') and amúgy (~'otherwise')*, doktori disszertáció.
Elérhető:
[https://dea.lib.unideb.hu/dea/bitstream/handle/2437/210101/disszertacio_abuczki_agnes_t.pdf?sequence=13&isAllowed=y]
- Altenberg, B. 1987. Predicting text segmentation into tone units. In: Willem Meijs (szerk.) 1987. *Corpus Linguistics and Beyond*. Amsterdam: Rodopi. 49–60.
- Bell, D. M. 1998. Cancellative discourse markers: a core/periphery approach. *Pragmatics*, 8:4: 515–541.

- Beňuš, Š. 2012. Prosodic forms and pragmatic meanings: The case of the discourse marker 'no' in Slovak. In: *Proceedings of CogInfoCom 2012*. 77–82.
Elérhető: [http://www1.cs.columbia.edu/~sbenus/Research/Benus_CogInfoCom_no-words_final_check.pdf]
- Dér, Cs. I. 2005. A diskurzusjelölők kialakulásáról a grammatikalizációs paraméterek tükrében. In: Mártonfi, A., Papp, K., Slíz, M. (szerk.) 2006. *101 írás Pusztai Ferenc tiszteletére*. Budapest: Argumentum. 83–88.
- Dér, Cs. I., Markó, A. 2007. A magyar diskurzusjelölők szupraszegmentális jelöltségéről. In: Gecső, T., Sárdi, Cs. (szerk.) *Nyelvelmélet–nyelvhasználat*, Székesfehérvár–Budapest: Kodolányi János Főiskola–Tinta Könyvkiadó. 61–67.
- Dér, Cs. I., Markó, A. 2010. Diskurzusjelölők használata az életkor és a nem függvényében. In: Gecső, T., Sárdi Cs. (szerk.) 2010. *Új módszerek az alkalmazott nyelvészeti kutatásban*. Székesfehérvár–Budapest: Kodolányi János Főiskola–Tinta Könyvkiadó. 78–83.
- Dér, Cs. I., Markó, A. 2017. A hát funkciói a prozódiai megvalósulás függvényében. *Beszédkutatás*, 2017: 105–116.
Elérhető: [<http://ojs3.mtak.hu/index.php/beszkut/article/view/385/197>]
- Fraser, B. 1999. What are discourse markers? *Journal of Pragmatics*, 31: 931–952.
- Gósy, M. 2000. A beszédszünetek kettős funkciója. *Beszédkutatás* 2000. 1–14.
- Hirschberg, J., Litman, D. 1993. Empirical Studies on the Disambiguation of Cue. *Computational Linguistics*, 19/3: 501–530.
- Iványi, Zs. 2001. A nyelvészeti konverzációelemzés. *Magyar Nyelvőr* 125: 74–93.
- Jefferson, G. 1984. On the organization of laughter in talk about troubles. In: Atkinson, J. M., Heritage J. (szerk.) 1984. *Structures of Social Action. Studies in Conversation Analysis*. Cambridge: Cambridge University Press. 346–369.
- Kawamori, M., Shimazu, A., Kawabata, T. 1996. A Phonological Study on Japanese Discourse Markers. In: *Proceedings of the 11th Pacific Asia Conference of Language, Information and Computation*. 297–306. Elérhető: [<http://www.aclweb.org/anthology/Y96-1031>]
- Levinson, S. C. 2013. Action formation and ascription. In Sidnell, J., Stivers, T. (szerk.) *The Handbook of Conversation Analysis*. Oxford: Wiley-Blackwell. 101–130.
- Markó, A., Dér, Cs. I. 2011a. Diskurzusjelölők használatának életkori sajátosságai. In: Navracscics, J., Lengyel, Zs. (szerk.) 2011. *Lexikai folyamatok egy- és kétnyelvű közegben*. Pszicholingvisztikai tanulmányok, II. Budapest: Tinta. 49–61.
- Markó, A., Dér, Cs. I. 2011b. Diskurzusjelölők használatának életkori sajátosságai. Elérhető: http://real.mtak.hu/26155/1/Marko-Der_pszicholingvisztika_2011.pdf
- Markó A. 2015. A spontán beszéd prozódiai szerkezete. Időzítés és beszéddallam. Nyelvtudományi Értekezések 166. Budapest: Akadémiai.
- Meyer, C-F. 1986. Punctuation practice in the Brown Corpus. *ICAME news*, 10: 80–95.
- Németh, Zs. 2014. Javítási műveletek jelöltségi hipotézise. *Jelentés és Nyelvhasználat*, I. 29–54. Elérhető: [<http://www.jeny.szte.hu/jeny/images/issues/2014/JENY-2014-NemethZs.pdf>]
- Sacks, H., Schegloff E. A., Jefferson G. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language*. 50/4: 696–735.
- Schegloff, E. A. 2007. *Sequence Organization in Interaction: A Primer in Conversation Analysis*. Cambridge: Cambridge University Press
- Schirm, A. 2007–8. A hát diskurzusjelölő története. *Nyelvtudomány*. III–IV., 185–201.
- Schirm, A. 2009. Partikula és/vagy diskurzusjelölő? In: Keszler, B., Tátrai, Sz. (szerk.) *Diskurzus a grammatikában - grammatika a diskurzusban*, Budapest: Tinta. 304–311.
- Schirm, A. 2011a. A diskurzusjelölők funkciói a számok tükrében. *Alkalmazott Nyelvészeti Közlemények*, VI/1, 185–197.
- Schirm, A. 2011b. *A diskurzusjelölők funkciói: a hát, az -e és a vajon elemek története és szinkron státusa alapján című doktori disszertáció*. Elérhető: [http://doktori.bibl.u-szeged.hu/759/1/schirm_anita_doktori_disszertacio.pdf]
- Schirm, A. 2017. A diskurzusjelölők és a szövegtípusok. *Magyar Nyelv*, 113 (3): 330–341.
- Statistical Tests in APA Format. 6. University of Washington, Psychology Writing Center. 2010. Reporting Results of Common.
Available: https://depts.washington.edu/psych/files/writing_center/stats.pdf.
- Tseng, C., Su, Z-j., Chang, C-H., Tai, C-H. 2006. Prosodic fillers and Discourse Markers – Discourse prosody and Text Prediction. In: *Proceedings of TAL 2006*. 109–114.
- Wichmann, A., Simon-Vandenberg, A. M., Aijmer, K. 2010. How prosody reflects semantic change: A synchronic case study of of course. In Da-vide, K., Vandelanotte, L., Cuyckens,

-
- H. (szerk.) 2010. *Subjectification, intersubjectification and grammaticalization*. Berlin–New York: De Gruyter–Mouton. 103–154.
- Zwicky, A. M. 1985. Clitics and particles. *Language*, 61: 283–305.
- Zsák, É. 2016: *A verbalitás és a vizualitás hatása gyermekek szövegértésére*. Elhangzott: Új utak a gyermeknyelvi kutatásokban. Budapest, ELTE, BTK. 2016. december 1.

Magyarországi cigány közösségek interakciós stratégiái: az iskolai nyelvi gyakorlatok és az identifikáció

Kresztyankó Annamária

ELTE BTK Nyelvtudományi Doktori Iskola
annamari915@gmail.com

Kivonat: Tanulmányom cigány gyermekek nyelvi szocializációját vizsgálja a társadalmi egyenlőtlenségek és feltételezett különbségek kialakulásának, valamint e különbségek iskolai kontextusban történő fenntartásának középpontba állításával. Kutatásom szociokulturális nyelvészeti keretbe (Bucholtz–Hall 2008) illeszkedve vizsgálja cigány közösségek komplex nyelvi gyakorlatait. A korábbi és kortárs nyelvi szocializációs kutatások között a kompenzáció, a tudatosítás és a gyakorlat elve szerint tesztek különbséget, és saját vizsgálatomat a gyakorlatelví kategóriába sorolom. Jelen dolgozat egy kulturális antropológiai terepmunka eredményeit mutatja be, melynek újszerűségét a szocializáció és a hierarchia kölcsönhatásának társas-konstruktivista és szociokulturális nyelvészeti megközelítésmódja, valamint a terepmunka módszertana, a nexuselemzés adja. Hipotézisem szerint a gyermekek interakciós stratégiái – a beszédjog, a csoporttagság-kategorizáció, a javítás, valamint ezen diszkurzív formákra érkező metanyelvi reflexiók és a proxémika – egyszerre értelmeződnek identitáskonstruáló és szankcionált nyelvi gyakorlatokként az intézményes nyelvi szocializáció során. Írásomban e kategóriák közül a csoporttagság-kategorizációra fókuszálok, de a beszédjog kiosztása és néhány metanyelvi reflexió is megjelenik az identifikáció leírása kapcsán.

1 Bevezetés

Alapvető kérdésként merülhet fel, hogy miért épp egy nyelvészeti dolgozat tárgyal társadalmi marginalizációs kérdéseket, és miért inkább etnikai, mintsem szocioökonómiai háttér alapján történt a témaválasztás. A magyarországi cigányok nyelvhasználati kérdéseit vizsgáló kutatások jellemzően a cigány – nem cigány különbségtétel nyelvi aspektusai mentén mutatják be a különböző cigány közösségek nyelvi gyakorlatait – mint a nyelveíró munkák nyelv/dialektus klasszifikációs törekvései (Erdős 1958/1989: 42–43; Heltai 2001; Kovalcsik 1993; Landauer 2009; Orsós 2012; Szalai 2007; Vekerdí 1999), a nyelvi kontaktusok hatásai (Kakuk 1997; Kresztyankó 2016; Schirm 2006), a nyelvi szocializáció (Bíró 2012; Orsós 2010; Réger 1974, 1984, 1987, Torgyik 2005), a nyelvválasztás és -vesztés okai és következményei (Bartha 2007; Bartha–Hámori 2011; Heltai–Kulcsár 2017; Kemény 2000; Réger 1974), a standardizáció nehézségei (Heltai 2015a, 2015b; Orsós 2015), valamint a beszédmódok és a nyelvi ideológiák szociokulturális beágyazottságú vizsgálatai (Bartha 2009,

Ludányi Zsófia, Krepsz Valéria, Gráczai Tekla Etelka (szerk.): *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2018*. Budapest: MTA Nyelvtudományi Intézet. 2018.

2015; Bodó 2016, Heltai 2017; Kresztyankó 2017; Szalai 2009). Munkám során a kritikai szociolingvisztika paradigmáját követve a nyelv interszjektív, diszkurzusokban alakuló jellegét emelem ki (vö. Heltai 2017). E paradigma a rendszerszemlélettől elszakadva a beszélés egy-egy módját nem egymással korrelatív eltérések halmazaként értelmezi (Eckert 2016: 69), hanem a fluiditás (Otsuji–Pennycook 2010) és a skála (Blommaert 2007) metaforáját állítja középpontba (vö. Heltai 2017: 111). A statikus metaforákat ugyanis (egyes nyelv, nyelvváltozat, egynyelvűség, kétnyelvűség, anyanyelv) az elméleti szociolingvisztika szintjén a beszélés leírására alkalmatlan, kiüresedett kategóriáknak tartják (Beck 2002; Heltai 2017: 110).

A vizsgálat során alkalmazott megközelítésmód arra a problémára világít rá, hogy az etnicitás kategóriájának tudományos leírásával épp egy olyan felosztáshoz (cigány – nem cigány beszéd, cigány–magyar beszéd) járul hozzá a kutatás, amelyről épp azért kívánok számot adni, mert diszkriminatívnak tartom.

1.1 A kutatás perspektívája

A hasonló – kulturális antropológiai és szociokulturális szemléletű – empiriára támaszkodó kutatások egyik tipikus „csapdája” lehet Kovai Cecília szerint, hogy az általuk feltárt közegek egyfajta zárt, belső logika szerint működő közösségként tűnnek fel. A másik nehézség, amelybe ezek a kutatások ütközhetnek, hogy a terepen tapasztaltak csak egy bizonyos közösség sajátosságainak látszanak (Kovai 2017: 12). Annak érdekében, hogy e „csapdákat” elkerüljem, elméleti fejezetemben ismertetem a nyelvszocializáció-kutatások irányzatait, valamint a magyarországi cigány közösségek nyelvi szocializációs vizsgálatait, így helyezve szélesebb történeti és társadalmi kontextusba kutatásomat.

2 Elméleti háttér

Kutatásom szociokulturális nyelvészeti keretbe (Bucholtz–Hall 2008) illeszkedve vizsgálja cigány közösségek komplex nyelvi gyakorlatait. A szociokulturális nyelvészeti megközelítés a következő öt pontban foglalható össze: (1) központi kérdése a nyelvi gyakorlatok és azok változatainak kutatása, (2) tárgya a kommunikáció elmélete, a nyelvi szerkezetek szociális (a gyakorlatközösség tagjai közti viszonyokat kifejező) jelentéseinek kontextuális vizsgálata, (3) bármiféle szociális jelentésű és rendszeres összefüggés (gyakoriság, tipikusság, stb.) iránt érdeklődik, az abszolút szabályt csupán a tendencia szélső esetének tekinti, (4) úgy véli, hogy az empirikus vizsgálatok nagyban hozzájárulnak a nyelvi gyakorlatok lényeges tulajdonságainak megértéséhez, (5) elemzési kerete a „természetes diskurzusⁱⁱ” (vö. Cseresnyési 2004: 17). Ezt pontosítja Mary Bucholtz és Kira Hall értelmezése, mely szerint a kutatások tárgya nemcsak a nyelvi gyakorlatok, hanem a gyakorlatközösségek identitáskonstrukciói is (Bucholtz–Hall 2005), hiszen az identitások magukban foglalják a makroszintű de-

ⁱⁱ Kutatásom során diszkurzusnak tekintek minden olyan emberi produktumot, amely a kommunikáció céljait szolgálja, legyen az a szó köznapi értelmében vett verbális megnyilatkozás vagy bármilyen egyéb kommunikációs termék: rajz, festmény, fénykép, film, zene, tánc, stb. – bármi, ami valamilyen szándékolt üzenetet hordoz (vö. Terestyéni 2006).

mográfiai kategóriákat, az ideiglenes és kölcsönhatáson alapuló sajátosságokat és a résztvevők szerepét, valamint a helyi, néprajzilag kialakuló kulturális pozíciókat, valamint ezen identitások nyelvileg indexikálható címkéken, kategóriákon, állításokon, stílusokon vagy nyelvi struktúrákon és rendszereken keresztül alakulnak. Szociokulturális nyelvészetben a szerzők a nyelv, a kultúra és a társadalom kapcsolataival foglalkozó széles körű interdiszciplináris területet értik. Ez a kifejezés magában foglalja a szociolingvisztika, a nyelvi antropológia, a diskurzuselemzés társasorientált formáit (például a beszélgetéselemzést és a kritikai diskurzuselemzést), valamint a nyelvorientált szociálpszichológiát is. Nem célja, hogy tagadja e diszciplínák különbségeit, inkább elismeri a nyelv és az identitás címszó alá eső munkák teljes skáláját, és felkínál közös hivatkozási alapokat e megközelítésekre. Ezen posztdiszciplináris szemlélet célja, hogy az identitás elemzéséhez mint központi nyelvi jelenséghez eszköztárat adjon (Bucholtz–Hall 2005: 586). Kutatásom során e szemléletmódot követem, és jelen tanulmányban Woolard (2008) nyomán amellet érvelek, hogy szükségtelen a különböző terminusok (például szociolingvisztika, nyelvészeti antropológia, antropológiai nyelvészet, etnolingvisztika, szociokulturális antropológia, nyelvészeti etnográfia, kommunikáció etnográfia) közti jelentésárnyalatok fejtegetése:

„I did not in fact have the labels sociolinguistics and linguistic anthropology directly aligned with any of more fundamental divides that I perceived, more often I conceptualised one as nested within or overlapping broadly with the other”ⁱⁱⁱ (Woolard 2008: 432–433).

E terminusok közötti különbségek önkényes és talán kevésbé produktív számbavétele helyett a különböző irányzatok nyelvhez való viszonyát tekintem központi jelentőségűnek, hiszen ahogy Woolard is kifejti, a szociolingvisztika olyan kérdéseket fejez ki, melyeket a nyelvészeti antropológia már évtizedek óta kutat (Woolard 2008: 433), ráadásul az 1960-as évektől kezdve e terminusok szorosan kapcsolódnak egymáshoz^{iv} (Duranti 2003).

Az interpretatív antropológia perspektíváját alapul véve, valamint a bourdieu-i hagyományra ('occasionalist fallacy') támaszkodva kijelenthető, hogy minden értelmezés keretét az interpretatív gyakorlatközösségek biztosítják, melyek nem *a priori* definiált rétegek vagy csoportok, hanem az a bármekkora méretű és bármilyen összetételű, formális vagy informális szerveződésű csoport, amelynek tagjai valamilyen dolgot adott időben és kontextusban azonos módon értelmeznek (Boglár 2005: 30).

Kutatásom során kísérletet teszek arra, hogy hangsúlyt fektessek az etnicitás szituatív és kontextuális karakterének bemutatására, a konstruktivista nézőpont érvényesítésére, mely során a nyelvi gyakorlatokat és a nyelvi szocializációt identitásszimbólumként értelmezem. Az etnikai identitásra társadalmilag meghatározott (kulturális) különbségként tekintek, a társadalmi-kulturális-etnikai rétegeket és egymásra hatásukat pedig a szocializációs folyamatok megfigyelésén keresztül vizsgálom. Ennek oka, hogy e kutatások fő kérdésfelvetése, hogy miként szerveződik a szociális és kulturális gyakorlatok elsajátítása a különböző közösségekben (Rogoff 2003). Ráadásul annak

ⁱⁱⁱ Nem teszek különbséget a szociolingvisztika és a nyelvészeti antropológia címkék között. (...) Sokkal inkább az egymásba ágyazottságukat, valamint az átfedéseket hangsúlyozom. (A szerző fordítása.)

^{iv} E terminológiai kérdéskör szintézisét Franz Boas a relativista, kultúra-centrikus antropológiai iskola alapítójaként már a 20. század első felében megfogalmazta, a nyelvészetet az antropológia egyik területeként értelmezte. Ennek értelmében a nyelvészet tárgya nem válik el a nyelvnek nevezett tárgytól (Agha 2007), hanem annak minél több aspektusát magába foglalja.

megfigyelése is lehetővé válik, hogy milyen nyelvi ideológiák irányítják a gyermekekhez intézett nyelvi gyakorlatokat intézményi szintereken (Ochs–Schieffelin 2011). E tanulmányban a nyelvi ideológiák explicit metanyelvi diskurzusként, illetve a nyelvről mint rendszerről és a nyelvhasználatról mint cselekvésről megfogalmazott közléseként értelmeződnek (vö. Laihonon 2008: 699). A nyelvhasználók egyes csoportjai számára számos nyelvi forma és cselekvés indexikus (Ochs 1991; Silverstein 2003; Szalai 2010), tehát „azon túl, vagy ahelyett, hogy hozzájárul a denotációs jelentéshez, társas jelentésre utal, vagy segít annak létrehozásában” (Johnstone 2008: 133). A nyelvi formák (a kiejtősváltozatoktól egészen más-más nyelvekig) vagy a diskurzív stratégiák társadalmi identitásra (pl. etnicitásra, lásd Fought 2006) utaló jellé válhatnak (vö. Szalai 2010). A kutatás során pedig nemcsak a gyermek–pedagógus, hanem a gyermekek egymás közti kommunikációja^v (vö. Sterponi 2005; de León 2011) is a vizsgálat tárgya.

2.1 A megnevezés problémája

A kutatás során az etnicitásra – az antropológiai és a szociológiai diskurzusok „többszólamú” etnicitás- és identitásmeghatározásait alapul véve – nem mint előre adott tényezőre tekintek. Emellett hangsúlyozandó az is, hogy az egyén identitásválasztása a legfontosabb tényező, és kiemelem annak lehetőségét, hogy ugyanaz az egyén több identitást is választhat önmaga számára. Az etnocentrikus ideológiai tradíción (vö. Andorka 2003; Csepeli 2014) továbblépve úgy vélem, hogy a kulturális különbségek társadalmilag és alkalomhoz kötötten szerveződnek, tehát kiindulópontom szerint a társalgások résztvevői az etnicitás kategóriáját és jellemzőit közösen hozzák létre, így az etnicitás egy állandóan változó, dinamikus jelenség, és elkerülve az esszencializmust és a csoportizmust (vö. Barth 1996; Brubaker 2004; Eriksen 2000), Kovai (2017) nyomán a cigányt és a magyart nem bizonyos népcsoportok, etnikumok megnevezésére használom, hanem mint egy adott térben és időben létező társadalmi pozíciót és jelentést, amely etnonimák stigmatizáló és/vagy identifikáló hatásai folyamatos reflexió tárgyai (Kovai 2017: 12–13) az iskolában is.

2.2 Nyelvi szocializáció és hierarchia

Elemzésem a skála és a skálaugrás (Blommaert 2007) metaforája mentén végzem. Blommaert ezt a fogalmat a történelem- és globalizációtanulmányokból, valamint a társadalomföldrajzból kölcsönzi. Ahogy Pachné Heltai Borbála disszertációjában (2017) kifejti, a korábbi metaforákhoz képest Blommaert a skála fogalmával azt kívánja hangsúlyozni:

...hogy a horizontális mellett a tér vertikális dimenziójára, azaz annak megosztottságára, hierarchikus jellegére is gondoljunk. E hierarchikusság szem előtt tartása ugyanis a társadalmi események hatalmi kérdéseire irányítja a figyelmet (Blommaert 2007: 4). Emellett a skála nemcsak térbeli, hanem időbeli eseményekre vonatkozó koncepció is. Összességében a skálákat mint téri-időbeli kereteket, szinteket vagy mint dimenziókat érthetjük, amelyeken bizonyos nyelvhasználati formák, normatív elvárások érvényesülnek (Blommaert 2010: 36).

^v Magyarországi cigány gyermekek egymás közti nyelvi gyakorlatainak empirikus elemzésével tudomásom szerint Réger Zita kutatásai óta nem került sor.

Amikor skálaugrás történik (a fogalmat Blommaert Uitermark 2002-re hivatkozva vezeti be), tulajdonképpen interakcionális elrendeződések közötti mozgásról beszélünk, amelyek eltérő normákat, elvárásokat és társadalmi pozíciót jeleznek. Bizonyos emberek vagy csoportok ilyen skálaugrást kommunikációjuk során véghez tudnak vinni, míg mások nem. Az adott interakció skálájáról történő kiugrás egy másikba, amelyet a partner nem ismer, hatalmi taktika is lehet. Az, hogy a skálák között tehát valaki tud-e mozogni, a diszkurzív erőforrásokhoz történő hozzáférés kérdése (Pachné Heltai 2017).”

Az elmúlt évtizedek hazai és nemzetközi nyelvviszocializáció-kutatásait elméleti és gyakorlati szemléletmódjuk mentén három kategóriába sorolom a skála mint értékelés középpontba állításával, melyek:

- (i) a kompenzáció elve,
- (ii) a beszédmodok különbségeinek tudatosításának elve,
- (iii) az interpretatív gyakorlat elve a nyelvi szocializáció kutatásában.

| | Kompenzáció elve | Tudatosítás elve | Gyakorlat elve |
|----------------------|---|--|--|
| Nyelvszemlélet | Az elvárt normatív nyelvvaltozattól való eltérést fejlődésbeli késésként, nyelvi deficitként értelmezi. | A kommunikatív kompetencia eltérése helyezi a hangsúlyt, deficit helyett másságként értelmezi a különböző beszédmodokat. | nyelvi különbséget színterekhez köthető nyelvi gyakorlatok konfliktusaiként értelmezi. |
| Kutatói pozíció | A kutató elvárása a normatív nyelvvaltozathoz igazodik, ezt képviselve értékeli a nyelvi gyakorlatokat. | A kutató nem kapcsolódik be a társas gyakorlatokba, de nem értékeli, hanem leírja a résztvevők nyelvi gyakorlatait. | A kutató (már jelenlétével is) bekapcsolódik a társas gyakorlatokba, a kutatói attitűd formálódik a vizsgálat során. |
| Hazai kutatások | Nemesné–Sajtosné (2010) | Stewart (1987) Réger (1987) Bartha (2002, 2015) | Bakó (2008) Boros (2016) Heltai (2017) |
| Nemzetközi kutatások | Bernstein (1971) | García-Sánchez (2011) Fader (2006) Howard (2011) Paugh (2005) | García (2009) Jaspers (2017) |

1. táblázat: A nyelvviszocializáció-kutatások típusai

Réger Zita meghatározását követve – mely szerint a nyelvi szocializáció a szocializáció folyamatának része és egyben eszköze^{vi} – a felnőtt-gyermek (és gyermek-gyermek) beszédkapcsolat alakulása, s ebben a nyelv társadalmilag elfogadott használati módjainak az átadása része és egyben eszköze is a kultúra átadásának, magának

^{vi} Réger Ochs–Schieffelin meghatározását vette alapul: „language socialization comprises socialization through the use of language and socialization to use language” (Ochs–Schieffelin 1984: 163). Azaz „a nyelvi szocializáció egyrészt a nyelvhasználat segítségével elsajátított társas viselkedéseket, másrészt a nyelvhasználati szabályok társas elsajátítását jelenti. (Fordítás a szerző által.)

a szocializációs folyamatnak: társadalmi magatartást, személyiségfejlődést, gondolkodást, világméretet formáló tényező (Réger 1990: 9–10). Réger kutatását a tudatosítás elve szerint értékelem, hiszen nem deficitként, hanem másságként értelmezi a különböző beszédmódokat. Az '60-as évektől kezdve a labovianus hagyomány a nyelvi gyakorlatokat a stigma és a presztízs fogalmai mentén értékelté, és a nyelvviszocializáció-kutatások is e gyakorlatokat követték, így kialakítva a nyelvi hátrány fogalmát. A nyelvi hátrány terminust Réger Zitához hasonlóan az elsődleges szocializációs színtér és az iskola nyelvi elvárásainak különbsége felől értelmezem, ám a nyelvi hátrány helyett az iskola és a gyermek nyelvi konfliktusai meghatározást használom, hiszen a skála metaforája szerint lehetséges, hogy egy nyelvi gyakorlat az iskolában stigmatizál és/vagy marginalizál, míg a családi társas közegben épp ellenkezőleg hat.

2.3 A magyarországi cigány közösségek nyelvviszocializáció-kutatása

A magyarországi cigány közösségek nyelvi szocializációjának kutatásáról az 1970-as évek végétől beszélhetünk. Pap Mária, Réger Zita és követőik (Kovalcsik Katalin, Forray R. Katalin, Bartha Csilla) a kétnyelvűségre és az ebből fakadó nyelvi konfliktusokra helyezik a hangsúlyt kutatásaikban. Réger értelmezése szerint a romani–magyar kétnyelvű közösségekben végzett vizsgálatok:

„...lényegét tekintve nyelvészeti antropológia »fejlődési perspektívában«: a kulturális-szociális identitás és a nyelvhasználat összefüggéseit vizsgálja az egyén fejlődésének látószögéből, egy-egy adott kulturális kontextuson belül” (Réger 1998: 135, idézi: Szalai 2009: 16).

A felnőtt–gyermek interakciókon túl Réger Zita a gyermekek egymás közötti kommunikációját is vizsgálta, és elsők között vallotta Magyarországon a gyermeknyelv-kutatás képviselőjeként, hogy nemcsak a gyermekek nyelvi gyakorlatait irányítják meghatározott szabályok, hanem a hozzájuk intézett gondozói beszédet is, ezért nem csupán a gyermekek nyelvi produkcióját érdemes vizsgálni, hanem figyelmet kell fordítani a nyelvet elsajátító kisgyermeket körülvevő társadalmi és interakciós kontextusra is (vö. Kassai 1992; Szalai 2009). Réger vizsgálatában 13 kétnyelvű magyarországi cigány közösség vett részt, a kutatás főbb megállapításai a következők (Réger 1987; Szalai 2009):

- (i) a kisgyermeknek szóló romani gondozói beszéd alkalmazkodik a gyermekek életkori sajátosságaihoz,
- (ii) megjelenik az egyszerűsített regiszter (Ferguson 1977) használata,
- (iii) megjelennek a roma közösség orális kultúrájára jellemző interakciós sajátosságok és a felnőtt közösségben releváns beszédmódok,
- (iv) a gyermekközösségek szóbeli kultúrájának része a párbeszédes jelleg, az interaktív szövegalkotás a gyermekekhez szóló beszédben és a laza, improvizatív szerkesztésmód,
- (v) gyakori stratégia a dialógusimprovizáció,
- (vi) a legfiatalabb életkortól jellemző a mesemondás, amely a kisgyermeket hallgatóként és szereplőként is bevonja a történetekbe,

(vii) a tesztkérdések^{vii} használatának fokozatos csökkenése az életkor előrehaladtával.

Réger Zita beszédmódkülönbség-tudatosító kutatásai alapján is megállapítható, hogy a szociális és érzelmi tényezők a nyelvelsajátítás szempontjából fontosabbak, mint az életkornak megfelelő értelmi képességek. Az anya–gyermek beszédkapcsolat sajátosságait az adott kultúra, társadalmi csoportra jellemző értékek, hiedelmek, elvárások, közösségi és nyelvi normák együttese határozza meg, így természetesen a cigány gyermekek nyelvelsajátítása során is fontos szerepet kapnak ezek a normák. Az is nyilvánvaló, hogy a társas környezet, az elérhető nyelvi minta a nyelv elsajátításának elengedhetetlen feltétele. A nyelv társadalmilag elfogadott használati módjainak az átadása része és egyben eszköze is a kultúra átadásának, magának a szocializációs folyamatnak. Az anyai beszéd igen sok tekintetben sajátosan alkalmazkodik a kisgyermek adott kommunikációs és nyelvi szintjéhez, s ez bizonyos mértékig a gyerek fejlődésére is visszahat: lassíthatja vagy gyorsíthatja a nyelvi fejlődés folyamatát (Réger 1990).

A magyarországi cigány közösségek jellemzésekor ma már megkérdőjelezhető a kétnyelvűség kizárólagossága, átfedő diglossziás helyzetük azonban eddigi terepmunkáim során néhány helyen megfigyelhető volt. Jelen írás a *kettős átfedő diglosszia* fogalmát használja (Fasold 1984: 45), melynek lényege az emelkedett és a közönséges nyelvváltozatok szituatív meghatározottsága, amely jól illeszkedik a szociokulturális nyelvészet társas-konstruktivista irányzatához.^{viii} Ráadásul a magyarországi és (cseh)szlovákiai romani tárgyú kutatások már több évtizede a romani- és beáspárú kétnyelvűséget pontosabb szociolingvisztikai, valamint nyelvkörnyezettani kutatások hiányában pontosítás nélkül egyszerűen diglossziának tekintik (Pintér 2009). Pintér elismeri, hogy a magyar nyelv iskolai oktatás körében történő elsajátítása ma is létezik, de hangsúlyozza, hogy ez egyre kevésbé jellemző jelenség (Pintér 2009: 398).

Bartha Csilla a cigány gyermekek nyelvi helyzetével számos munkájában foglalkozik. Támaszkodva Réger Zita kutatásaira, a nyelvi egyenlőtlenségek vizsgálatán keresztül értekezik az oktatás szerepéről, miszerint az iskolába kerülő gyermekek több tekintetben hasonlóak egymáshoz: egyrészt az iskola minden tanuló számára új nyelvi közeget jelent, emellett az iskolában jóval nagyobb az esélye az eltérő nyelvi, kulturális, társadalmi háttérű gyermekek érintkezésének, mint korábban. A gyermekeknek először kell tudatosan megnyilvánulniuk formális beszédhelyzetekben, először találkoznak új regiszterekkel, a szókészlet eddig nem ismert egységeivel, szerepekkel és kommunikációs műfajokkal. A gyermekek különböző csoportjai a valóságban azonban eltérő kompetenciákkal, készségekkel, más-más tárgyi és nyelvi ismeret, viselkedési és interakciós minták, fogalmi készlet birtokában kerülnek az iskolába (Bartha 2002). A nyelvi hátrány tehát nem az iskolában alakul ki, a gyermekek már úgy kezdik meg tanulányaikat, hogy esélyeik nem egyenlőek, s erre épül az iskola

^{vii} A tesztkérdések olyan kérdések, amelyekre a kérdező személy már eleve tudja a választ, és a beszédpartnerétől ezt az előre megszabott választ várja el (Réger 1987: 59).

^{viii} Fasold tág diglossziameghatározása sok tekintetben rokonságot mutat Ferguson szemléletével. Gumperz a '80-as években bevezette a kettős beágyazott diglosszia ('double-nested diglossia') terminust a nyelvváltozatokon belüli különböző stílusok presztízsértékének jelzésére. A kétnyelvű helyzeteknek egy másik funkcionális felosztása található Haugennél (Haugen 1972: 334), aki különbséget tesz kiegészítő és helyettesítő típusok között az érintkező nyelvek szerepkörei alapján (Bartha 1999: 74).

merev elvárásrendszere, amely sokszor nemhogy csökkenti, inkább elmélyíti az egyenlőtlensége(ke)t. Nagy problémát jelent az is, hogy sokszor helytelenül ítélik meg a gyermekek tanulási képességeit (Bartha 2002: 85), tehát az oktatás nyelvi kihívásokkal találja szembe magát, melyet nem a megfelelő módon ismer fel.

3 A kutatás módszertana: a nexuselemzés

Kutatásom a nexusanalízis (Scollon–Scollon 2004) módszertani lépéseit követi, melynek során a társadalmi gyakorlatokat komplex rendszerben, a tágabb történelmi és társadalmi folyamatok, illetve az egyének kölcsönhatásaiban szerepet játszó helyi faktorok párhuzamos figyelembevételével vizsgálja. Ezen interdiszciplináris stratégia az etnográfia, a kritikai diskurzuselmélet és az interakcionális szociolingvisztika elméleti és módszertani elemeit használja fel. A Scollon házaspár alaszakai kutatásuk alapján dolgozta ki ezt az elemzéstípust a *Nexus Analysis. Discourse and the emerging Internet c.* programadó munkájukban, pontos fogalmi és módszertani lehetőségeket adva a társadalomkutatásokhoz, melyet egyre több nemzetközi kutatásban alkalmaznak (Hult 2014, Scollon–De Saint-Georges 2012, Pietikäinen 2013; Lane 2014). Magyarországi példa Bodó Csanád és Zabolai Margit 2016-os esettanulmánya és Pachné Heltai Borbála 2017-es doktori disszertációja, amelyben a szerző az analízis részletes bemutatását is vállalja. A módszer kiinduló gondolata szerint a társadalmi cselekvések három fő dimenzió találkozásában, metszéspontjában (nexusában) jönnek létre (1) a vizsgált szituációban megjelenő diskurzusok összessége, (2) az elrendeződés (műfaj, szituáció), amelyben a cselekvés történik, (3) a résztvevők magukkal hozott élettapasztalatai, tulajdonságai, cselekvéseikbe már beépült nézetei (Scollon–Scollon 2004: 19–20). A terepmunka a Scollon házaspár leírása szerint három lépésből áll: az első szint a nexusba való bekapcsolódás^{ix}, a második a társadalmi gyakorlat navigálása, a harmadik pedig a társadalmi gyakorlat megváltoztatása.

4 Identitáskonstruáló nyelvi gyakorlatok

E fejezetben az iskolai kontextusban vizsgált diszkurzustípusokat szemléltetem a kirajzolódó nexusok szerint. A kiválasztott szempontokat a 2016-ban elvégzett próbakutatásomra és a társalgáselemzés fő vizsgálati kérdéseire (Clift 2016) építettem, s e diszkurzustípusokat (beszédjog, csoporttagság, javítás, metanyelvi reflexiók, proxémika) fogom a jövőben összehasonlítani a gyermekek családi interakcióiban megjelenő diszkurzusaival. Jelen tanulmány a beszédjog, a csoporttagság, valamint a metanyelvi reflexiók mentén vizsgálja az identitás hitelesítésének folyamatát.

Vizsgálatom 2017 októberében kezdődött, 50 tanóra hospitálási jegyzőkönyvét készítettem el, a megfigyelt tanórák száma pedig 2018 márciusáig 80 tanóra. A megfi-

^{ix} E terminus Pachné Heltai Borbálánál (2017) nexussal való elköteleződés, Bodó Csanádnál és Zabolai Margitnál (2016) pedig bevonódás a gyakorlat nexusába fordításban jelenik meg. Kutatásom során a bekapcsolódás fogalmat használok annak jelzésként, hogy a társas jelentésteremtés folyamata a kutatás lezárásával e kapcsolódást is csökkenti.

gyelések első időszakában nem készítettem jegyzeteket, mert a megfigyelői pozícióm kialakítására helyeztem a hangsúlyt.

A következőkben a **csoporttagság-kategorizáció** két típusát szemléltetem, ugyanis az előzetes tanári interjúk során kialakított álláspont – mely szerint nem artikulálódik, „láthatatlan marad” az iskolában a tanulók etnikai identitása – már a megfigyelés elején sem igazolódott. Mind a megfigyelésre adott reflexiók, mind a nexusba való bekapcsolódás során előfordultak a gyermekek részéről olyan nyelvi gyakorlatok, melyek az identitás és az ezzel összekapcsolódó hierarchiális pozicionálást jelezték. Az első típusba azokat a gyakorlatokat sorolom, melyek a (1) megfigyelői pozícióra reflektáltak. Ebben az esetben mind a magyar–cigány, mind a tanár–diák oppozíció előfordult:

(1.a.) diák: (nekem) *lehet, hogy meg fog haragudni, de maga magyar vagy cigány?*

(1.b.) diák: *elmondja, mér van itt?*

én: *persze, csak késtél, és nem hallottad. na, én is iskolába járok, és dolgozatot kell írnom. azt nézem meg, hogy hogy beszéltek.*

diák: *és ha csúnyán beszélünk?*

én: *akkor csúnyán beszéltek.*

diák: *kérsz enni?*

(1.c.) diák1: *hazamész, vagy baszni mész?*

diák2: (rám célozva) *nem vagytok ám egyedül!*

diák1: *ne cigánykodj már!*

A második típus az (2) „elmerülés” metaforáján keresztül ragadható meg, mely terminust Williams vezette be a romológiai kutatásokba. E fogalommal írja le az egyénesen cigánynak mondott kultúrák jelenlétét a gádzsó világban. Az elmerülés – mint társadalmi gyakorlatok alapvető mintája – nyelvi jelenségként ragadható meg (Prónai 2006: 285): az antropológusok valamiféle fokozott érdeklődésről beszélnek, amit az általuk vizsgált csoportok tagjai a nyelvhasználattal kapcsolatos kérdések iránt tanúsítottak. Beszámol arról, hogy a rom cigányok, mielőtt megkérdeznék valakitől, hogy kicsoda, és a situációtól függően közösséget vállalnának vele, vagy elhatárolódnának tőle, az általa használt nyelv alapján azonosítják az illetőt, tehát a közösségek identitásstratégiáinak nyelvi alapon történő interpretációi, valamint a nyelvi gyakorlatok performatív tulajdonságai is identifikálnak.

(2.a.) diák: *van férje?*

én: *az élettársammal élek.*

diák: *hol?*

én: *a szerelmemmel? albérletben.*

diák: *úgy beszél, mint mi!*

(2.b.) én: *érti valaki ezt a számot? (zeneszöveg)*

diákok: (többiek kiabálva) *Dorina!*

én: *Dorina, milyen nyelveket tudsz?*

diák: *oláh, de csak értem.*

diákok: (többiek kiabálva) *nem igaz, tanár néni, beszél is, Dorina oláh.*

Ezen példák (2.a. és 2.b.) szorosan kapcsolódnak a **metanyelvi reflexiók** vizsgálati szempontjához. Az iskolai metanyelvi ideológiák – melyek szerint a standard az

egyetlen „helyes” beszédmód, a „csúnya beszéd” pedig szankcionált (vö. Szabó 2013) – megjelennek a diákok egymás közti interakcióiban is:

- (3) diák1: (5. osztályban nyelvtan óra közben) *adj egy kis szotyit!*
 diák2: *mér, te mit adtál, cigány?*
 diák1: *én se adtam.*
 diák2: *faszt adtál a szájba.*
 diák1: *ne beszéljé' má' csunyan!*

Megfigyeléseim szerint azonban a pedagógusok nem reflektálnak explicit módon a diákok nyelvi tevékenységében megjelenő ideológiákra és e beszédmódra, mely szempont már a **beszédjog kiosztásával** is szorosan összefügg, ahogy ezt a következő példa szemlélteti:

- (4) tanár: (7. osztályban) *a mai történelem órán feleléssel kezdünk.*
 diákok: (a hátsó sorban lefekszenek a földre és nyögnek)
 diák1: *tanár bácsi, ma ne. hátul másra gyakoroltak a cigányok.*
 tanár: (mondja a felelő nevét) *jöjjön ki a táblához.*

Az iskolai kontextusban megfigyelt identifikáció tehát eddig azt mutatja, hogy a társas gyakorlatok korábbi gyakorlatokba ágyazódnak, és a társas jelentésteremtés (térben, időben és kontextuálisan) állandóan változó, dinamikus jelenség, melyet a társas gyakorlatok résztvevői közösen hoznak létre.

5 Összegzés

A kutatás nyelvtudományi eredményei annak megértéséhez járulnak hozzá, hogy milyen tényezők mentén konstruálódik az etnicitás és az identitás a cigány közösségekben, ha a nyelvi szocializáció tükrében tekintünk e jelenségekre.

Kvalitatív kutatásom során az általam kiválasztott általános iskolán keresztül a későbbiekben a gyermekek családját is bevonom, így a gyermekek nyelvi gyakorlatai és a skálaugrás képessége mindkét szocializációs szintéren megfigyelhetővé válik.

Irodalom

- Agha, A. 2007. *Language and Social Relations*. Cambridge: Cambridge University Press.
 Andorka R. 2003. *Bevezetés a szociológiába*. Budapest: Osiris Kiadó.
 Bakó B. 2008. Cigány női normák egy erdélyi kis közösségben. In: Bakó B., Tóth E. Zs. (szerk.): *Határtalan nők – Kizártak és befogadottak a női társadalomban*. Budapest: Nyitott Könyvműhely Kiadó. 129–151.
 Barth, F. 1996. Régi és új problémák az etnicitás elemzésében. *Regio*. 1: 3–25.
 Bartha Cs., Hámori Á. 2011. Cigány közösségek, nyelvi sokszínűség és az oktatás nyelvi kihívásai. Magyarországi helyzetkép. *Európai tükrök*, 2011(3): 107–131.
 Bartha Cs. 2002. Nyelvi hátrány és iskola. *Iskolakultúra*, 6(7): 84–93.
 Bartha Cs. 2007. Nyelvcseré két magyarországi oláh cigány közösségben. „A régiek a régit beszélnek, mi már kavarjuk”. In: Bartha, Cs. (szerk.): *Cigány nyelvek és közösségek a Kárpát-medencében*. Budapest: Nemzeti Tankönyvkiadó. 244–266.
 Bartha Cs. 2009. Út a többnyelvűség felé? Nyelvideológiák, attitűdök és nyelvcseré: a kétnyelvűséggel kapcsolatos elképzelések szerepe a kisebbségi nyelvek megőrzésében. In: Borbély A., Vanconé Kremmer I., Hattyár H. (szerk.): *Nyelvideológiák, attitűdök és sztereotípiák*. 15. Élőnyelvi konferencia tanulmánykötete. 542.

- Bartha Cs. 2015. Nyelvi hátrány, avagy a cigány gyermekek oktatásának elmulasztott lehetőségei. In: Antalné Szabó Á. – Major É. (sorozatszerk.) *Szakpedagógiai körkép I. Anyanyelv- és irodalompedagógiai tanulmányok*. Bölcsész- és Művészetpedagógiai Kiadványok 3. Budapest: ELTE. 28–45.
- Beck, U. 2002. The cosmopolitan Society and its enemies. *Theory, Culture and Society*, 19(1-2): 17–44.
- Bernstein, B. 1971. Társadalmi osztály, nyelv és szocializáció. *Valóság*. 11. 47–57.
- Bíró B. 2012. Cigány gyerekek a gádzsó idő fogságában. *Képzés és Gyakorlat*, 10(3–4): 132–139.
- Blommaert, J. 2007. Sociolinguistic Scales. *Intercultural Pragmatics*, 4-1: 1–19.
- Blommaert, J. 2010. *The Sociolinguistics of Globalization*. Cambridge: Cambridge University Press.
- Bodó Cs., Zabolai M. E. 2016. A csángó műrokon és a keresztanya: Nem és nemzet a moldvai magyar nyelv revitalizációs programjában. *Társadalmi Nemek Tudománya. Interdiszciplináris eFolyóirat* 6:1, 159–180. Elérhető: http://ntefjournal.hu/vol6/iss1/bodo_zabolai.pdf
- Bodó Cs. 2016. *Nyelvi ideológiák és különbségek*. Cluj-Napoca: Institutul Pentru Studierea Problemelor Minoritatilor Nationale.
- Boglár L. 2005. *A tükör két oldala*. Budapest: Nyitott Könyvműhely Kiadó.
- Boros J. 2016. A roma/cigány tanulók támogatása. *Szociálpedagógia*, 3–4. 73–87.
- Brubaker, R. 2004. *Ethnicity Without Groups*. Cambridge, MA: Harvard University Press.
- Bucholtz, M., Hall, K. 2005. Identity and interaction: a sociocultural linguistic approach. *Discourse Studies*, 7. 585–614.
- Bucholtz, M., Hall, K. 2008. All of above: New coalitions in sociocultural linguistics. *Journal of Sociolinguistics*, 12: 401–432.
- Clift, R. 2016. *Conversation Analysis*. Cambridge: Cambridge University Press.
- Csepeli Gy. 2014. *Szociálpszichológia mindenkiben*. Budapest: Kossuth Kiadó.
- Cseresnyési L. 2004. *Nyelvek és stratégiák avagy a nyelv antropológiája*. Budapest: Tinta Könyvkiadó.
- De León, L. 2011. Language Socialization and Multiplay Participation Frameworks. In: Duranti, A., Ochs, E., Schieffelin, B. B. (szerk.) 2011. *The Handbook of Language Socialization*. Wiley-Blackwell. 81–113.
- Duranti, A. 2003 Language as culture in U.S. anthropology: Three paradigms. *Current Anthropology*, 44. 323–347.
- Duranti, A., Ochs, E., Schieffelin, B. B. (szerk.) 2011. *The Handbook of Language Socialization*. Wiley-Blackwell. A John Wiley & Sons, Ltd., Publication.
- Eckert, Penelope 2016. Variation, meaning and social change. In: Coupland, N. (szerk.) *Sociolinguistics. Theoretical Debates*. Cambridge-New York: Cambridge University Press, 68–85.
- Erdős K. 1958/1989. A magyarországi cigányság. In: Vekerti J. (szerk.) 42–56. (Első közlése: Néprajzi Közlemények 3. 152–173.)
- Eriksen, Th. H. 2000. Ethnicity and Culture: A Second Look. In: Bendix, R., Roodenburg, H. (szerk.) *Managing Ethnicity. Perspectives from folklore studies, history and anthropology*. Amsterdam: Het Spinhuis. 185–205.
- Fader, A. 2006. Language Socialization and Morality. In: Duranti, A., Ochs, E., Schieffelin, B. B. (szerk.) 2011. *The Handbook of Language Socialization*. Wiley-Blackwell. 322–341.
- Fasold, R. 1984. *The Sociolinguistics of Society*. Oxford: Basil Blackwell.
- Ferguson, Ch. 1977. Baby talk as a simplified register. In: Snow, C. E. – Ferguson, Ch. (szerk.) *Talking to children*. Cambridge: Cambridge University Press, 209–235.
- Fought, C. 2006. *Language and Ethnicity*. Cambridge, UK: Cambridge University Press.
- García, O. 2009. Education, multilingualism and translanguaging in the 21. century. In: Skutnabb-Kangas, T. (et al. szerk.) *Social Justice through Multilingual Education. Multilingual Matters*, Cromwell. 140–158.
- García-Sánchez, I. M. 2011. Language Socialization and Exclusion. In: Duranti, A., Ochs, E., Schieffelin, B. B. (szerk.) 2011. *The Handbook of Language Socialization*. Wiley-Blackwell. 391–421.
- Haugen, E. 1972. *The Ecology of Language*. Essays by Einar Haugen. Stanford: Stanford University Press.
- Heltai J. I. 2001. A magyarországi cigány lakosság által beszélt nyelvek. Elérhető: <http://www.nytud.hu/pp/heltai.html>

- Heltai J. I. 2015a. A romani és a beás nyelvi tervezés lehetőségei Magyarországon. *Regio*, 23(4): 25–27.
- Heltai J. I. 2015b. Roma/cigány gyerekek kétnyelvűségének néhány maganyelv-pedagógiai vonatkozása. *Anyanyelv-pedagógia* 2015/2. Elérhető: <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=568>
- Heltai J. I. 2017. A nyelvi hátrányos helyzet paradigmájának kritikája: zombik az iskolában. In: Benő A., Fazakas N. (szerk.) *Élőnyelvi kutatások és a dialektológia: Válogatás a 19. Élőnyelvi Konferencia - Marosvásárhely, 2016. szeptember 7–9. előadásaiból*. Konferencia helye, ideje: Marosvásárhely, Románia, 2016.09.07-2016.09.09. Kolozsvár: Erdélyi Múzeum-Egyesület (EME) 110–121.
- Heltai J. I., Kulcsár J. 2017. A romani beszéd és az óvoda. *Anyanyelv-pedagógia* 2017/4. Elérhető: <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=706>
- Howard, K. M. 2011. Language Socialization and Hierarchy. In: Duranti, A., Ochs, E., Schieffelin, B. B. (szerk.) 2011. *The Handbook of Language Socialization*. Wiley-Blackwell. 341–365.
- Hult, F. 2014. Drive-thru linguistic landscaping: Constructing a linguistically dominant place in a bilingual space. *International Journal of Bilingualism*, 18(5): 507–523.
- Jaspers, J. 2017. The transformative limits of translanguaging. *Working Papers in Urban Language and Literacies*. Paper 226.
- Johnstone, B., Kiesling, S. F. 2008. Indexicality and experience: Exploring the meanings of /aw/- monophthongization in Pittsburgh. *Journal of Sociolinguistics*, 12: 5–33.
- Kakuk M. 1997. *Cigány közvetítés jövevényszavak a magyar nyelvben*. Jász-kun-ság füzetek 10. Szolnok.
- Kassai I. 1992. *Út a nyelvhez – esély az életre*. BUKSZ 1992. 37–41.
- Kemény I. 2000. Anyanyelvi csoportok. In: Keményi, I. (szerk.) *A magyarországi romák*. Budapest: Press Publika. 25–30.
- Kovai C. 2017. *A cigány – magyar különbségtétel és a rokonság*. Budapest: L'Harmattan.
- Kovalcsik K. 1993. A beás cigányok népzenei hagyományai. In: Barna G. (szerk.) *Cigány népi kultúra a Kárpát-medencében a 18–19. században*. Salgótarján: Mikszáth Kiadó. 231–239.
- Kresztyankó A. 2016. Cigány jövevényszavak nyelvtanórán. In: Bagyinszki Sz., P. Kocsis R. (szerk.) *Anyanyelviünk évszázadai 2*. Budapest. ELTE BTK. 215–224.
- Kresztyankó A. 2017. *Cigány tanulók és tanáraik nyelvi attitűdjei*. *Anyanyelv-pedagógia* 10:3. 49–59.
- Laihonen, P. 2008. Language ideologies in interviews: A conversation analysis approach. *Journal of Sociolinguistics*, 12. 668–693.
- Landauer A. 2009. *A ticsánokról – kutatás közben*. Előadás A cigány kultúra néprajzi és történeti kutatása a Kárpát-medencében c. konferencián. Gödöllő, 2009. február 18–20.
- Lane, P. 2014. Nexus Analysis. In: Östman – Verschueren (szerk.) *Handbook of Pragmatics: 2014 Installment*. John Benjamins.
- Nemesné Kis Sz. – Sajtosné Csendes Gy. 2010. Hátrányos helyzetű tanulók szövegértési problémái. *Anyanyelv-pedagógia*. 2010/1. Elérhető: <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=242>
- Ochs, E. 1991. Indexing gender. In: Duranti, A., Goodwin, Ch. (szerk.) *Rethinking Context: Language as an Interactive Phenomenon*. Cambridge: Cambridge University Press. 335–358.
- Ochs, E., Schieffelin, B. B. 1984. Language acquisition and socialization. In: Schweder, R. A., LeVine, R. A. (szerk.) *Culture theory: Essays on mind, self and emotion*. New York: Cambridge University Press.
- Orsós A. 2010. A nyelvi szocializáció szerepe a hátrányos helyzetű tanulókatoktatásában. In: Németh Sz. (szerk.) *HH-s gyermekek a közoktatásban. Kézikönyv hátrányos és halmozottan hátrányos gyermekek neveléséhez és oktatásához*. Budapest: Raabe Kiadó.
- Orsós A. 2012. *Metszéspontok: a beás nyelv megőrzésének lehetőségeiről*. Pécs: Virágmandula Kft.
- Orsós A. 2015. A beás nyelv. In: Uő. (szerk.) *A romológia alapjai*. Pécs: Wlislöck Henrik Szakkollégium, PTE BTK. 151–180.
- Otsuji, E., Pennycook, A. 2010. Metrolingualism: fixity, fluidity and language in flux. *International Journal of Multilingualism* 7(3): 240–254.
- Pachné Heltai B. 2017. *Nyelvcseré és migráció egy többnyelvű magyarországi közösségben*. Doktori disszertáció. Kézirat.

- Paugh, A. 2005. Local Theories of Child Rearing. In: Duranti, A., Ochs, E., Schieffelin, B. B. (szerk.) 2011. *The Handbook of Language Socialization*. Wiley-Blackwell. 150–169.
- Pietikäinen, S. 2013. Heteroglossic Authenticity in Sámi Heritage Tourism. In: Pietikäinen, S., Kelly-Holmes, H. (szerk.) *Multilingualism and the Periphery*. New York: Oxford University Press. 78–94.
- Pintér T. 2009. Észrevételek a cigányok diglossziájáról. In: Borbély A. –Vanconé Kremmer I. – Hattvár H. (szerk.) 2009. *Nyelvideológiák, attitűdök és szereoptíák. 15. Élnyelvi Konferencia*. Tinta Könyvkiadó. Budapest. 393–401.
- Prónai Cs. (szerk.) 2006. *Cigány világok Európában*. Budapest: Nyitott Könyvműhely Kiadó.
- Réger Z. 1974. A lovari-magyar kétnyelvű cigánygyermek nyelvi problémái. *Nyelvtudományi közlemények* 76: 229–255.
- Réger Z. 1984. A cigányság helyzetének nyelvi és iskolai vonatkozásai – Álláspontok és viták. *Szociálpolitikai értekezések* 1984(2): 140–173.
- Réger Z. 1987. Nyelvi szocializáció és nyelvhasználat magyarországi cigány nyelvi közösségekben. In: Bartha Cs. 2007. (szerk.) *Cigány nyelvek és közösségek a Kárpát-medencében*. Budapest: Nemzeti Tankönyvkiadó. 80–116.
- Réger Z. 1990/2002. *Utak a nyelvhez. Nyelvi szocializáció – nyelvi hátrány*. Budapest: Akadémiai Kiadó.
- Rogoff, B. 2003. *The Cultural Nature of Human Development*. Oxford University Press.
- Schirm A. 2006. A magyar nyelv cigány eredetű jövevényszavai. *Nyelvtudomány* 2. 149–163.
- Scollon, S. W., de Saint-Georges, I. 2012. Mediated discourse analyses. In: Gee, J. P., Handford, Michael (szerk.) *The Routledge Handbook of Discourse Analysis*. London – New York: Routledge. 66–79.
- Scollon, R., Scollon, S. W. 2004. *Nexus Analysis. Discourse and the Emerging Internet*. London: Routledge.
- Silverstein, M. 2003. Indexical order and the dialectics of sociolinguistic life. *Language & Communication* 23: 193–229.
- Sterponi, L. 2005. Literacy Socialization. In: Duranti, A., Ochs, E., Schieffelin, B. B. (szerk.) 2011. *The Handbook of Language Socialization*. Wiley-Blackwell. 227–247.
- Stewart, M. 1987. „Igaz beszéd” – avagy miért énekelnek az oláh-cigányok? Valóság. 30(1): 49–64.
- Szabó, T. P. 2013. Az iskolai metanyelvi szocializáció dinamikusa. Anyanyelv-pedagógia. 2013/3. Elérhető: <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=469>.
- Szalai A. 2007. Egységesség? Változatosság? A cigány kisebbség és a nyelvi sokféleség. In: Bartha Cs. (szerk.) *Cigány nyelvek és közösségek a Kárpát-medencében*. Budapest: Nemzeti Tankönyvkiadó. 20–52.
- Szalai A. 2009. *Átok, feltételes átok és társadalmi nem erdélyi roma közösségek nyelvi ideológiájában és gyakorlataiban*. Pécsi Tudományegyetem. Doktori disszertáció.
- Szalai A. 2010. Nyelvi ideológiák és a társadalmi határok kijelölése egy erdélyi gábor roma közösségben. In: Feischmidt M. (szerk.) *Etnicitás: különbségteremtő társadalom*. Budapest: Gondolat Kiadó, MTA Etnikai-Nemzeti Kisebbségkutató Intézet. 418–429.
- Terestyéni T. 2006. *Kommunikációelmélet. A testbeszédtől az internetig*. Budapest: Typotex Kiadó.
- Torgyik J. 2005. Nyelvi szocializáció és oktatás. *Új Pedagógiai Szemle* 2005(3): Elérhető: <http://epa.oszk.hu/00000/00035/00090/2005-03-ta-Torgyik-Nyelvi.html>
- Uitermark, J. 2002. Re-scaling, ‘scale fragmentation’ and the regulation of antagonistic relationships. *Progress in Human Geography* 26(6): 743–765.
- Vekerdi J. 1999. Cigány. In: Fodor I. (szerk.) *A világ nyelvei*. Budapest: Akadémiai Kiadó.
- Woolard, K. A. 2008. Why *dat* now?: Linguistic-anthropological contributions to the explanation of sociolinguistic icons and change. *Journal of Sociolinguistics* 12: 432–452.

Az N-nAk köszönhetően mint névutói szerkezet

Pomázi Bence

ELTE BTK Nyelvtudományi Doktori Iskola
pomazi.bence@gmail.com

Kivonat: A dolgozat amellett érvel, hogy a köszönhetően datívuszragos N-t (névszót) vonzó névutóvá grammatikalizálódott. A grammatikalizációs folyamat során fogalmi jelentése gyengült, és a 'vmi révén' viszonyjelentést vette fel. Ezt kollostrukturális elemzéssel igyekszik bizonyítani a tanulmány, vagyis egy korpuszból lekérdezett véletlen mintán nézi át, hogy az eredetileg pozitív orientációjú N-ekkel előforduló határozóragos melléknév hogyan veszít preferenciájából, és fogad el maga mellett negatív értékjelentésű N-eket is, ezzel grammatikaibb jelentésűvé alakulva. A korpusznyelvészeti módszerek mellett funkcionális kognitív-konstrukciós elemzés vizsgálja a névutóvá válás eredményét.

1 Bevezetés

A dolgozat az N-nAk köszönhetően szerkezettel foglalkozik, és a köszönhetően névutói szerepe mellett érvel. A dolgozat 2. pontja a névutó definiálásának problémáját tárgyalja főképp 20. századi, a hagyományos leírást követő magyar szakirodalom kritikai feldolgozása alapján azzal a céllal, hogy világossá tegye a névutó kategória jellemzőit. A kategória prototípuselvű értelmezése megalapozza azt a döntést, hogy a köszönhetően kifejezés beletartozik a névutók állományába. Kitérek a névutókra jellemző formális és funkcionális kritériumokra az egyes leíró nyelvészeti munkákban, illetve tanulmányokban. Az N + névutó szerkezet konstrukcióként leírható kompozitum (Tolcsvai Nagy 2013: 268; Langacker 2008: 60), amelyben a névutó a profilmeghatározó komponens. A kompozitum az elemi mondat fogalmi magjában lévő folyamat másodlagos szereplőjét dolgozza ki (N), a névutó a szereplők viszonyát profilálja. Ennek megfelelően a dolgozat 3.1. pontjában az N-nak köszönhetően szerkezetet is konstrukcióként mutatom be. Mivel a ragos névszók lexikalizálódása és névutóvá válása grammatikalizációs folyamat eredménye (vö. Dér 2004, 2012, 2013a, 2013b), a 3.2. pont a grammatikalizáció felől közelít a problémához.

A dolgozat empirikus módszertannal dolgozik, amelyet a 3.3. pontban mutatok be. A példamondatok az MNSz.² (Oravecz–Váradi–Sass 2014) korpuszának adatai, amelyek egy véletlenszerű adatokat felsorakoztató 191 mondatos adathalmaz elemei. Azokat a mondatokat, amelyek célzott keresés eredményei, a nyelvi adatok melletti (CK) címkével jelölöm.

Ludányi Zsófia, Krepsz Valéria, Grácsi Tekla Etelka (szerk.): *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2018*. Budapest: MTA Nyelvtudományi Intézet. 2018.

2 A névutók

A névutókat a *Magyar grammatika* (MGr., amelynek koncepciója szerint Balogh 2000, Keszler 2000a és 2000b egységesen viszonyul a kategóriához) és Sebestyén Árpád monográfiája is a viszonyozók kategóriájába sorolja, tehát önálló, fogalmi szójelentéssel nem rendelkező kategóriaként tartják számon (vö. Balogh 2000; Sebestyén 1965). A „morfológiai természetű” kategorizáció a névutót a főnévi viszonyragokkal rokonítja, ez összhangban van azzal, hogy a kifejezésnek grammatikalizálódott jelene van.

Sebestyén azt is megjegyzi, hogy bizonyos névutók szórendje megfordítható, ez főképp a ragvonzó névutóknál fordulhat elő, hiszen ott a rag és az őt vonzó névutó közötti szoros és jelölt szemantikai kapcsolat lehetővé teszi, hogy a kettőt együtt dolgozzuk fel akkor is, ha nem névszó-névutó sorrendben állnak a mondatban. Ezzel szemben a *Magyar grammatika* pozícióhoz is köti a névutóságot.

A *Magyar grammatika* és Sebestyén Árpád monográfiája is három-három típust különít el alaktanilag a névutók között. A MGr. alapján ez a három csoport a ragtalan névszókhoz járó névutók, a ragvonzó névutók, illetve a birtokos személyjelet tartalmazó névutók (Balogh 2000: 263). Sebestyén hasonló szempontok alapján osztja fel a szófaji osztályt: egyszerű, ragvonzó és ragos névszói, s emellé egy kiegészítő kategóriát is felvesz: az esetenként névutószerű kifejezéseket, amelyeknek a névutóval mutatott funkcionális hasonlóságát emeli ki (Sebestyén 1965: 169), vagyis hogy a tartalmi jelentés kezd viszonyjelentéssé olvadni. Dér Csilla Ilona 2012-es tanulmányában ragvonzó és nem ragvonzó névutókról ír (Dér 2012: 12).

A *vminek köszönhetően* nem szerepel a MGr., Rác és Szemere (1985), valamint Sebestyén egyetlen névutó-kategóriájában sem, ezek a szakirodalmak nem számoltak még azzal, hogy ez a szerkezet egy ragvonzó névutós szerkezet lehet. Ligeti-Nagy Noémi korpusznyelvészeti vizsgálatában a „névutószerű viselkedést produkáló” szavak közé sorolja (Ligeti-Nagy 2018: 61) a lexémát. Úgy ítéli, hogy mondatbeli elemzésük módját még tükrözi morfológiai felépítésük, viszont jelentésük és pozíciójuk alapján felveendők a névutói kategóriába.

A névutók prototipikusan az N névszó után helyezkednek el. (N alatt dolgozatban azt a névszót értem, amellyel együtt morfológiai típusú szószerkezetet alkot a névutó.) Erre utal a MGr., Sebestyén Árpád és Dér Csilla Ilona is (vö. Balogh 2000: 261; Sebestyén 1965: 11; Dér 2004: 37–39; Dér 2012: 11; Dér 2013a: 13). Károly Sándor a *nézve* határozói igenévnél a névutóvá válásához kritériumként használta a megfelelő szórendet, névutónak csak a névszó-névutó szórendi helyzet kialakulása után tekintette a szót (Károly 1953: 461–462).

Dérnek a korpuszokon folytatott vizsgálataiban azonban azt mutatták, hogy a ragvonzó névutók kétféle pozíciót is felvehetnek: N-t követő hátulsó pozíciót és az N vagy a névszói szerkezet előtti (Dér 2012: 11).

Dolgozatomban megvizsgálom, hogy a *köszönhetően* esetében mennyire érvényesül a prototipikus sorrend, illetve hogy a különböző mintázatokhoz milyen egyéb szemantikai, strukturális, illetve szövegtípussal összefüggő sajátosságok köthetők.

Összefoglalásképpen elmondható, hogy a korábbi szakirodalom definitív erejű jellemzőnek tartotta: általánosan 1) a mondatrészszerre (önálló határozói szerepre) való alkalmatlanságot; részlegesen 2) a kötött (N névutó) szórendet (Károly 1953); 3) a határozószóéval rokonítható jelentést, és a hasonló grammatikalizációs folyamatot (ragos névszóból lexikalizálódás) a névutók bizonyos csoportjaiban. A korábbi mun-

kák a *köszönhetően* kifejezést sem névutóként, sem névutószerű kifejezésként nem említik.

3 Az esettanulmány elméleti háttere és módszertana

3.1 Kognitív-konstruktív nyelvészeti megközelítés

A *vminek köszönhetően* névutós szerkezetet úgy magyarázom, mint a KÖSZÖNETMONDÁS forgatókönyvi modelljét követő forráskonstrukcióból eredő szerkezetet. A konstrukcióról Goldberg kiemeli, hogy minden grammatikai szinten előforduló forma és (szemantikai vagy diszkurzív) funkció párosokról van szó. Ebben az elméleti keretben minden olyan nyelvi mintázat konstrukciónak számít, amely valamilyen szerkezetben fordul elő (Goldberg 2006: 5). Egy-egy nyelvi mintázat több konstrukciót is felidézhet, és különböző elvontsági szintű sémával is összefüggésbe hozható (pl. a *vág* ige elsődlegesen 'éles tárggyal való darabolás'-t jelent, de ugyanez az ige az OKOZOTT MOZGÁS konstrukciójában is előfordul, l. *bevágta a labdát a kapuba*).

Az ige típusjelentése szereplői viszonyokat határoz meg (Goldberg terminusaival: argumentumstruktúra-konstrukciót hív elő), az ige specifikálja a szemantikai/szintaktikai információt, amelyet a konstrukció hordoz. A *(meg)köszön* háromargumentú ige, három szereplő viszonyát dolgozza fel, a három szereplő három különböző, láncot alkotó függőségi viszonyban dolgozható ki (a *megköszön* alapjelentésének megfelelő konstrukcióban *(meg)köszön vki vmit vkinek/vminek*). A lánc fogalma itt Imrényi (2017: 240) értelmezését követi, vagyis szavak folyamatos kombinációját jelöli. A *(meg)köszön* ige konstrukciójában a következő tematikus szerepek fordulnak elő: ágens (a KÖSZÖNŐ, aki megköszön valamit), beneficiens (akinek szól a köszönet, a köszönetmondás CÍMZETTje) és témaként az, amit megköszönnek, vagyis a MEGKÖSZÖNT DOLOG. Az alábbi mondat szolgál példaként:

[1] *Ahogy szóhoz jutottam, megköszöntem Ferinek a segítségét (CK).*

A mondatban a konstrukció szereplői viszonyai szempontjából az ágens az egyes szám első személyű, alanyként reprezentált beszélő, a datívuszragos főnév, *Ferinek* pedig egy olyan referenciális szereplőt jelöl, amely felveszi azt a tematikus szerepet, akinek a részére történik a megköszönés, a *segítség* pedig a megköszönt dolog. Ez egy aktív, cselekvő konstrukció, a teljes szerkezet a köszönést mint folyamatot profilálja, az ige bővítményei a fenti módon kidolgozzák az ige tematikus alszerkezeteit (vö. Langacker 2008: 360; Tolcsvai Nagy 2013: 178–179).

A *vminek köszönhetően* igenévi alaptagú szerkezet, amely az igei alaptagtól eltérően valósítja meg a KÖSZÖNETMONDÁS forgatókönyvi sémájának nyelvi reprezentációját.

A *-hető* képzőbokor a *köszön* igéből passzív jelentésű melléknévi igenevet képez:

[2] *a Ferinek köszönhető segítség*

A konstrukció a *-hető* képzőbokor műveletének eredményeként aktívból passzív jelentésű lesz, a szereplői viszonyok (az alapszerkezethez képest) átrendeződnek: a megköszönt dolog 'folyamat' jelentésű, a szerkezet alaptagjaként realizálódik az új szerkezetben; az alapszerkezet köszönést fogadó, beneficiensét változatlanul N-nAk

szerkezet dolgozza ki potenciális beneficiensként (vö. [2]). Az ágens nem dolgozható ki az igenévi szerkezeten belül.

A melléknévi igenévből konverzióval válik melléknév. Az így létrejövő melléknévekről Palágyi László megjegyzi, hogy gyakran ÁLLAPOTként profilálódnak (Palágyi 2016: 229–230), ahogy a *köszönhetően* is. Melléknévi igenévi tulajdonsága, hogy részlegesen megőrzi az igei bővítményeit, melléknévi tulajdonsága viszont az, hogy a folyamatot egy dolog JELLEMZŐjeként reprezentálja, és ebben a jelentésben lehetséges a JELLEMZŐ megjelenítése határozói viszonyokban, ez pedig együtt jár határozórag felvételével. A *köszönhetően* morfológiai konstrukció funkciója az, hogy a KÖSZÖNETMONDÁS fogalmi szerkezetét úgy jelenítse meg, mint egy előtérben álló esemény körülményét (lásd [3]). A szereplői viszonyok részlegesen megegyeznek az igenévnél leírtakkal, de a konstrukció absztraktabb.

[3] *Ferinek köszönhetően* [magmondat]

A *Ferinek* határozó feldolgozódik a lehetséges köszönetmondás címzettjeként, a mondat fogalmi magjához (Tolcsvai Nagy 2013: 282) képest azonban a *Ferinek köszönhetően* kompozitum a pozitív eredményhez (okozott, előidézett fejleményhez) vezető okot, az oksági jellegű körülményt dolgozza ki. A konverziós elemzés magyarázatot tud adni arra, hogy miért jelenhet meg az *-en* rag a szerkezetben. Problémája azonban, hogy a kategóriaváltás nem egyértelmű, mert – mint látjuk – a kifejezésnek igeire-igenévre jellemző bővítménye van. A *köszönhetően* absztrahálódása azonban lehetővé teszi, hogy az N-nAk komponens ne (vagy ne csak) folyamat végpontjaként legyen feldolgozva, hanem a fogalmi mag felől nézve egészlegesen, ok jellegű kiindulópontként (is), ahonnan a magmondatot (vö. Imrényi 2011) mint annak fejleményét értjük meg.

3.2 Grammatikalizáció és a konstrukció átrendeződése

A névutóvá válás folyamata grammatikalizáció. A grammatikalizáció során (általában) egy lexikai elem jelentése grammatikaiabbá válik (Dér 2013b: 17). A *köszönhetően* nem prototipikus (deverbális) melléknév. Az absztrahálódás (a *(meg)köszön* ige jelentéséhez képest) a 'potencialitás' és a 'jellemző' jelentések révén az igenévképzés esetében is megfigyelhető. A dolgozat amellet érvel, hogy ez a jelentésváltozás belépés a grammatikalizációs folyamatba, amelynek az eredményeként a *köszönhetően* névutóvá válik, tehát kategóriaváltás következik be.

A továbbiakban a ragos melléknév → névutó lexikalizációval járó grammatikalizációs folyamatot a konstrukciók viszonyában, konstrukciós változásként is értelmezzük.

A dolgozat tehát, mint említettem, a megköszönés konstrukciójából indul ki. A *megköszön valamit valakinek* alapvetően pozitív polaritású szerkezet. A köszönés konstrukciójából kiindulva itt úgy változik a konstrukció. 1. A köszönés a mondatban háttérbe kerül, az előtérben lévő folyamatot a mondat állítmánya közvetíti, és általában ez az a dolog, amelyet meg lehet köszönni, de ez az új konstrukcióban eseményszerű. 2. A névutós szerkezet *-nak/-nek* ragos névszója megtartja a tematikus szerepét, az ő részére történik/történne a megköszönés. 3. A megköszönő szerepét továbbra is az alany viszi, de már nem ő az, aki pozitívként értékeli az eseményeket.

Az egyik korpuszbeli mondat ezt jól mintázza:

[4] *Köszönöm a srácoknak, hiszen nekik köszönhetően ünnepelhetek felhőtlenül.*

A [4] mondat konstrukcióját az alábbi, 1. ábra szemlélteti:



1. ábra.

A konstrukció átalakulása a [4] példa alapján

Az 1. ábra az alábbiak szemléltetésére szolgál. Legfelül az aktív konstrukció, alul pedig az *N-nAk köszönhetően* konstrukciója. Az ágenst mindkettőben az egyes szám első személyű igei személyrag (*Köszönöm*, *ünnepelhetek*) által reprezentált szereplő képviseli (a példában mindkettő a megnyilatkozóra referál). A beneficienst az első tagmondatban a *srácoknak* (*N-nAk*), a második tagmondatban az ezzel koreferens *nekik* (az *ők* személyes névmás datívuszi alakja) dolgozza ki. A megköszönt dolog azonban másképp konstruálódik meg a két szerkezetben. A pontozott vonal a szaggattal szemben azt hivatott jelölni, hogy a két tagmondatban nem ugyanúgy konstruálódik meg a MEGKÖSZÖNT DOLOG. Amíg a fenti konstrukcióban ezt egy főnévi természetű szerkezet hivatott képviselni, tehát 'dolog' jelentésű, tárgyiasított (reifikált) N, addig az *N-nAk köszönhetően* konstrukcióban ezt egy 'folyamat' jelentésű igei szerkezet teszi.

A [4]-es mondatban az egyes szám első személyű alany a megköszönő, a *srácok* datívuszragos alakja pedig az a szereplő, akinek köszönetet mond a megnyilatkozó. A megköszönés folyamata előtérben van, a tárgya viszont kidolgozatlan. Ehhez képest a második tagmondatban a köszönés folyamata háttérbe kerül, miután az előtérben lévő folyamatot a tagmondat állítmánya, az *ünnepelhetek* képviseli. Ahogy az 1. ábra is mutatja, a datívuszragos névszó (itt a *nekik* névmás) tematikus szerepe nem változik. A megköszönt dolog, a téma csak az aktív konstruálású első tagmondatban jelenhetne meg tárgyként, de implicit; a *köszönhetően* mellett ez a 'dolog' jelentésű szerep nem jelenik meg a mondatban, hanem metonimikus anaforával teremt vele kapcsolatot a 'folyamat'. Bybee is felhívja a figyelmet rá, hogy a grammatikalizáció folyamatával már meglévő konstrukciókból új konstrukció jön létre (Bybee 2010: 106).

3.3 Az empirikus vizsgálat

Mivel a grammatikalizációs folyamatban az értékjelentés mint a tartalmas jelentés része halványulhat, hipotézisem szerint a *köszönhetően* is elő fog fordulni semleges és

negatív értékjelentésű N-hez kapcsolódva. Ezek alapján feltételezem, hogy a *köszönhetően* névutóvá vált.

A kutatást a Magyar nemzeti szövegtár (MNSz.²) korpuszán végeztem. A teljes szövegtárban kerestem rá olyan datívuszragos főnevekre, amelyek hétszavas környezetében előfordul a *köszönhetően* szóalak. Ebből szűrtem le egy véletlenszerűen válogatott kétszáz adatos listát, a dolgozatban szereplő statisztikai adatok erre a véletlenszerű listára vonatkoznak. Továbbá egy másik keresést is elvégeztem nagy kezdőbetűs *Köszönhetően* kifejezésre is, hogy meg lehessen állapítani, a teljes szövegtárban arányosan hány olyan szerkezet fordul elő, amelyben a *köszönhetően* mindenképpen megelőzi a ragos N-t. A kollokálódás vizsgálatához lekérdeztem a kétszáz-as lista, illetve a teljes szövegtár kollokációs listáját is. Stefanowitsch és Gries kollostrukurális analízisnek nevezik azt a korpuszbeli keresési és adatelemzési módszert, amely arra világít rá, hogy bizonyos konstrukciók bizonyos helyein melyek azok a lexémák, amelyek gyakrabban fordulnak elő a többinél, tehát a *köszönhetően* kollexémái (Stefanowitsch–Gries 2003: 214–215).

A dolgozatban olyan adatokat is közlök, amelyek nem a véletlenszerűen kiválasztott kétszáz-as listában fordultak elő, de ezeket az adatokat a kvantitatív elemzésben nem használtam fel.

4 A kutatás eredményei

A teljes szövegtárban az olyan datívuszragos főnevek száma, amelyek hétszavas környezetében előfordul a *köszönhetően* vagy a *Köszönhetően* szavak valamelyike, 90 034. A nagykezdőbetűs *Köszönhetően*-nel kezdődő kifejezések száma 413. Ennek az aránya tehát a korpuszban 0,46%. A véletlenszerű kétszáz-as listában 9 mondat ismétlődő adat volt, tehát 191 értékes adat származott innen. Azoknak az adatoknak a száma, ahol a *köszönhetően* megelőzte az N-t, 12. Ez 6,28 %. Ha a két adatot összevetjük, akkor azoknak a mondatoknak az aránya, amelyeknél megelőzi (akár mondatkezdő pozícióban, akár a mondat belsejében) a névutó az N-t, az összes mondathoz képest 6,74%. Ez tehát azt jelenti, hogy a prototipikus szórendi mintázat alapján a *köszönhetően* követi az N-nak kifejezést, tehát a névszó-névutó sorrendnek tipikusan megfelel.

Egy másik, jelentéstani változás, amikor a *köszönhetően* -nak/-nek ragos főneve negatív jelentésű. A szerkezet prototipikusan pozitív polaritású névszóval vagy pozitív polaritású mondatban fordul elő. Például:

- [5] *Sikerének köszönhetően Radwanska a következő heti ranglistán a harmadikról a második helyre lépett előre, elérve karrierje eddigi legelőkelőbb pozícióját.*

Ebben a mondatban a *köszönhetően* által vonzott datívuszragos névszó is pozitív értékjelentésű, illetve az a folyamat, amelyet ez kivált, szintén előnyös, a tágabb értelemben vett kotextus polaritása az *előrelép* és a *legelőkelőbb* szavakból is megállapíthatóan pozitív.

Ezek után érdemes megnézni a *köszönhetően*-nel előforduló kollexémákat. A 191 adat között az 1. táblázatban ismertetett főnevekkel (lemmatikus keresésben, tehát lexémákkal) fordul elő többször is a *köszönhetően* névutó (levonva a mutató és vonatkozó névmásokat is, amelyek koreferens szerepben utalószóként vagy kötőszóként szerepelnek összetett mondatokban).

| Lemma | Előfordulás |
|---|-------------|
| fejlesztés, változás, munka, győzelem | 3 |
| szabályozás, rezsicsökkentés, növekedés, játék, együttműködés, támogatás, politika, kapcsolat | 2 |

1. táblázat. A *köszönhetően* lexéma mellett 1-nél többször előforduló lemmák a 191-es listában

Philip Dilts (2010) felhívja arra a figyelmet, hogy egyes szavak, szerkezetek, illetve konstrukciók gyakran fordulnak elő tipikusan bizonyos orientációjú szavakkal, ami azt jelenti, hogy ezek a nyelvi egységek szemantikai preferenciákkal rendelkeznek (pl. OKOZÁS konstrukciója negatív értékjelentésű kollokációs szerkezetekkel). Itt a szemantikai orientációt úgy definiálja, mint ami alapján a beszélők ítéletet hoznak szavakról vagy fogalmakról. Az ANEW (Affective Norms of English Words) metódusát használva beszélőkkel tesztelték bizonyos lexémák szemantikai orientációját a kontextusukból kiragadva (Dilts 2010: 104–105), majd vizsgálták azok preferenciáját a korpuszban.

Ezek főként pozitív jelentésű szavak, amelyeknek a polaritása nem dönthető el egyértelműen, az pedig pozitív jelentésű kontextusban szerepel. Egy pár példával bemutatva az alábbi mondatokban fordulnak elő:

- [6] *A folyamatos fejlesztéseknek köszönhetően már a szélessávú internet hálózatot is kibővítettük.*
- [7] *A bíróság szerint a vádirati tényállások többsége a „rendkívül alapos, szakszerű nyomozati munkának” köszönhetően igaznak bizonyult*
- [8] *Húsz éves a Magyar Máltai Szeretetszolgálat mentőszolgálat. Ez az állandó pénzhiány ellenére, viszont a folyamatos önkéntes munkának köszönhetően van így.*
- [9] *Szigorú monetáris politikájának [...] köszönhetően a dél-amerikai államba áramlani kezdett a külföldi tőke.*
- [10] *A ZTE győzelmének köszönhetően az első helyen áll a tabellán.*
- [11] *A most bejelentett változásoknak köszönhetően, javítani tudjuk a vállalat költséghatékonyságát.*
- [12] *Stirlinget ezek után őrnaggyá léptették elő és új rangjának valamint kapcsolatainak köszönhetően el tudta érni, hogy az SAS jobb és nagyobb mennyiségű felszerelést kapjon.*

Azoknál a példáknál, amelyeknél önmagában a kollexémákból nem lehet eldönteni a szerkezet polaritását, ott segítségül szolgál a szövegkörnyezet, ezekben az esetekben mindig pozitív jelenetről van szó. Kijelenthető tehát, hogy azokban az esetekben, amelyekben a kétszázas véletlenszerű listában egynél többször előforduló tartalmas szavak kollexémaként viselkednek a *köszönhetően* névutóval, ott ezek mind pozitív jelentésű lexémák vagy pozitív jelentésű mondatkörnyezetben fordulnak elő.

A névutóvá válás útját jól jelzi, ha a *köszönhetően* melléknév veszíteni kezd abból a preferenciájából, hogy pozitív értékjelentésű N-nel alkot egy szerkezetet. Ezt negatív jelentésű főnevek mellett lehet megfigyelni, és jelentheti azt, hogy a melléknévi jelentéstartalom kezd üresedni, vagyis a *köszönhetően* ezekben az esetekben viszony-

szói funkcióban fordul elő. Továbbá erre utalhatnak azok az adatok is, amelyeknél értékjelentésük eldöntéséhez egyre nagyobb ko(n)textuális ismeretre van szükség.

A szemantikai tartalom változására utaló jelenség az ironikus hangnem megjelenése is. Ennek egy első állomása lehet (írásban), amikor ezt az ironikus használatot a mondat megalkotója például idézőjelekkel jelöli. A következő mondatokban fordult ez elő:

[13] *az előrejelzések szerint különösen a női COPD [= krónikus obstruktív tüdőbetegség] előretörésének „köszönhetően” a COPD néhány éven belül leszorítja a tüdőrákot a dobogóról.*

[14] *az elbocsátásoknak „köszönhetően” tanárhiány alakult ki (CK)*

A [13]-as és [14]-es mondat megkonstruálójá itt már negatív polaritású jelenetekben alkalmazza a *köszönhetően*-t mint névutót. Az idézőjel kitétele a megnyilatkozás írójának metapragmatikai reflexiója, hiszen prototipikusabbnak gondolja, hogy az adott szerkezetet inkább pozitív értéktartalommal használják a beszélők.

A negatív értéktartalom megjelenése azonban a grammatikalizációs folyamat egyik jele, hiszen jelzi, hogy a *köszönhetően* preferenciája változik, és az N orientációja már többféle lehet. Joan Bybee a grammatikalizációról megjegyzi, hogy amikor egy lexikai elem grammatikaibbá válik, annak disztribúciója és funkciója a folyamat közben megváltozik (Bybee 2010: 106–107). Dér Csilla rámutat a folyamattal kapcsolatban, hogy a grammatikalizáció során egy nyelvi egység egyszerre veszít referenciális jelentéséből és nyer funkciójelentést (Dér 2013b: 25).

A fenti szerkezetekben ez a következőképpen figyelhető meg. A *köszönhetően* lexikai jelentése a *köszön* ige pozitív jelentéstartalmából származik, ahogy azt az [1]-es példamondat is szemlélteti. A *köszönhetően* ebből *-hető* képzőbokkal képzett és határozóraggal ellátott melléknév, tehát a referenciális jelentéstartalma csak pozitív jelentést engedne meg. Azonban ez a jelentéstartalom a [13]–[14]-es példákban veszít fogalmiságából, grammatikaibbá válik, éppen azért, hogy funkciója változott. Ez a funkcióváltozás disztribúciós változással is együtt jár, amikor megjelenik az egyértelműen negatív jelentéstartalmú *elbocsátás* lexémával is.

A 191-es listában 16 alkalommal fordult elő negatív jelentéstartalommal rendelkező mondatokban az N-nAk *köszönhetően* szerkezet. Íme néhány példa:

[15] *az elképesztő számításerejű otthoni gépek rohamos terjedésének köszönhetően néhány tíz-száz fertőzött gépről elementáris támadást lehet indítani bármilyen hálózatos szolgáltatás ellen.*

[16] *A bekerített német erőknek így repülőgépen kellett az utánpótlást eljuttatni, de ezek zöme megsemmisült a szovjeteknek köszönhetően, így magukra maradtak.*

[17] *Ennek, illetve a jég és a hó által okozott albedo [=napsugárzást visszaverő képesség] változásnak köszönhetően az északi félteke is gyorsabban melegszik, mint a déli.*

[18] *Vezetni szeretne mindenki, arra az útra terelni az eseményeket, amelyet helyesnek lát, ennek köszönhetően azonban csak eltávolodnak egymástól.*

[19] *a teljes gyógyulásnak feltételei vannak, amit én a betegségemnek köszönhetően nem tudok megoldani.*

Ezekben a mondatokban a *köszönhetően* már elveszítette a fogalmi, tartalmas jelentését, és az N-nAk *köszönhetően* szerkezet nagyjából 'vmi révén' jelentésben értelmezhető. A konstrukció megváltozik, ahogy Bybee is írja, egy meglévőből egy másikra, nemcsak azáltal, hogy a tárgyragot felvevő páciens szerepét megjelenítő névszó 'folyamat' jelentésű igeiként konstruálódik meg az új szerkezetben, hanem a funkció oldalán is: a datívuszragot felvevő N már nem feltétlenül olyan szereplője a jelenetnek, aki egy kedvezményezettet valami számára előnyös dologhoz juttat.

A megváltozott konstrukcióban az alábbi szereplők érhetők tetten: egy okozott dolog, a folyamat (állítmányként); illetve az okozó, amelyet a *-nak/-nek* ragos névszó és annak névutója képvisel (a kompozitumszerkezeten belül N dolgozza ki), a klasszikus grammatika kategorizációja alapján okhatározóként.

Itt tehát azt figyelhetjük meg, hogy a szerkezet polaritása az eredetihez képest változott, a *köszönhetően* orientációs preferenciája gyengül az egyértelműen pozitív értékjelentésű N-ektől, és a szerkezet polaritásának megállapításához egyre nagyobb mértékben a ko(n)textusra kell támaszkodni. A fogalmi jelentés leginkább talán ott üresedik ki, ahol nem lehet egyértelműen eldönteni, hogy az okozóként fellépő N-nAk *köszönhetően* negatív vagy pozitív hatást ért el.

[20] *Az elmúlt napok viharos időjárása során a heves záporoknak és felhőszakadásoknak köszönhetően az ország több területén is szokatlanul nagy mennyiségű csapadék hullott.*

[21] *A gondos előkészületeknek köszönhetően számos bankot sikerült kirabolnia.*

Ezekben a mondatokban tulajdonképpen olyan ténymegállapításokat olvashatunk, amelyek polaritásának eldöntése problémákba ütközhet. A [20] mondat orientációja nem állapítható meg egyértelműen, [21]-é pedig a tágabb kontextustól függ: az sem mindegy, melyik szereplő nézőpontjából tekintjük az eseményeket, hiszen a bankrabló sikeres rablása a bűnöző szempontjából jó eredmény, társadalmi megítélése viszont negatív. A grammatikalizációs folyamatban a legelőrehaladottabb lépcső az, ha arányosan egyre nagyobb azoknak a típusoknak a száma, amelyekben negatív vagy semleges orientációjú kifejezés fordul elő, főleg, ha – mint a [20]–[21]-es mondatokban is – a prototipikus névszó-névutó sorrendet is hozzák.

A névutósodás egy másik sarkalatos kérdése a szórend. Erre vonatkozóan a lekérdésben az alábbi arányokat találtam: a 191 mondatból 179 alkalommal a *köszönhetően* követte az N-t, míg 12 olyan eset volt, ahol N hamarabb fordult elő a mondatban, tehát az esetek mindössze 6,28%-ában fordult elő a nem prototipikus szórend. Ezekben az esetekben gyakran távolabb került a *köszönhetően* a datívuszragos névszójától, viszont mivel vonzza a ragot, a kettőt egységként kezelni nem okoz gondot. Az alábbi példákban fordult elő az atipikus sorrend:

[22] *Az exportmennyiség 12,5 százalékkal nőtt a harmadik negyedévben, köszönhetően az export Ázsiában elért 26 százalékos növekedésének.*

[23] *A csapat végül megmenekült, köszönhetően a télen kinevezett új mesternek, Hans Meyernek*

[24] *egyre nehezebb lényeges újításokat mutatni az előző részekhez képest, de a FIFA 2003 esetében ez várhatóan sikerül, köszönhetően a játék új motorjának.*

[25] *Szerencsére és köszönhetően egy „Örök Fradistának”, Szécsi Antal keramikusművésznek, a szívélyes fogadtatást izléses ajándékokkal [...] tudtuk megköszönni.*

Érdekes tapasztalat volt megfigyelni, hogy a névutó N sorrend az előforduló adatok közül leggyakrabban a sporttal kapcsolatos szövegtípusokban fordult elő (l. [23], [24], [25] mondatokat), a 12 eset 25%-ában. Ez a sorrend éppen amiatt, hogy nem tipikus, a nézőpont fókuszába állítja az N-t, és úgy bizonyult, hogy a sporttal kapcsolatos történetekben kulcsszerephez jut az, aki a sikerért vagy sikertelenségért, tehát az okozott dologért felelős.

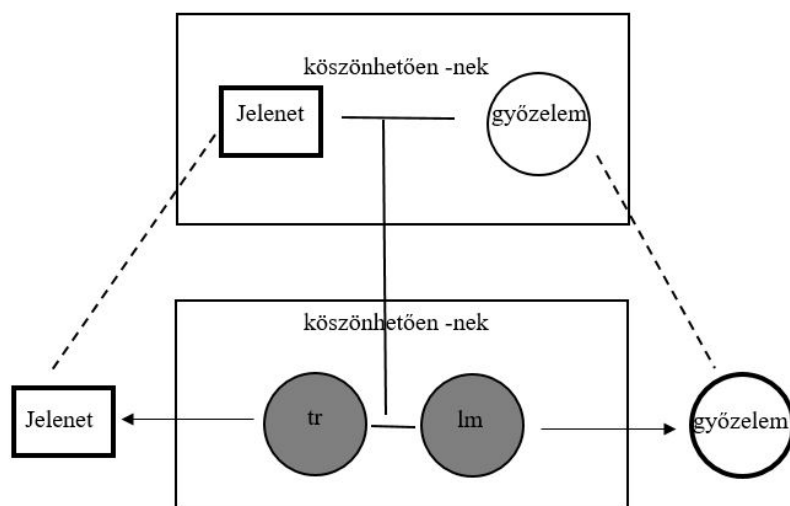
4.1 Az N-nak köszönhetően mint kompozitumszerkezet

Az N-nAk köszönhetően szerkezet egységei komponensszerkezetként vesznek részt a névutói kompozitumszerkezet létrehozásában (vö. Tolcsvai Nagy 2013: 268). A kompozitumszerkezet profilmeghatározója a *köszönhetően*, ennek sematikus alszerkezetét dolgozza ki a -nAk ragos főnév. A rag kijelöli a szerkezetben azt az N-t, amely az okozásra képes, ez lesz az aktív zónája N jelentésszerkezetének (vö. Tolcsvai Nagy 2013: 270). Vegyük példaként a következő mondatot, amelyet a 2. ábra mutat be:

[26] *A ZTE győzelmének köszönhetően első helyen áll a tabellán.*

Az *első helyen áll a tabellán* egy olyan jelenetet, állapotot ábrázol, amelynek elsődleges figuráját a ZTE dolgozza ki. A mondat előterében álló jelenete (*első helyen áll*) tehát már nem a köszönetmondáshoz kapcsolódó folyamat. Ennek az állapotnak az okozóját jelöli ki a *győzelmének köszönhetően* kompozitumszerkezet. Az N-nAk köszönhetően sematikus szerkezet olyan oksági viszonyt fejez ki, amelyben az N-nAk egy irányultságot kifejező komponensszerkezet (ebben a -nAk a profilmeghatározó, ennek kidolgozási tartománya a landmark, kidolgozója pedig főnévi kifejezés). A datívuszrag által kifejezett irányultsági viszony egy trajektor és egy landmark viszonya, amelyek közül a trajektor a mondat fogalmi magját alkotó folyamatként reprezentálódik a mondatban, a landmark pedig N (itt a győzelem, lásd *győzelmé-nek*).

Az ábra eltekint a konstituencia modellálásától, nem építi össze a *győzelmének* kompozitumát (-nek profilmeghatározó sematikus alszerkezettel, és ennek *győzelm-é-kidolgozásával*) a *köszönhetően* sematikus fogalmi szerkezetével. Az ábra a *köszönhetően-nek* viszonyt veszi alapul. A Jelenet feliratú doboz tartalmazza az *első helyen áll a tabellán* állapotot, amelyet ki lehetne bontani, de a konstrukció ábrázolásához most nem szükséges. Ennek és a *győzelemnek* a kapcsolatát dolgozza ki a -nAk köszönhetően szerkezet. A teljes hatókör része a konceptualizáló, aki nyelviileg kidolgozottan nem szereplője a jelenetnek, és konceptualizálóként sem jelenik meg explicit módon a mondatban. A *köszönhetően* névutó azonban implicit módon kifejezi egy értékelő szubjektum (a konceptualizáló) működését (szubjektívizáció, lásd Langacker 2008; Tolcsvai Nagy 2013: 166). A megnyilatkozó, a jelenet megkonstruálója szubjektívaltán teszi hozzáférhetővé, hogy az okozott eredményt, a jelenetet pozitívan értékeli, kedvezőnek tartja.



2. ábra.

A -nek köszönhetően kompozitumszerkezete

5 Összegzés

A dolgozat azt kívánta bemutatni, hogy a *köszönhetően* grammatikalizációs folyamat névutóvá vált. A *köszönhetően* alapszavának konstrukciójából kiindulva [vki_{AG} köszön vmit_{TÉMA} vkinek_{BEN}] a dolgozat leírta az N-nAk *köszönhetően* konstrukciós sémáját, és a jelentésváltozással összefüggésben feltárta a konstrukciós séma átrendeződését.

A *köszön* alapszó pozitív értékjelentésével összefüggésben a *köszönhetően* névutó is pozitív orientációjú, a korpuszadatok alapján a pozitív preferencia erős tendencia. A grammatikalizáció ösvényére lépve (a jelentés absztrahálódásával) azonban az értékjelentés (orientáció) meggyengül, a pozitív értékjelentése semlegesülhet, és ennek eredményeként egy nagyobb fokban grammatikalizálódott viszonyjelentés, 'vmi révén' kezd körvonalazódni. A kollostrukturális adatelemzési módszerrel az bizonyosodott be, hogy a szerkezet megjelenik ironikus használatban, illetve negatív polarítású jelenetek okának/okozójának kidolgozásában is, valamint olyan helyzetekben, ahol a konstrukció polarítását csak nagyobb ko(n)textuális ismeretek bevonásával lehet megállapítani.

A szórendi helyzettel kapcsolatban a korpuszadatok kvantitatív elemzése – az ismertetett mintán – azt mutatja, hogy a szerkezet szórendi mintázata az esetek döntő többségében megfelel a főnév névutó prototipikus sorrendnek.

Források

MNSz.² = Magyar nemzeti szövegtár v2.0.4. Elérhető: <http://corpus.nytud.hu/mnsz>. Letöltve: 2018. március 30.

Irodalom

- Balogh, J. 2000. A névutó. In: Keszler, B. (szerk.): *Magyar grammatika*. Budapest: Nemzeti Tankönyvkiadó. 261–265.
- Bybee, J. 2010. *Language, Usage and Cognition*. New York: Cambridge University Press.
- Dér, Cs. I. 2004. A grammatikalizáció „tettenérése” – Néhány határozói igenévből keletkezett, határozóragot vonzó névutóknak kialakulásáról. In: Ladányi, M., Dér, Cs., Hattyár, H. (szerk.) 2004. „... még onnan is eljutni túlra...” *Nyelvészeti és irodalmi tanulmányok Horváth Katalin tiszteletére*. Budapest: Tinta Könyvkiadó. 37–43.
- Dér, Cs. I. 2012. Mennyire prototipikus névutók a ragvonzó névutók a magyarban? In: Tolcsvai Nagy, G., Tátrai Sz. (szerk.) 2012. *Konstrukció és jelentés*. Budapest: Eötvös Loránd Tudományegyetem. 11–30.
- Dér, Cs. 2013a. A kontextus szerepe a magyar ragvonzó névutók viselkedésében. In: Csepregi, M., Kubinyi K., Sivonen J. (szerk.) *Grammatika és kontextus. Új szempontok az uráli nyelvek kutatásában III. (Uralisztikai tanulmányok 20.)*. Budapest: ELTE BTK Finnugor Tanszék. 9–19.
- Dér, Cs. I. 2013b. Mi számít grammatikalizációnak? Kísérlet a fogalom újradefiniálására. In: Forgács, T., Németh M., Sinkovics B. (szerk.) 2013. *A nyelvtörténeti kutatások újabb eredményei VII*. Szeged: Szegedi Tudományegyetem Magyar Nyelvészeti Tanszék. 17–28.
- Dilts, P. 2010. Good nouns, bad nouns: what the corpus says and what the native speakers think. In: Gries, S. Th., Wulff S., Davies M. (szerk.) *Corpus-linguistic applications. Current studies, new directions*. Amsterdam, New York: Rodopi. 103–117.
- Goldberg, A. 2006. *Constructions at Work. The Nature of Generalization in Language*. Oxford: Oxford University Press.
- Imrényi, A. 2011. *A magyar mondat viszonyhálózati modellje*. Doktori disszertáció. ELTE BTK. Elérhető: <http://doktori.btk.elte.hu/lingv/imrenyandras/diss.pdf>
- Imrényi, A. 2017. Egymásbaágyazottság a magyar mondat hálózatos szerkezetében. *Magyar Nyelvőr*, 141(2): 239–249.
- Károly, S. 1953. A nézve névutó történetéhez. *Magyar Nyelv*, 49(3–4): 461–463.
- Keszler, B. 2000a. A mai magyar nyelv szófaji rendszere. In: Keszler, B. (szerk.): *Magyar grammatika*. Budapest: Nemzeti Tankönyvkiadó 67–76.
- Keszler, B. 2000b. A határozószó. In: Keszler, B. (szerk.): *Magyar grammatika*. Budapest: Nemzeti Tankönyvkiadó 209–222.
- Langacker, R. W. 2008. *Cognitive Grammar. A Basic Introduction*. Oxford: Oxford University Press.
- Ligeti-Nagy, N. 2018. *Névutók, előre! Korpuszvezérelt elemzés a névutószerű elemekről*. In: Vincze, V. (szerk.) 2018. *XIV. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged: Szegedi Tudományegyetem, Informatikai Intézet. 52–63.
- Oravecz, Cs., Váradi T., Sass B. 2014. *The Hungarian Gigaword Corpus*. In: Calzolari, N., Choukri K., Declerck T., Loftsson H., Maegaard B., Mariani J., Moreno A., Odijk J., Piperidis S. (szerk.) *Proceedings of Ninth International Conference on Language Resources and Evaluation (LREC 2014)*. Elérhető: http://www.lrec-conf.org/proceedings/lrec2014/pdf/681_Paper.pdf. Letöltve: 2018. március. 30.
- Palágyi, L. 2016. A magyar főnév és melléknév kognitív szerveződésének alapjai és konverziós viszonyai. *Nyelvtudományi Közlemények*, 112: 207–242.
- Rác, E., Szemere Gy. 1985. *Mondattani elemzések*. Budapest: Tankönyvkiadó.
- Sebestyén, Á. 1965. *A magyar nyelv névutórendszere*. Budapest: Akadémiai Kiadó.
- Stefanowitsch, A., Gries S. Th. 2003. Collocations: Investigating the interactions of words and constructions. *International Journal of Corpus Linguistics*, 8: 209–243.
- Tolcsvai Nagy, G. 2013. *Bevezetés a kognitív nyelvészetbe*. Budapest: Osiris Kiadó.

Köztesnyelvi udvariassági választások a kérés beszédaktusában

Sitkei Dóra

ELTE BTK Nyelvtudományi Doktori Iskola
sitkeidora@gmail.com

Kivonat: Jelen kutatás során a nyelvtanulók köztesnyelvi választásait vizsgálom: hogyan változtatják meg kérésstratégiájukat a különböző szociopragmatikai és beszédaktus-specifikus tényezők hatására, az enyhítő módosítók mely kombinációjával élnek a kérés beszédaktusában. A kutatás során összevettem a köztesnyelvi kérési megnyilatkozásokat, kérésstratégiákat – az ugyanezen szituációkra adott – célnyelvi (magyar L1) és anyanyelvi (lengyel L1) kérésekkel is. Az összevetés eredményeképp sajátos köztesnyelvi kommunikációs stratégiákat (például a nyelvtanulói költség-haszon elv) lehetett megfigyelni a vizsgált anyagban, amelyek eltérnek a nyelvtanuló előtt álló lehetséges (cél- és anyanyelvi) mintáktól.

1 Bevezetés

A kérést – a szituáció tényezői alapján – egy adott udvariassági szinten fejezi ki a beszélő (l. Fraser és Nolan 1981), azzal a céllal, hogy így biztosítsa teljesítését, s egyúttal ne változzon a felek között a kapcsolat. Az udvariassági szint elérésének, „beállításának” eszközei többek között a belső és külső módosítók. Ezen módosítók az adott beszédszándék, például a kérés kifejezésének a másokra gyakorolt negatív hatását csökkenthetik (Alcón-Soler–Safont-Jordà–Martínez-Flor 2005), enyhíthetik, vagy a negatív hatással kevésbé törődve például kiemelhetik a kérés fontosságát, sürgősségét – így biztosítandó a kérés sikerességét (Blum-Kulka–Olshtain 1984; Trosborg 1995).

A jelen kutatás során köztesnyelvi produktumokban az enyhítő módosítók használatát vizsgálom a pragmatikai kutatások általános keretén belül (l. pl. Szili 2004). Választ keresek arra a kérdésre, hogy az egyes beszélők a különböző szituációkban az adott szociopragmatikai és beszédaktus-specifikus tényezők alapján a módosítók mely kombinációját választják, a tényezőkben történő változás mennyiben befolyásolja kérésstratégiáik megváltoztatását (vö. Blum-Kulka–House 1989; Szili 2004: 110–112).

2 A kutatás alapfogalmai

2.1 A köztesnyelvről

A nyelvtanulás kreatív folyamat, amely során a nyelvtanulók a környezetükkel való interaktív kapcsolat révén létrehozzák a megfigyelt szabályosságok egyfajta belső reprezentációját, ami dinamikusan változik a nyelvtanulási folyamat alatt (Corder 1992). Ezen belső reprezentáció, nyelvi rendszer megnevezésére a köztesnyelv (interlanguage, rövidítve: IL) fogalmát Selinker vezette be 1972-ben. Jelen vizsgálat során Selinker (1972) köztesnyelvi kutatási elképzelését követem: a magyar köztesnyelvi produktumokat összevetem az – egyazon szituációra adott – magyar (Selinker fogalmával: célnyelvi) és lengyel anyanyelvi megnyilatkozásokkal.

2.2 A kérés beszédaktusáról

A kérés a nyelvtanulók számára a köszönés mellett a legfontosabb beszédaktus (Suh 1999), pragmatikailag megfelelő véghezvitele a sikeres kommunikáció egyik kulcsa, így azt megfelelően elsajátítva sikerélményt tudnak szerezni a magyarul tanulók.

A kérésben megjelenő beszédzándék (Gordon–Lakoff 1975) a hallgató viselkedésének befolyásolására vonatkozik (Reichenbach 1947), azt szeretné elérni a beszélő, hogy a hallgató megtegyen valamit (Blum-Kulka 1991). Trosborg (1995) definíciójában pontosít a kérés tárgyával kapcsolatban: teljesítése hasznot hoz a beszélőnek.

A hallgató számára „költéssel” jár a kérés teljesítése (House–Kasper 1987: 25), s a beszélő tudatában van, hogy – részben ezen költség miatt – a hallgató el is utasíthatja (Tracy et al. 1984: 514). Holtgraves és Yang szerint (1992: 246) minél nagyobb a kérés, a beszélő annál udvariasabban fejezi ki, Kasper (1994: 3209) szavaival pozitív korreláció áll fenn a kontextuális tényezők (társas/szociális távolság, hatalom és a kényszerítő jelleg), valamint azon befektetés között, hogy a megnyilatkozás udvarias legyen.

2.3 A belső és külső módosítók rendszere

A kérést nyelvi szinten udvariassá az adott kultúrában többek között az erre a célra elfogadott nyelvi eszköztárral teheti a beszélő, amelyre Blum-Kulka–Olshtain (1984) a belső és külső módosítók rendszereként utal. A kérésekben használt módosítók egyfelől enyhíthetik a kérés negatív oldalát, másfelől fokozó elemként, például sürgethetik a kérés teljesítését. Fraser (1978) szerint enyhítéskor a beszélő az üzenetben szereplő beszédzándék erejét árnyalja, csökkenti, fokozás esetén pedig növeli.

A belső módosítók a sok esetben komplex kérésstratégián belül, a kérést leginkább kifejező (tag)mondat, a főcselekmény részeként fejtik ki hatásukat (Szili 2004). A magyar nyelv esetén enyhítő belső módosító lehet például a hatóige-képző, a *tud* segédige, a tagadás és a magázás. A külső módosítók, a támogató elemek – a főcselekmény közvetlen nyelvi kontextusában találhatóak (Blum-Kulka–Olshtain 1984). Külső módosító lehet a kérés bevezetéseként elhangzó elnézés-kérés, a segítségkérés vagy a magyarázat adása.

2.4 Az udvariasságról

Mindenkinek van egy elképzelése arról, hogy milyen az udvarias viselkedés vagy egy udvarias megnyilatkozás. Nézzük meg, hogy nyelvészeti pragmatikai keretben – a teljesség igénye nélkül – időrendben haladva milyen udvariassági megközelítések fogalmazódtak meg.

Az udvariasság a beszédpartnerek közötti társas kötelék megerősítését szolgálja (Goffman 1967), csökkenti a súrlódást a társas kapcsolatokban (Lakoff 1973), a konfliktus elkerülésében játszik szerepet (Leech 1983). Leech hat udvariassági alapelve (a tapintat, a nagylelkűség, a szerénység, a helyeslés, az egyetértés és a szimpátia) az udvariasságot skálaszerűnek tekinti, ahol a skálák a következők: költség-haszon, választhatóság, indirektség, tekintély és társadalmi távolság. A jelen kutatás szempontjából fontos kiemelni a költség-haszon skálát, amely azt vizsgálja, hogy a beszélőnek „mennyibe kerül” az adott megnyilatkozás, és ez „mennyi hasznot hoz” neki vagy a hallgatónak (Balogh 2002). Clark és Schunk (1980) Leech-hez képest pár évvel korábban megfogalmazta „költség-haszon” udvariassági elvét, amely szerint olyan mértékben udvarias a beszélő – ami adott költséggel jár számára –, amennyire haszna származik belőle.

Fraser és Nolan (1981) az udvariasságot egyfajta társalgási szerződés betartásához hasonlítja, másfelől akkor udvarias egy megnyilatkozás, ha a kontextuális tényezők figyelembe vételével, azok alapján megfelelőnek tekinthető.

2.5 A konvencionizált nyelvi formákról

Kérés megfogalmazásakor konvencionális formával élve biztosíthatja a beszélő, hogy a szituációnak megfelelő udvariassági szinten fejezze ki beszédszándékát, míg az attól eltérő formák pragmatikailag motiváltak lesznek (Bańcerowski 1999).

Diskurzusaink során beszédszándékaink verbalizálásakor sok esetben nyelvi mintákat követünk (vö. Goffman 1953: 1), konvencionizált, rutinszerű formákkal élünk (vö. Ferguson 1976: 140–141). Becker (1975) szituációs megnyilatkozásnak nevezi a lexikális frázisok azon csoportját, amelyek természetüket tekintve többnyire kevés variációt megengedő teljes mondatok, funkciójuk alapján pedig olyan megnyilatkozások, amelyekről a beszélők tudják, hogy ez az, amit mondani kell (ez a helyénvaló, az elfogadott) bizonyos kontextusokban. Hymes (1968) nyelvi rutinként utal a konvencionizált nyelvi formákra, használatuk egyik céljának pedig a viselkedés kiszámíthatóvá tételét jelöli ki. Terkourafi (2015) akkor tekint egy kifejezést konvencionizáltnak, ha a kifejezés egy adott szituációs kontextusban sűrűn használt, s így egy adott beszédszándékot, interakciós célt meg tud vele valósítani a beszélő, másfelől a hallgató számára is egyértelműbb ilyenkor az elhangzott megnyilatkozásban rejlő beszédszándék (Copestake–Terkourafi 2010: 125). Manes és Wolfson (1981) a bókokat vizsgálva állapította meg, hogy a konvencionális bókok használata megvédi a beszélőt a félreértésektől (egy, a konvencionalistól eltérő, kreatív bók esetleg nem okoz jó érzést a hallgatóban).

3 Módszer

3.1 A DCT-s vizsgálati módszerről

A DCT (diskurzuskiegészítő teszt) a köztesnyelvi pragmatikai kutatások egyik fő eszköze, a rögzített szociolingvisztikai tényezők alapján ez az adatgyűjtési forma praktikus a kontrasztív vizsgálatokhoz (Beebe–Cummings 1996). Beebe–Cummings az elutasítás beszédaktusát vizsgálva bebizonyították, hogy az adatközlők az élőnyelvi változatokhoz hasonló diskurzuskiegészítő tesztbeli válaszokat adtak az alkalmazott szemantikai formák tekintetében. Vagyis a módszer előnye, hogy „hozza” a leggyakoribb, standardizált, sematikus válaszokat (Blum-Kulka 1987; Beebe–Cummings 1996), viszont a módszer negatív oldalaként említik a pszichológiai tényezőt: nem biztos, hogy éles szituációkban is ezen stratégiákat alkalmazzák az adatközlők, a teszt nem tudja érzékeltetni a beszélgetésben részt vevők közötti interakció pszichoszociális dinamikáját (Beebe–Cummings 1996). A kutatásban részt vevő azt írja le, amit gondol, hogy mondana (Babbie 1998: 274), de nem biztos, hogy ezt mondaná „az életben” (Boxer 1996) – a spontán beszédhez képest több idő jut a válaszolónak át-gondolni, hogy mi lehet a legjobb válasz. Kasper (2000: 329) szerint a DCT megfelelő adatgyűjtési módszer, ha a vizsgálat célja a beszélő pragmalingvisztikai és szociopragmatikai tudásának feltérképezése (vagyis az adott szituáció kontextuális tényezői közepette mely stratégiai és nyelvi formákat tart megfelelőnek az adott kommunikatív nyelvi cselekedet végrehajtásához).

A diskurzuskiegészítő teszt alkalmazásával így tehát megismerjük azon köztesnyelvi megnyilatkozásokat, amelyet a nyelvtanuló a legmegfelelőbbnek (tipikusnak, l. Weizman 1989) gondol az adott szituációban attitűdje, nyelvtudása, illetve célnyelvi ismeretei alapján (Kasper 2000: 330; Rose–Ng Kwai-Fun 2001: 155).

3.2 A vizsgálat módszertani jellemzői

Nyílt végű diskurzuskiegészítő tesztre 80 lengyel anyanyelvű, magyarul tanuló és legalább középfokon tudó egyetemistától kaptam magyar nyelvű válaszokat. A kérdőíven a szituációkat magyarul és zárójelben angolul adtam meg. A válaszokat összevettem egyazon szituációkra 40 magyar anyanyelvű egyetemista által – magyar nyelvű kérdőívre – adott magyar, és 30 lengyel anyanyelvű hallgató – tegező stílusban megfogalmazott, lengyel nyelvű kérdőívre adott – lengyel nyelvű válaszával. A köztesnyelvi válaszokat papíralapon, az anyanyelvi válaszokat internetes felületen adták az adatközlők. A köztesnyelvi válaszolóknak körülbelül a 15%-a részt vett a lengyel nyelvű válaszok adásában is.

A kérdőívbeli szituációk leírása: italautomatánál szeretnél pénzt váltani, akihez fordulsz: (a) egyetemi hallgató, (b) 5 évvel fiatalabb személy, (c) 10 évvel idősebb személy; illetve egy könyved kéred vissza (d) osztálytársadtól, (e) húgodtól.

(A szituációk megadásánál a szakirodalomban előzetesen vizsgált helyzetek közül [vásárlás előtt bevásárlókocsi használatához pénzváltás kérése és kölcsönadott tankönyv visszakérése, l. Szili 2004] a leggyakrabban előforduló szituációkat adtam meg, melyeket nem jellemzett szélsőséges kontextuális tényező.)

Az egyes szituációkban a szociopragmatikai (életkor, ismeretség foka, valamint státusz, ez utóbbi a 10 évvel idősebb beszédpartner esetére vonatkozik, több adatközlő

megadta, hogy egyetemen 10 évvel idősebb beszédpartner csak tanár lehet) és a beszédaktus-specifikus tényezők (a kérdés súlya, a szívesség és a jogosság mértéke) alapján a beszélő választhatott a külső és a belső módosítók eszköztárából, illetve ezen módosítók kombinációjával is élhetett.

A nyelvtanulók választásainak jelen vizsgálata a következőképp történt: első lépésként beazonosítottam, hogy a beszélő mely enyhítő belső és külső módosítókkal élt megnyilatkozásaiban. Ezt követően az első („a”, „b”, „c” szituációk) és a második kéréskörben („d” és „e” szituációkban) külön-külön figyeltem meg a beszélők által használt belső és külső módosítók adott mintázatát. Mintázatnak nevezem az egyes kéréskörökben adott módosítók adott rendszer szerinti jelenlétét a beszélők megnyilatkozásaiban (vö. Blum-Kulka–Olshtain 1984). Mindkét kéréskör első szituációja referenciaszituációnak tekinthető. Az első kéréskör esetén, ahol három szituáció szerepel, vagyis a referenciaszituációt két attól eltérő szituáció követ, azért döntöttem mindhárom szituáció egyszerre történő – mintázatként való – figyelembe vétele mellett, mert az adatközlő is a három szituációval ezen sorrendben „találkozott” a kérdőív kitöltése során, és maga is mindig az előző szituációhoz hasonlíthatta aktuális választát, attól térhetett el az új tényezőket figyelembe véve.

A tanulmány során a megfigyelt tipikus köztesnyelvi mintázatokat mutatom be. Tipikusnak tekintetem azon nyelvi formákat, illetve mintázatokat, amelyek 10%-os relatív gyakoriság felett fordultak elő.

4 Megfigyelések

Köztesnyelven a kérési megnyilatkozások tipikus enyhítő módosítói a következők:

- (i) külső módosítók: üdvözlés, elnézésekérés, felkészítő elem, magyarázat adása;
- (ii) belső módosítók: szintaktikai és morfoszintaktikai elemek: nem tegező formák, *tud* segédige, hatóige-képző, feltételes mód és a kérdő mondat.

Az első kéréskörben – a szívességkérés-közeli szituációkban – a következő módosítók fordultak elő a tipikus köztesnyelvi mintázatokban: az elnézésekérés, a nem tegező forma, a kérdő mondat és a *tud* segédige. Ezen szituációkban három fő tipikus mintázatot lehetett megkülönböztetni, mindegyiknek jellemzője, hogy az egyes szituációkban kérdő mondattal kér a beszélő, így ezt a jellegzetességet a továbbiakban nem emeltem ki.

Az 1. tipikus köztesnyelvi mintázat vonásai: az „a” és a „b” szituációban tegező formában áll az ige, és a kérést nem vezeti be elnézésekéréssel a beszélő, míg a „c” szituációban nem tegező formát használ, és elnézésekéréssel indítja megnyilatkozását (a mintázat előfordulási gyakoriságához l. az 1. táblázatot):

- [1] IL a,b: *Felválthatod a fémpénzre?*^x / c: *Bocsánat, felválthatja-e a fémpénzre?*
- [2] IL a,b: *Tudnál-e felváltani papír pénzt fémpénzre?* / c: *Elnézést, szeretném kérdezni, hogy tudna-e felváltani 10 zlotyt fémpénzre?*

^x A nyelvi példákat az adatközlők által megadott formában idézem. Az egyes példánál /-lel választom el az egyazon beszélő által a különböző szituációkban adott megnyilatkozásokat.

| | a | b | c | IL | L1(m) | L1(l) |
|-----|--------|--------|---------------------------|-----|-------|-------|
| I/1 | tegező | tegező | elnézéskérés + nem tegező | 10% | 8% | 13% |

1. táblázat. Az első kéréskör első tipikus köztesnyelvi mintázatának előfordulása a három beszélői csoportban

A 2. tipikus köztesnyelvi mintázat jellemzői: az „a” és „b” szituációban tegező, a „c” szituációban nem tegező formában áll az ige, továbbá az „a” és a „c” szituációban elnézéskéréssel vezet be kérését a beszélő. (Az előfordulási gyakoriságokhoz l. a 2. táblázatot.)

[3] IL a: *Bocsi, tudsz váltani nekem fémpénzre?* / b: *Tudsz váltani nekem fémpénzre?* / c: *Bocsánat, tud váltani papírpénzem fémpénzre?*

| | a | b | c | IL | L1(m) | L1(l) |
|-----|-----------------------|--------|---------------------------|-----|-------|-------|
| I/2 | elnézéskérés + tegező | tegező | elnézéskérés + nem tegező | 14% | 0% | 4% |

2. táblázat. Az első kéréskör második köztesnyelvi tipikus mintázatának előfordulása a három beszélői csoportban

Magyar L1-en nem volt erre a mintázatra példa, lengyel L1-en pedig nem tekinthető tipikusnak az előfordulása (4%). A „b” szituációban az „a” és a „c” szituációhoz képest nem alkalmazza az elnézéskérést a beszélő, valószínűleg a szituáció szociopragmatikai tényezőit figyelembe véve, itt az „a” referenciaszituációhoz képest 5 évvel fiatalabb a beszédpartner. Mivel a kérések általában komplex kérésstratégiának tekinthetők, külön meg lehet vizsgálni, nem kompenzálta-e a beszélő más enyhítő módosítóval az elnézéskérés elhagyását a „b” szituációban. Jelen vizsgálatban a köztesnyelvi beszélők 6%-a a „b” szituációban elhagyta az elnézéskérést, de üdvözölte beszédpartnerét („*Szia, ...*”), s ők az „a” és a „c” csoportban nem kezdték üdvözléssel kérésüket. Ilyenkor az üdvözlés a „b” szituációban enyhítő elemként szolgálhatott a kérésstratégia részeként. A 2. mintázatot felmutató többi produktumban az „a” referenciaszituációbeli megnyilatkozáshoz képest a „b” szituációban kompenzáció nélkül elmaradt az elnézéskérés.

A 3. tipikus köztesnyelvi mintázat jellemzői: az „a” és „b” szituációban tegező, a „c” szituációban nem tegező formában áll az ige, s a beszélő mindegyik szituációban elnézéskéréssel vezet be kérését. A 3/a. tipikus mintázat a 3. mintázat alelete, jellemzői a következők: a 3. mintázat jellemzőin túl mindhárom szituációban a *tud* segédige szerepel feltételes mód nélkül. A kontrollcsoportokban (magyar és lengyel L1) feltételes módban állt a *tud* segédige az egyéb tekintetben megegyező jellemzőket mutató mintázatban. (Az előfordulási gyakoriságokhoz l. a 3. táblázatot.)

A „c” szituációban, amikor 10 évvel idősebb személyt szólít meg a beszélő, a tipikus köztesnyelvi mintázatokban nem tegező formában áll az ige. A tipikus mintázatoktól eltekintve, a teljes vizsgált anyagról azt mondhatjuk, hogy a „c” szituációban a köztesnyelvi válaszolók 80%-ban fejezték ki kérésüket nem tegező formával, a lengyel L1 válaszolók 87%-ban, a magyar L1 adatközlők pedig 40%-ban (Sitkei 2017: 132). Részben emiatt alacsonyabb magyar L1-en a köztesnyelven megfigyelt tipikus mintázatok előfordulása.

[4] IL a,b: *Bocsi, tudsz pénzváltani?* / c: *Bocsánat, tud pénzváltani?*

| | a | b | c | IL | L1(m) | L1(l) |
|-------|---|---|---|-----|-------|-------|
| I/3 | elnézéskérés + tegező | elnézéskérés + tegező | elnézéskérés + nem tegező | 19% | 18% | 37% |
| I/3/a | elnézéskérés + <i>tud</i> segédige (tegező) | elnézéskérés + <i>tud</i> segédige (tegező) | elnézéskérés + <i>tud</i> segédige (nem tegező) | 12% | 0% | 0% |

3. táblázat. Az első kéréskör harmadik köztesnyelvi tipikus mintázatának előfordulása a három beszélői csoportban

A második kéréskörben, a jogos kérés szituációiban a következő tipikus mintázatok lehetett megfigyelni.

Az 1. tipikus köztesnyelvi mintázat jellemzői: a „d” szituációban kérdő mondatban fogalmazza meg kérését a beszélő, az „e” szituációban pedig felszólító módban (5. példa). Ezen mintázat alete az 1/a tipikus mintázat, amely során a „d” szituációban megjelenik a hatóige-képző (6. példa). (Az előfordulási gyakoriságokhoz l. a 4. táblázatot.)

[5] IL d: *Visszaadj a könyvemet?* / e: *Adj vissza a könyvemet!*

[6] IL d: *Adhatod vissza a könyvemet?* / e: *Add vissza a könyvemet!*

| | d | e | IL | L1(m) | L1(l) |
|--------|---------------------------------|-------------|-----|-------|-------|
| II/1 | kérdő mondat | felszólítás | 23% | 25% | 44% |
| II/1/a | hatóige-képző + kérdő mondat | felszólítás | 13% | 0% | 20% |

4. táblázat. A második kéréskör első köztesnyelvi tipikus mintázatának előfordulása a három beszélői csoportban

(Lengyel L1-en a *móc* segédige ragozott alakjai szerepeltek az 1/a mintázatban, ez a segédige lehetőséget, képességet, engedélyt és valószínűséget fejez ki [Swan 2002; Divjak–Szymor–Socha 2015], amely jelentések megtalálhatók a magyar hatóige-képző jelentései között is [Csonka 1998]).

A 2. tipikus köztesnyelvi mintázat jellemzői: mindkét szituációban felszólító módban áll az ige. Jelen mintázatban az enyhítő módosítók hiánya tekinthető mintázat-képző tényezőnek. (Az előfordulási gyakoriságokhoz l. az 5. táblázatot.)

[7] IL d, e: *Add vissza nekem a könyvemet.*

| | d | e | IL | L1(m) | L1(l) |
|------|-------------|-------------|-----|-------|-------|
| II/2 | felszólítás | felszólítás | 10% | 7% | 4% |

5. táblázat. A második kéréskör második köztesnyelvi tipikus mintázatának előfordulása a három beszélői csoportban

Egy kérés megnyilatkozás, amelyet egy adott szituációra közölt a beszélő, a változó egy halmazát öleli fel, pl. tartalmazhatja a *tud* segédigét, a hatóige-képzőt, az ige állhat nem tegező formában, a főcselekmény lehet kérdő mondat, elnézésekérés kapcsolódhat hozzá stb. Az első kéréskörben három szituáció kérés megnyilatkozásai tettek ki együttesen a mintázatot, a második kéréskörben két szituációé. Matematikai megközelítésben, az első kéréskör három szituációjának változóhalmazából kisebb

valószínűséggel jöhetnek létre tipikus, vagyis legalább 10%-os relatív gyakorisággal előforduló, azaz adott paramétereiben megegyező mintázatok a második kéréskör két szituációjához képest, ha a változók teljesen véletlenszerűen jellemeznék a kérésstratégiákat. A változók (vagyis a módosítók) a véletlenszerűséghez képest adott rendszer szerint kapcsolódnak a megnyilatkozásokhoz. Jelen vizsgálatban a mintázatok a beszélőknek ezen rendszerre vonatkozó – adott szociopragmatikai tényezők szerinti – döntéseit követik le, modellezik. Abból kiindulva, hogy az első kéréskörben három szituációból több tipikus mintázat volt megfigyelhető, s azok nagyobb lefedettséget mutattak a megnyilatkozások között (10% + 14% + 19%), mint a második kéréskör két szituációjának tipikus mintázatai (23% + 10%), az alábbi következtetésre juthatunk. Az első kéréskörben, vagyis szívességkéréskor a beszélőnek kevésbé van szabad mozgástere arra, hogy milyen módosítókat használhat kérése megfogalmazásakor adott szociopragmatikai tényezők mellett: a szívességkérő szituációkban jobban „előírt” a módosító adott halmazának használata. Ezekben a szituációkban a beszélők anyanyelvükön is konvencionális kifejezésekkel élnek (amely adott módosítóhalmazzal leírható), s köztesnyelven is igyekeznek a nyelvtanulók előhívni a célnyelvi konvencionális formákat.

További megfigyelésként érdemes megemlíteni, hogy köztesnyelven gyakoribb mindkét kéréskör esetén az azonos megnyilatkozásokkal való kérés – a különböző szituációkban – a változó szociopragmatikai tényezőktől függetlenül, míg az anyanyelvükön kérést kifejezők – az „a”, „b” szituációban alkalmazott kéréshez képest – a „c” szituációban mind lengyelül, mind magyarul minden esetben külön kérési stratégiával fogalmazták meg kérésüket, s ugyanígy a „d”, „e” szituációkban is. Az „a” referenciaszituációhoz képest a „b” szituációban 5 évvel fiatalabb személytől kér az egyetemista beszélő, s ekkor a magyar L1 adatközlők 68%-a, míg a lengyel L1 adatközlők 84%-a mind az „a”, mind a „b” szituációban teljesen megegyező kérési megnyilatkozással élt. Vagyis mindkét nyelvről elmondható, hogy a körülbelül 18-23 éves korosztályú magyar és lengyel anyanyelvű személyek anyanyelvükön 5 évvel fiatalabb beszédpartnerrel szemben általában nem változtatnak az alkalmazott udvariassági szinten, vagyis ilyen irányú és mértékű tényezőváltozásra kismértékben szenzitív a két nyelv. Azonos megnyilatkozásokat (l. a 6. táblázat) a köztesnyelvi adatközlők 43%-a adott az „a” és a „b” szituációban, míg 8%-uk mindhárom első körös szituációban, 5%-uk pedig a „d”, „e” szituációkban adott megegyező kérési formákat, vagyis ezen beszélők figyelmen kívül hagyták a szociopragmatikai tényezők változását.

| a | b | c | d | e |
|------------------------------------|---|---|--------|---|
| IL: 43%, L1(m): 68%, L1(l): 84% | | | | |
| IL: 8% | | | IL: 5% | |

6. táblázat. Az egyes szituációkban megegyező megnyilatkozást adók relatív gyakorisági értékei az egyes beszélői csoportokban

- [8] IL a, b, c: Tudsz váltani egy papír pénzt fémpénzre?
 [9] IL a, b, c: Fel tudsz váltani papír pénzemet fémpénzre?
 [10] IL a, b, c: Válthatsz azt fémpénzre?

Az „a” és a „b” szituációkban azonos választ adó anyanyelvi beszélők relatív gyakorisági értékei (magyar 68%, lengyel 84%) mellett 43%-ban fordult elő ez a jelenség

köztesnyelven, vagyis az anyanyelvi kérésekhez képest nagyobb gyakorisággal változtat a beszélő a kérésstratégiáján a referenciaszituációhoz képest fiatalabb beszédpartnertől való kéréskor.

5 A megfigyeléseken alapuló kutatási eredmények

Az első kéréskörben, a szívességkérő szituációkban több tipikus mintázatot lehetett felfedezni, s azok nagyobb mértékben lefedték a produktumokat a második kéréskör, a jogos kérések szituációihoz képest. Szívességkéréskor enyhítő módosítókkal igyekszik udvariasabbá tenni kérését a beszélő, s ezt leginkább konvencionális formákkal tudja elérni. A köztesnyelvi beszélők is törekedtek a célnyelvi konvencionális formák alkalmazására, sok esetben kevés választotta el őket annak megvalósításától, például a *tud* segédigét nem tették feltételes módú alakba (1. a 3/a mintázat).

Az első kéréskör 2. mintázata sajátos köztesnyelvi kommunikációs stratégiát mutat be, az 5 évvel fiatalabb beszédpartnertől való szívességkéréskor a köztesnyelvi beszélő kevesebb enyhítő módosítót alkalmaz, nem kezdi elnézés-kéréssel megnyilatkozását (illetve 6%-uk – annak kiváltására – nyelvtanulói szempontból egyszerűbb formát választ, a köszönést [*Szia*], 1. a Tarone-féle (1983) óvatos nyelvhasználati mód során a nyelvtanuló olyan nyelvi formával él, mellyel biztosnak érzi, hogy nem követ el hibát). Ez a stratégia a kontrollcsoportoknál egyedül a lengyel L1-ben fordult elő 4%-os relatív gyakorisággal. Kasper (1994: 3209) szavaival pozitív korreláció áll fenn a kontextuális tényezők, valamint azon befektetés között, hogy a megnyilatkozás udvarias legyen (1. 2.2-es pont). Máshogy fogalmazva a nyelvtanulói költség-haszon elv is magyarázhatja ezen mintázat előfordulását, amely szerint ha a szociopragmatikai tényezők alapján (például fiatalabb személlyel való interakció során) kevésbé szükséges udvariasan megnyilatkozni a nyelvtanulónak, azaz a megnyilatkozást udvariasná tevő plusz „költség”, nem eredményezne annak megfelelő „hasznot”, akkor a nyelvtanuló az előtte álló mintákhoz képest (esetünkben lengyel és magyar L1) kevesebb enyhítő módosítót alkalmaz, végeredményben kevésbé udvarias megnyilatkozással él. Az adatközlők a többi szituációban tudták a módosítókat használni, így nem indokolható nyelvi kompetenciabeli hiányossággal az enyhítő módosítók elhagyása.

További sajátos köztesnyelvi stratégia a szociopragmatikai tényezők figyelmen kívül hagyása, amikor a nyelvtanuló az adott kéréskörben változatlan módon fejezi ki kérését, nem törődve a referenciaszituációtól eltérő hallgatói paraméterekkel. Mindazonáltal a jelenség „ellentéte” is megfigyelhető, amikor az első kéréskörben az „a” és a „b” szituációk között az anyanyelvi beszélők (a magyar 68%, a lengyel 84%-os relatív gyakorisági értékkel) azonos kérésstratégiákkal viszik véghez a kérésüket, viszont köztesnyelven lecsökken ez az érték 43%-ra, vagyis az anyanyelvi kérésekhez képest nagyobb előfordulással változtat a beszélő a kérésstratégiáján a referenciaszituációhoz képest, amikor fiatalabb beszélőtől kér. Ez a megfigyelés részben a nyelvtanulói költség-haszon elvhez kapcsolódik, de sok esetben nem kevesebb módosítóval élt a beszélő, hanem változtatott a kérésstratégiáján, a referenciaszituációban használt módosító(k)tól eltérőt választott. Ez utóbbi jelenség szintén sajátos köztesnyelvi stratégiának tekinthető, mivel eltér a beszélő előtt álló (magyar és lengyel L1-beli) mintáktól.

6 Zárógondolatok, összegzés

A korábbi eredményekhez (Sitkei 2017) csatlakozva, ahol kérés-szituációnként vizsgáltam a köztesnyelvi kérésstratégiák jellemzőit, a jelen megfigyelések segítik megismerni a köztesnyelvi beszélők választásait. Továbbá azt, hogy a különböző szociopragmatikai tényezők hatására hogyan változtatják meg kérésstratégiájukat: a nyelvtanulói költség-haszon elv során fiatalabb beszédpartnerrel kevésbé udvariasan fejezi ki kérését a nyelvtanuló az előtte álló (anyanyelvi és célnyelvi) mintáktól függetlenül, emellett megfigyelhető volt a szociopragmatikai tényezők figyelmen kívül hagyása, illetve figyelembevétele olyankor, amikor cél- vagy anyanyelven nem teszik a beszélők. Ezen kutatási eredmények hozzájárulhatnak a sajátos köztesnyelvi kommunikációs stratégiák megismeréséhez segítve a MID-tanári munkát, de a köztesnyelvi vizsgálatoknak is alapját képezik, mert ezen információk birtokában lehet továbblépni és következtetéseket levonni.

A jelen kutatás során köztesnyelvi mintázatokat azonosítottam be, majd megnéztem, hogy mennyire egyeznek a magyar és a lengyel anyanyelvi mintázatokkal. Mivel a köztesnyelvi mintázatok nem feleltek meg teljesen a magyar L1 mintázatoknak, ezért további kutatási lehetőségként érdemes lenne megismerni a magyar anyanyelvi beszélők választásait, beazonosítani tipikus mintázataikat, hogy megismerjük adott szociopragmatikai tényezők alapján hogyan változtatják meg kérésstratégiájukat, ami hasznos lehet mind a magyar, mind specifikusan a magyar mint idegen nyelvi kutatások, illetve oktatás szempontjából is.

Köszönetnyilvánítás

Ezúton is szeretném megköszönni mindazok segítségét, akik valamely módon hozzájárultak a kutatásom megvalósításához. A köztesnyelvi válaszokat a 2016/2017-es őszi tanévben a krakkói, a poznańi és a varsói magyar szak hallgatói, valamint a budapesti Balassi Intézet és az ELTE MID tanszékén tanuló hallgatók adták. A magyar anyanyelvi válaszokat a SZIE GTK és az ELTE BTK hallgatói, a lengyel anyanyelvi válaszokat a poznańi magyar szak és a Varsói Egyetem hallgatói adták.

Irodalom

- Alcón-Soler, E., Safont-Jordà M. P., Martínez-Flor A. 2005. Toward a typology of modifiers for the speech act of requesting: a socio-pragmatic approach. *Revista Electrónica de Lingüística Aplicada*, 4: 1–35.
- Babbie, E. R. 1998. *The practice of social research* (8th ed.). Belmont, CA: Wadsworth.
- Balogh K. 2002. A nyelvészeti udvariasság-elméletek jelentősége a nyelvoktatás számára. *BGF Tudományos évkönyv 2002*. 336–342.
- Bañcerowski J. 1999. A nyelvi jelentés és az emberi tudás. *Magyar Nyelvőr*, 123(2): 214–217.
- Becker, J. D. 1975. The Phrasal Lexicon. In: Shank, R., Nash-Webber, B. L. (szerk.): *Theoretical issues in natural language processing*. Cambridge, MA: Bolt Beranek & Newman. 60–63.
- Beebe, L. M., Cummings M. C. 1996. Natural speech act versus written questionnaire data: How data collection method affects speech act performance. In: Gass, S. M., Neu J. (szerk.): *Speech Acts across Cultures: Challenges to communication in a second language*. Berlin: Mouton de Gruyter. 65–86.

- Blum-Kulka, S. 1987. Indirectness and politeness in requests: same or different? *Journal of Pragmatics*, 11: 131–146.
- Blum-Kulka, S. 1991. Interlanguage pragmatics: The case of requests. In: Phillipson, R., Kellerman E., Selinker L., Sharwood Smith M., Swain M. (szerk.): *Foreign/Second Language Pedagogy Research*. Clevedon, UK: Multilingual Matters. 255–272.
- Blum-Kulka, S., House J. 1989. Cross-Cultural and Situational Variation in Requestive Behavior. In: Blum-Kulka, S., House J., Kasper G. (szerk.): *Cross-cultural pragmatics*. Norwood, NJ: Ablex. 123–154.
- Blum-Kulka, S., Olshtain E. 1984. Requests and apologies: A cross-cultural study of speech act realization patterns (CCSARP). *Applied Linguistics*, 5(3): 196–213.
- Boxer, D. 1996. Ethnographic interviewing as a research tool in speech act analysis: the case of complaints. In: Gass, S. M., Neu J. (szerk.): *Speech act across cultures: Challenges to communication in a second language*. Berlin: Mouton de Gruyter. 217–240.
- Clark, H. H., Schunk D. H. 1980. Polite responses to polite requests. *Cognition*, 8(2): 111–143.
- Copestake, A., Terkourafi M. 2010. Conventional speech act formulae: from corpus findings to formalization. In: Kühnlein, P., Benz A., Sidner C. L. (szerk.): *Constraints in Discourse 2*. 125–140.
- Corder, S. P. 1992. A Role for the Mother Tongue. In: Gass, S. M., Selinker L. (szerk.): *Language Transfer in Language Learning*. Amsterdam, Philadelphia: John Benjamins Publishing Company. 18–31.
- Csonka Cs. 1998. A *-hat/-het* képzőről. *Intézeti Szemle*, 20(1): 5–13.
- Divjak, D., Szymor N., Socha-Michalik A. 2015. Less is more: possibility and necessity as centres of gravity in a usage-based classification of core modals in Polish. *Russian Linguistics*, 39(3): 327–349.
- Ferguson, C. A. 1976. The Structure and Use of Politeness Formulas. *Language in Society*, 5(2): 137–151.
- Fraser, B. 1978. Acquiring social competence in a second language. *RELC Journal*, 9: 1–21.
- Fraser, B., Nolan W. 1981. The association of deference with linguistic form. In: Walters, J. (szerk.): *The Sociolinguistics of Deference & Politeness*. The Hague: Mouton. 93–111.
- Goffman, E. 1953. *Communication Conduct in an Island Community*. PhD Dissertation. Chicago: University of Chicago.
- Goffman, E. 1967. *Interaction ritual: essays on face to face behavior*. Chicago: Aldine Publishing Company.
- Gordon, D., Lakoff G. 1975. Conversational Postulates. In: Cole, P., Morgan J. L. (szerk.): *Syntax and Semantics 3: Speech acts*. New York: Academic Press. 83–106.
- Holtgraves, T., Yang J.-N. 1992. Interpersonal underpinnings of request strategies: General principles and differences due to culture and gender. *Journal of Personality and Social Psychology*, 62: 246–256.
- House, J., Kasper G. 1987. Interlanguage pragmatics: Requesting in a foreign language. In: Lörscher, W., Schultze R. (szerk.) *Perspectives on language in performance. Festschrift für Werner Hüllen*. Tübingen: Narr Verlag. 1250–1288.
- Hymes, D. H. 1968. The ethnography of speaking. In: Fishman, J. A. (szerk.): *Readings in the sociology of language*. The Hague, Paris: Mouton. 99–138.
- Kasper, G. 1994. Politeness. In: Asher, R. E., Simpson J. M. Y. (szerk.): *The Encyclopedia of Language and Linguistics*. Oxford: Pergamon Press. 3206–3211.
- Kasper, G. 2000. Data collection in pragmatics research. In: Spencer-Oatey, H. (szerk.): *Culturally Speaking. Managing Rapport through Talk across Cultures*. London, New York: Continuum. 316–341.
- Lakoff, R. T. 1973. The logic of politeness: or, minding your p's and q's. In: Corum, C., Smith-Stark C. T., Weiser A. (szerk.): *Proceedings of the ninth regional meeting of the Chicago Linguistic Society*. Chicago: University of Chicago Press. 292–305.
- Leech, G. N. 1983. *Principles of Pragmatics*. London: Longman.
- Manes, J., Wolfson N. 1981. The compliment formula. In: Coulmas, F. (szerk.): *Conversational Routine: Explorations in Standardized Communication Situations and Prepatterned Speech*. The Hague: Mouton Publishers. 116–132.
- Reichenbach, H. 1947. *Elements of symbolic logic*. New York: The Macmillan Company.
- Rose, K. R., Ng Kwai-Fun C. 2001. Inductive and deductive teaching of compliments and compliment responses. In: Rose, K. R., Kasper G. (szerk.): *Pragmatics in language teaching*. Cambridge: Cambridge University Press. 145–170.
- Selinker, L. 1972. Interlanguage. *International Review of Applied Linguistics*, 10: 209–231.

- Sitkei D. 2017. A kérés beszédaktusáról magyarul tanuló lengyel hallgatók köztesnyelvében. In: Koutny, I., Piotrowiak-Junkiért K., Kornatowski P., Németh Sz. (szerk.): *A hungarológia ma: nyelvészet, kultúra és oktatás. Hungarologia dzisiaj: językoznawstwo, kultura i dydaktyka*. Poznań: Rys Kiadó. 129–139.
- Suh, J.-S. 1999. Pragmatic perception of politeness in requests by Korean learners of English as a second language. *IRAL*, 37(3): 195–214.
- Swan, O. E. 2002: *A Grammar of Contemporary Polish*. Bloomington, IN: Slavica.
- Szili K. 2004. *Tetté vált szavak. A beszédaktusok elmélete és gyakorlata*. Budapest: Tinta.
- Tarone, E. 1983. On the Variability of Interlanguage Systems. *Applied Linguistics*, 15: 285–295.
- Terkourafi, M. 2015. Conventionalization: A new agenda for im/politeness research. *Journal of Pragmatics*, 86: 11–18.
- Tracy, K., Craig R. T., Smith M., Spisak, F. 1984. The Discourse of Requests: Assessment of a Compliance-Gaining Approach. *Human Communication Research*, 10(4): 513–538.
- Trosborg, A. 1995. *Interlanguage pragmatics. Requests, Complaints and Apologies*. Berlin: Mouton de Gruyter.
- Weizman, E. 1989. Requestive Hints. In: Blum-Kulka, S., House J., Kasper G. (szerk.): *Cross-cultural pragmatics*. Norwood, NJ: Ablex. 71–95.

Linguistic realizations of research article abstracts written by Ecuadorian and North American academic authors

Rodrigo Tovar-Viera

University of Szeged, Hungary
Technical University of Cotopaxi, Ecuador
victov912@hotmail.com

Abstract: The way in which article abstracts introduce their research has increasingly become the subject of great interest among researchers and scholars. The abstract, indeed, after the title, is the most read section freely available online. The study investigates the linguistic features, such as verb tenses, clauses, modals and stance markers that categorize abstracts and abstracts moves in North American and Ecuadorian journals. The genre analysis of English and Spanish texts is based on a corpus of 120 abstracts from the fields of humanities and sciences. Results show that present and past tense, and hedges are the most frequent categories. Present tense predominantly occurs in M1 and M2, and sometimes in M5. The usage difference concerning the linguistic realization of abstracts moves might be due to the nativeness or non-nativeness of authors however, such rhetorical and style choices remain quite unclear, in terms of conventional patterns or authors' preferences.

1 Introduction

The way in which abstracts present their research articles has led to extensive attention being paid to the rhetorical and writing style across disciplines and languages. Indeed, much research recognizes the importance of examining how research article abstracts (RAAs) written by native and non-native authors textually organize the language when communicating the gist of the whole paper (e.g. Çakir 2015, Can et al. 2016, Fallatah 2016; Fan and Song 2017, Kosasik 2018, Lee 2017). Although abstracts significantly vary in the organization of the texts namely, in the rhetorical and writing style as well as language use, they have to present faithful and accurate information. In fact, much credit and merit is given nowadays to those RAAs that provide readers with a clear picture of the information content. In other words, they function as “screening devices”, helping readers decide whether to read the complete article or not (Huckin 2001:93). An abstract, therefore, is the brief description and the pure reflection of the entire article (Bhatia 1993). Abstracts, particularly RAAs, are quite important because they constitute the first mini-texts that academics will encounter freely available online.

Ludányi Zsófia, Krepsz Valéria, Grácsi Tekla Etelka (szerk.): *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2018*. Budapest: MTA Nyelvtudományi Intézet. 2018.

However, writing publishable RAs that include accurate abstracts is not easy. It is because writing like any other advanced skill requires the pre-existing knowledge of the writing system, as formats, vocabulary and sentence structure of those second or foreign languages to be successfully acquired. In this process, the effective selection of the rhetorical and linguistic elements to persuade readers determines the readability of the written texts (Ren and Li 2011). In effect, according to the writing style, texts may be categorized as sources that contain good or vague information. In academic writing therefore abstracts play an essential role in disseminating the new knowledge, particularly the way of how writers show readers that the paper is worth reading and deserves their attention (Hyland 2000). What is more, scholarly journals, publishing articles written either in English or in other languages, require abstracts accompanying the original research studies (Martin 2003:27). From this point of view, abstracts create not only a road map of the full paper, but also are the key section of the article, which, according to Salager-Meyer (1992), are read more receptively by scholarly and non-scholarly readers, either national or international.

Abstracts then in the globalized world of scientific communication and information explosion are the gateways to knowledge production (Swales and Feak 2009). Thus, non-English medium journals, under pressure to disseminate the research findings and hook an international audience, include an English version of the original RAAs (Martin-Martin 2005). Despite the fact that the English version of the RAAs is a great chance for non-native authors to disseminate their articles internationally, the English translated version seems not to be of much interest for many writers, indeed, they perceive it as a mere requirement set by the journal editors (Lorés 2014:87). Accurate abstracts, in fact, draw readers' attention and increases opportunities of indexing and citation. Nevertheless, some writers are not aware of the potential impact that abstracts have when translating the scope of the paper into the foreign language version (Lorés 2014). Underlying the relevance of rhetorical and writing style in academic writing, a great number of the RAs summarize their works by emphasizing the purpose, methodology, results and conclusions (Hyland 2000). The present contrastive study examines the linguistic realizations of 'moves' —discoursal or rhetorical units to determine the communicative purpose of the texts (cf. Swales 1990, 2004) in RAAs published in North American and Ecuadorian journals. It explores the linguistic features, namely verb tenses, clauses, modals, and stance markers, that characterize the RAAs written in the disciplines of education, sociology, electronics and agronomy in English and Spanish languages.

1.1 Theoretical background

In the genre analysis of written texts, research studies have increasingly attempted to show how academic writers intervene in their texts and build solidarity with their readers (Hyland and Tse 2005:124). The analysis of texts then, particularly research articles (RAs), has gone beyond the structural and linguistic examinations aiming to understand the context in which the genre is produced. Much research therefore has been carried out on how language is textually organized in different sections of RAs, such as the introduction (Swales, 1990, 2004), results (Yang & Allison, 2003), methodology (Lim, 2006) and abstracts (e.g. Pho, 2008; Işık-Taş, 2018). This research interest, in the last two decades, has led to the proliferation of considerable studies in the genre analysis of RAAs published in particular disciplines, for example

linguistics, applied linguistics, education, literature, law, experimental science, sociology, among others (e.g. Doró 2013, Kafes 2012, Lau 2004, Samraj 2002, Suntara and Usaha 2013, Ghasempour and Farnia 2017, Loutayf 2017, Ebrahimi and Motlagh 2015, Hyland 2015) and genres (Bordet 2014, Hyland 2000, Kosasik 2018, Sayfour, 2010) across languages (e.g. Behnam and Golpour 2014, Divason and León 2006, Lorés, 2004, Martin-Martin 2003, van Bonn and Swales 2007).

As a result of the above studies, research outcomes show that abstracts may vary, in terms of rhetorical organization and how language is produced in the surrounding context of publication. For instance, Martin-Martin (2003), Lorés (2004) and Loutayf (2017) found that RAAs written in Spanish traditionally reflect the introduction, method, result and discussion (IMRD) conventional structure, but with variation in the frequent occurrence of result moves throughout the abstracts. Similarly, Behnam & Golpour (2014), notice that abstracts written in English and Persian follow the four-move structure, as purpose, method, result and conclusion (PMRC). This result is in line with those of Kafes (2012), Suntara and Usaha (2013), Samraj (2002) and Dos Santos (1996) where introduction, unlike conclusion, was optional. Nonetheless, Lau (2004) notices that abstracts written by expert authors showed a frequent occurrence of a five-move pattern compared to the abstracts written by PhD student authors.

Kafes (2015), in contrast, found that Turkish abstracts hold their local discursive practices, but follow the international English writing conventions to construct the rhetorical structures, reporting a three-move conventional pattern such as PMR. Çakır (2015), similarly reveals that the majority of Turkish and English RAAs adopt the three-move schema, however, Turkish abstracts describe the research purpose without referring to the place, whereas English ones define it when announcing the purpose of the study. These findings are in connection with those of Doró (2013), Hyland (2000), Kafes (2012), Pho (2008), Fallatah's (2016), Lorés (2014), Farzannia and Farnia (2017), Al-khaswneh (2017), Kosasih (2018) in which introduction and conclusion were optional, except in Farzannia and Farnia's study, and Fallatah's international RA abstracts, where conclusion had 90% of occurrence. This rhetorical divergence may be connected to writers' different local practices and international discourse conventions.

Taking the impact of the above descriptive framework, academics have turned their interest to analyzing the linguistic features that constitute those RAAs written in different disciplines and genres. Mur-Dueñas (2011) thus investigates the linguistic and cultural differences of research articles published in English and Spanish journals. Topicalizers and endophoric markers were more common in the Spanish articles than in the English ones. Nonetheless, English texts showed a strong use of interactional meta-discourse features to balance the writer–reader positioning as the agent or owner of knowledge. In the same way, Loutayf (2017) found that the majority of abstracts written by Argentinean authors in five disciplines adopt the impersonal style (70%) when communicating the content of the entire article. It seems that Argentinean writers hold a secondary role to avoid negative criticism within the academy. Similarly, Poveda (2007) notices that modal verbs occurred in English abstracts, and adverbial expression and subjectivity in Spanish ones, nonetheless, both types of texts used hedging strategies, such as depersonalization, modesty, solidarity, deference, tentativeness and justification to control the level of precision when presenting the information. Writers then show objectivity rather than authority or arrogance.

Tseng (2011), examining the linguistic realization of abstracts moves written in English discovered, that the present tense frequently occurred in M1, M2 and M5, and

past tense sometimes appeared in the moves: M3 and M4. Nonetheless, native texts used the present tense more than their nonnative counterparts did. The findings also revealed that the present perfect was used in almost over half percentage of the abstracts. Tseng's findings are in line with those of Salager-Meyer (1992), observing that verb tenses perform different functions in different sections of the texts and their moves, for instance, the present tense enhances and emphasizes the discussion in the conclusion, recommendation and synthesis of the data. In contrast, the past tense was commonly used in purpose, methods, results and case presentation moves. Ghasempour & Farnia (2017) and Kosasik (2018), meanwhile, found that present tense and past tense frequently occurred in the five moves identified. Likewise, Pho (2008) reports linguistic variation, in terms of verb tenses. Present tense and present perfect were the most frequent categories across the abstracts; however, their occurrence varied according to the context of publication. Pho also notices that writers have the tendency to open the scope of work by mentioning previous researchers or studies in the field, that is, the reference subject frequently occurred in the introduction move. Similarly, Ghasempour & Farnia (2017), Kosasik (2018) and Pho's studies, Lorés (2014) found dissimilar lexico-grammatical patterns in English and Spanish abstracts. Thus, active voice with textual nouns as subjects occurred in both languages; nevertheless, Spanish texts favor the use of the passive form 'se', which is used to refer to the process without mentioning the agent. These results show the linguistic differences between languages, in terms of the writing style.

Hu and Cao (2011), investigating the stance markers in RAAs published in English and Chinese-medium journals, found that hedges were most frequent in English texts and boosters in Chinese ones. This study is in line with the results of Divasson & Leon (2006), in which boosters were frequently used in non-native texts. Hu and Cao conclude that writers whose first language is not English sometimes show preference for using boosters when presenting their claims and arguments. On the other hand, Ebrahimi & Motlagh (2015), studying the textual devices and discourse functions in RAAs written by native and non-native writers of English, notice that the choice of textual devices, frequencies and discourse functions, to some extent respond to particular discursive practices and conventions across disciplines. This stylistic variation in Swales and Feak's (2012) view is affected by the genre, discipline and discourse community practices, which somehow are closely connected to the audience, purpose, organization, and style. For example, in Ebrahimi & Motlagh (2005), native authors tend to employ more contrastive textual devices than non-native ones to validate their research merits and to keep the readers' attention on their findings. Alonso (2014), on the other hand, revealed that evidential lexical items were common in law and medicine abstracts written in English and Spanish while epistemic modals were used in the English law and medicine texts.

The rhetorical and linguistic differences discussed above may be attributed to different linguistic and cultural conventions among writers (Connor 2004). Indeed, when authors position their claims about what others have said and what they actually say, they seem to have their own preferred rhetorical and textual strategies. Following Connor's view, Hyland and Tse (2005) highlight that written texts contain rhetorical patterns or moves, which are essential to analyze the rhetorical and textual organization of particular type of texts. In this vein, van Bonn and Swales (2007) argue that such rhetorical patterns are valuable sources to identify not only the socio-cultural influence, but also the discursive practices employed in different scientific communities across languages. Taking abstracts as the corpus for the analysis, the aim

of this study is to investigate the linguistic features —verb tenses, clauses, modals and stance markers— that categorize abstracts and abstract moves written in the fields of humanities —education, sociology— and natural science —electronics and agronomy— published in North-American and Ecuadorian journals.

1.2 Research Questions

In order to portray a clearer picture of the linguistic features used in English and Spanish RAAs, the present study addresses the following three research questions:

- (i) What are the most frequent linguistic features in articles abstracts published in the two journal types?
- (ii) What are the linguistic features that characterize each abstract move in the fields of humanities and science in both languages?
- (iii) Are there any significant differences in the linguistic realizations of moves across the four disciplines?

2 Methodology

Sentences are the analytical unit for the linguistic realizations of moves. Nevertheless, given the fact that abstracts are condensed texts, and a move pattern may likely be within one or more sentences (e.g. Lau, 2004; Pho, 2008), linguistic signals, for instance, *in a recent work...*, *the present study examined...*, *the purpose of this investigation is...*, *the article analyzes data from...*, *the data for this research...*, *the findings reveal...*, *the results of the study suggest...*, *the article concludes...*, *within this context, differences ...*, *in this project, we...* are used to identify the communicative goals of each move and its textual organization. With this in mind, the analysis focuses on (a) verb choice, namely present-past tense, perfect and passive structures, (b) clauses such as coordinated and subordinated, (c) modals, particularly auxiliary and non-auxiliary, and (d) stance markers, as hedges and boosters.

Following Tseng' (2011) procedure, the move analysis consisted of three steps: thus, if a move was realized in a sentence, then the main verb of that sentence determined the verb tense of such move, as in examples 1 and 2. Then, if a move was realized in a clause or a phrase, for instance, the main clause ruled the sentence structure from the analysis, such as in examples 3 and 4. Finally, if a move was embodied by several sentences, which had different linguistic realizations, then the features, namely, verb tense, voice and modals were counted as individually ones, see examples 5 and 6. The researcher, namely, human coder supports the reliability of the coding of this procedure, which consists of a double round of coding with an interval in between. Then the coding was compared to see to what extent this match, and to establish the final results.

[1] *The control of atmospheric emissions is relevant in fragile ecosystems as the Galapagos Islands* (M1). Present [Eng. electronics – NNS]

[2] *El propósito de esta investigación ha sido recoger datos sobre la aplicabilidad de la concepción simple de la lectura a la lectura del español.* (M2) Pretérito perfecto /perfect/ [Span. education – NSS]

- [3] *Even though the pesticide was tested in two different crops, it **needs** a further research to see its effects on the product.* (M5) Subordinate and present [Eng. agronomy – NS]
- [4] *Using canonical correlation analysis with 475 white Catholic and Protestant students, results **showed** religious beliefs and white privilege attitudes...* (M4). Subordinate and present [Eng. sociology – NS]
- [5] *La tecnología aplicada al desarrollo de prótesis **ha experimentado** grandes avances desde los diseños hasta la manufactura y sistemas de control. La tendencia **es** a crear diseños que permitan al amputado tener más grados de libertad de lo que una prótesis rígida normalmente ofrece* (M1). Pretérito perfecto /perfect/ and presente /present/ [Eng. electronics – NSS].
- [6] *The results **showed** that student teachers' learning from experiences is a process involving many interrelated personal and social aspects. This process **could include** past and present experiences gained in multiple situations and contexts over time. Four chains of activities and experiences **were distinguished** in their learning processes.* (M4) Past, non-auxiliary verb and passive [Span. education – NS]

Once each move was clearly set down and differentiated from one to another, the study quantifies the frequency of occurrence of the above-mentioned linguistic features in order to analyze to what extent native and non-native RAAs show writers' frequent preferences for these linguistic features when constructing the rhetorical moves of the abstracts. The data analysis includes the most and the least frequent features in both languages across the four disciplines.

2.1 Corpus

The study embraces the analysis of 120 RA abstracts published in North American and Ecuadorian peer-reviewed online/print journals in the fields of humanities and sciences. Eighty English abstracts written by native English and Spanish speakers of the language, and forty Spanish texts written by native Spanish speakers published in the periods of 2010–2016 constitute the sample of the research study. The native and non-native authorship, as authors' names and affiliation, were verified based on the bio-data section presented at the beginning of the article of each journal. Nonetheless, as many speakers from different countries speaking different languages may use the same names usually common in English-speaking countries (Sayfour 2010), in the current study, the nativeness or non-nativeness of the context in which the related journal/article was published is the matter and not the nativeness or non-nativeness of the author. Thus, the genre analysis focuses on the publication content of the abstract rather than on the writers themselves.

Although the linguistic realizations of RAAs, for instance, in the disciplines of linguistics, applied linguistics and literature have been widely investigated, the aforementioned disciplines, as far as I know, are scarce in Ecuadorian journals compared to education, sociology, electronics and agronomy. Moreover, there are few contrastive studies carried out in Ecuador, particularly in the genre analysis of abstracts between Ecuadorian and North American journals. With this in mind, the

corpus of the Spanish RAAs is composed of 6,886 tokens, with 6,179 ones in the English translated version and 6,303 running words in the English L1 RAAs. The corpus was distributed across the four disciplines, for instance, 1,850 words in education, 1,331 in sociology, 1,533 in electronics and 2,172 ones in agronomy constitute the Spanish sub-corpora. Additionally, 1,688 tokens in education, 1,284 in sociology, 1,298 in electronics and 1909 ones in agronomy form the sub-corpora of the English version of RAAs published in Ecuadorian journals. On the other hand, 1,574 running words, 1,371 in sociology, 1,357 in electronics and 2001 ones in agronomy are the sub-corpora of English RAAs published in North American journals.

The data set includes 40 English article abstracts published in North American journals, such as the *American Journal of Education (AJE)*, *Journal of Teacher Education (JTE)*, *American Journal of Sociology (AJS)*, *Sociological Perspectives*, *Journal of Electronic Materials (JEM)*, *Journal of Electronic Packaging (JEP)*, *Journal of Agricultural and Food Chemistry (JAFC)* and *Journal of Agricultural Science (JAS)*. Each discipline such as sociology, education, electronics and agronomy encompasses 10 abstracts. Similarly, the Ecuadorian corpus consists of 40 English article abstracts and their original Spanish versions (40) published in Ecuador in the following journals: *INNOVA Research Journal*, *Alteridad*, *Axioma*, *UTCiencia*, *Revista Tecnológica ESPOL*, *Amazonica*, *Analitika*, *Procesos*, *Perfiles*, *Avances*, *Ingenius and*, *La Granja*. All four disciplines, namely sociology, education, electronics and agronomy, comprise of 10 abstracts each.

The selection criteria used in choosing the journals were their indexation. Both Ecuadorian and North American journals are indexed either in a regional or international medium contexts. Ecuadorian journals, for instance, are indexed in latindex (regional cooperative online information system for scholarly journals from Latin America, the Caribbean, Spain and Portugal), as well as in *elsevier*, *dialnet*, *GoogleScholar*, *REDIB*, and *e-rivist@s*. These journals are hosted by Ecuadorian Universities. North American journals are indexed in *ERIC*, *EBSCOhost*, *elsevier*, *American statistical association*, *SciSearch*, *SCOPUS*, *Applied Science and Technology*, and *GoogleScholar*. Unlike North American journals, Ecuadorian ones are mixed journals, either online or print, which devote special sections and space to the aforementioned disciplines.

3 Results and Discussions

The contrastive study of the genre article abstracts has revealed that there are significant differences derived from the lexico-grammatical choices in the three linguistic and cultural practices, namely, L1-English, EFL-English, and L1- Spanish. The table 1 below shows the frequent use of each linguistic category in the abstracts accompanying the articles. It can be seen that present tense is one of the most frequent linguistic features used in English and Spanish abstract published in native and non-native-medium contexts. Hedges and past tense are reported to be the second and third highest features used in the majority of the RA abstracts, either in Spanish or in English. Unlike hedges and past tense, boosters and subordinate clauses are the fourth and fifth highest frequent linguistic category used among article abstracts written in both languages. Perfect tenses, coordinate clauses, non-auxiliary and passive

structures are the least frequent linguistic elements throughout the Spanish and English sub-corpora.

The research outcomes of the present study are in line with those obtained by Tseng (2011), and Hu and Cao (2011). Tseng, on the one hand, found that present and past tense were the most frequent linguistic categories in almost over the majority of 90 RA abstracts. Hu and Cao, on the other hand, report that hedges were more frequent in English texts than in their non-native English counterparts. Although in the current research the past tense is the third most frequent structure compared to Hu and Cao's, and Tseng's rhetoric and linguistic realization of verb tense, the findings further support its common use in RAAs. That is, in the construction of abstracts, native and non-native abstracts written in English and Spanish show the frequent use of such a linguistic feature. Thus, present and past tense, and hedges were the typical pattern used in both descriptive and informative article abstracts when introducing the scope of the research papers. In effect, the genre analysis reports that from 120 RAAs analyzed, 248 times (26%) the present tense was used in the texts, distributed across native and non-native English and Spanish written corpora. Additionally, past tense was used 124 times (13%), and hedges occurred 155 times (16%) in the RA abstracts. Moreover, boosters significantly appear 107 times (11%) in the whole corpora.

| Ling. Features | English corpus. | | | | Spanish corpus | | |
|----------------------|-----------------|----|------|----|----------------|----|-------|
| | NES | % | NNES | % | SPaS | % | Occur |
| Present tense | 90 | 25 | 77 | 21 | 81 | 33 | 248 |
| Past tense | 37 | 10 | 38 | 10 | 49 | 20 | 124 |
| Perfect tense | 11 | 3 | 9 | 2 | 9 | 4 | 29 |
| Passive voice | 33 | 9 | 37 | 10 | 7 | 3 | 77 |
| Coordinate | 18 | 5 | 12 | 3 | 11 | 5 | 41 |
| Subordinate | 38 | 10 | 52 | 14 | 13 | 5 | 103 |
| Auxiliary | 29 | 8 | 20 | 5 | 1 | - | 50 |
| Non-auxiliary | 14 | 4 | 27 | 7 | 3 | 1 | 44 |
| Hedges | 53 | 15 | 53 | 14 | 49 | 20 | 155 |
| Boosters | 41 | 11 | 47 | 13 | 19 | 8 | 107 |
| TOTAL | 364 | | 372 | | 243 | | |

Table 1. Linguistic features in English and Spanish abstracts

All the linguistic features mentioned above are more frequent in L1 English than in Spanish and in the English translated abstracts. Yet, the English translated version from Spanish outperformed the total use of the linguistic features in the sub-corpora. The parallel analysis of abstracts written in English by non-native users of the language (NNES) and Spanish ones written by native speakers of Spanish (SPaS) shows that Spanish texts significantly varied the occurrence of linguistic features when presenting the gist of the original RAAs in the foreign language translated version (see Table 1 above). For instance, whilst present tense in the Spanish sub-corpora occurs 81 times (33%), in the English one it does 77 times (21%). Similarly, Spanish abstracts outperformed English texts in the frequent use of the past tense with approximately 49 times (20%) of occurrence. Passive structures, subordinate clauses and boosters tend to be more common in English RA abstracts than in Spanish ones. Hedges tend to be more frequent in English (13%) than in Spanish (8%) texts. Auxiliary verbs were the least frequent linguistic feature in Spanish texts while

perfect tenses rarely occur in NNES and SPaS abstracts. This linguistic variability would seem to be due to the fact that Spanish authors have to adapt their texts to the English discourse conventions and writing system.

The data analysis of the frequency of the verb tense in each move shows that the present tense was the preferred tense in Move 1 (introductions), Move 2 (purposes), and M4 (results), and that the past tense was used more frequently in Move 3 (methods) in abstracts written in Spanish. Among the 80 abstracts analyzed (40 in English and 40 in Spanish), M2 sections show the highest occurrence of the present tense in Spanish and in the English translated version. Passive voice, in contrast, frequently occurs in M3 in the English texts translated from Spanish. The further analysis revealed that Spanish abstracts and their English translated versions seem to be less informative and less organized compared to their L1 English counterparts. However, despite the fact that English and Spanish languages, linguistically and culturally, in terms of linguistic base and language identity are quite different, these linguistic and cultural differences do not mean that Spanish speakers are not able to produce informative and well-structured information as natives do. The discrepancy arises when first language (L1) features are translated into foreign language writing to communicate the gist of the original paper. For example, the below excerpts, (7-9) discuss the lop-sided punctuation and mechanics, wordiness and lack of concreteness throughout the texts.

- [7] *'[...] impunemente, a través de los micromachismos (mM), que es el término con el que se conoce a las cotidianas agresiones machistas de baja intensidad...'* (SPaS)
 /...[...], impunity a Through the (mM) **micromachismos, which is the term that is known to everyday...that is known to everyday Low Intensity **macho** aggression.../ M1 (Sociology, English translated version).
- [8] *'La habilidad empresarial fué establecida como una variable dependiente, y como variables independientes...'* (SPaS).
 /...The entrepreneurship **ability was established as a dependent variable, and as **independants** variables.../ M3 (Sociology, English translated version).
- [9] *'Este artículo explora las razones por las cuales las manifestaciones de violencia simbólica son socialmente toleradas y reproducidas, en la mayoría de los casos...'* (SPaS)
 /This article explores the reasons why the manifestations of symbolic violence **son and reproduced socially tolerated in the majority of cases.../ M1(Sociology, English translated version)

The examples above alert writers, particularly those to whom English is not their first language, so that they pay more attention when constructing the information content of any type of text, mainly those of RAAs. In the sense that a text poorly written in any language remains poor in its recipient/target language after being translated unless the translator re-writes the whole text (Wallwork, 2016). In other words, writers are virtually 100% responsible for readers understand or not understanding the text. Then, in order to make reading a pleasant experience, writers should include clear and concrete information to introduce the gist of the complete article. Otherwise, it affects the comprehension and interpretation of the information content. As a result, the content may be judged in terms of concreteness and

wordiness, in the sense that such information is neither informative nor descriptive text (Tovar, 2016).

The non-parallel linguistic analysis between abstracts written in English by native and non-native academic authors reported that boosters (13%) and subordinates (14%) appear more often in the RAAs published in non-native-medium journals than in those of native ones (11% vs. 10%, respectively). Nonetheless, native English texts exceeded non-native ones in the frequent use of the present tense (see Figure 1 below). Past tense and hedges appear to have similar frequency of occurrence in both the native and non-native English sub-corpora. The results are consistent with Hu and Cao (2011), who noticed that hedges were frequently used in English texts. The genre-related difference in this study is that boosters markedly occurred in English abstracts written by Ecuadorian speakers, whereas in Hu and Cao (2011), boosters frequently appeared in Chinese texts. The lexico-grammatical choices between native and non-native English abstracts shed lights on to assume that non-native writers adapt their texts to the rhetorical and international English writing style. Despite the fact that non-native RAAs, particularly those of Ecuadorian authors, follow English writing conventions, a further analysis suggests that there are stylistic aspects such as word choice that needs to be polished when translating from one language into another (see the examples a, b, c above).

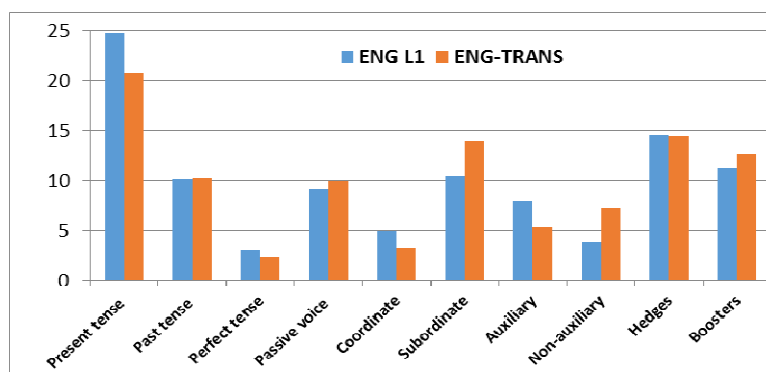


Figure 1.

Non-parallel linguistic analysis of native and non-native English corpora.

When comparing the results of the present study with those of Hu and Cao (2011), similarities and differences were discovered. In effect, the data in figure 1, clearly show that in the present study and that of Hu and Cao, writers whose first language is not English sometimes show preference for using boosters to present claims. That is, non-native English speakers tend to express a higher degree of certainty and confidence when interpreting the research findings, while native writers, in contrast, appeared to be more prudent at the moment of discussing and presenting arguments in M4 (results).

Results of the lexico-grammatical choices in each move across RAAs written in English by native and non-native speakers show differences in the linguistic realization of the abstract moves. The data suggest some linguistic variation in the realization of moves that constitute abstracts in the field of humanities. The statistical analysis reveals that non-native English texts more often use the present tense when

presenting the purpose, M2 (31%), introducing the topic, M1 (25%) and describing methodology, M3 (24%) than their native English counterparts. Moreover, present tense was the linguistic feature in which non-native English abstracts devoted much more space when constructing the purpose (M2). Unlike coordinates, auxiliaries occurred in almost all English abstract moves written in the field of humanities. Although passive voice indicated regular move occurrence in both NE and NNE English RA abstracts (8%), this feature did not occur in M1 (introductions) and M5 (conclusions) throughout native and non-native texts, respectively. Similarly, non-auxiliaries were rarely used in the moves M2, M3 and M5 written by native speakers of English. Moves 2 and 4 had the highest occurrence of using all the linguistic categories, in effect, M2 in NNE abstracts accounted for 48 times while M4 registered 53 times in those texts authored by NE speakers of the language.

On the other hand, the relevant research outcome derived from the contrastive analysis between English and Spanish RAAs written in the fields of humanities and science indicate the common occurrence of the present tense in both languages. Nonetheless, the Spanish abstracts published in Ecuadorian journals, in the two fields have the tendency to use the present tense to communicate their research studies, in essence, 33% corresponds to science and 30% to humanities. The genre analysis so far also revealed that in English abstracts authored by Ecuadorian writers, boosters were common in move M4 in the fields of science and humanities. Meanwhile, hedges in M2 and subordinate clauses in M1 frequently occurred in RAAs written in the field of humanities. Conclusion (M5) was the move reporting the least use of the linguistic categories, particularly in those of Spanish texts, which were used 13 and 11 times in both fields. Passive voice was rarely used in Spanish texts; however, few had the use of the passive form 'se' (see examples 10-13 below), which is used to refer the process without mentioning the agent. This result is in line with Lorés (2014, 2016), who found this linguistic category in sociology texts. The usage differences concerning verb tense, clauses and markers might be due to the nativeness or non-nativeness of authors when producing abstracts (Lau 2004). Moreover, non-native English texts with 183 occurrences in humanities and 189 in science exceeded Spanish abstracts (115 vs. 128, respectively) in the usage of the linguistic features within the five-move pattern. Yet, the frequent occurrence of hedges and present tense significantly varied from move to move. Subordinates and non-auxiliaries occur in conclusion move (M5) in English abstracts published in Ecuadorian journals; nonetheless, these categories were absent in Spanish texts. Although the linguistic variation of the linguistic features discussed earlier was clearly identified across RAAs, the causes for such linguistic choice in each abstracts' move remains quite unclear, in terms of conventional or authors' preferences.

- [10] *En la actualidad casi toda la electricidad que se consume en Galápagos se genera en centrales equipadas con motores de combustión interna.* (Electronics M1)

'Currently, almost all the electricity consumed in the Galapagos is **generated** in plants equipped with internal combustion engines.'

- [11] *Se presenta una variedad de propuestas existentes en la bibliografía sobre modelación, optimización y pruebas para dispositivos solares tipo termosifón, considerando la posible aplicación en el diseño de dispositivos que se adecuen a la geofísica de la zona ecuatorial.* (Electronics M2)

'**The paper presents** existing proposals in the literature on modeling, optimization and Thermosiphon solar devices testing, considering the possible application in the design of devices that fit the equatorial geophysics.'

- [12] *En la deshidratación se aplicó el método de secado en bandejas a tres temperaturas...* (Agronomy M3)

'In the dehydration the drying trays method **was applied** in three temperatures,...

- [13] *Dentro de dicho contexto (...) en el presente artículo se describe la adaptación y descripción de los valores obtenidos en el experimento Go/No-Go dentro de una muestra de estudiantes.* (Sociology M2)

'In this context (...) **the article describes** the adaptation and description of the values obtained in the Go / No-Go experiment in a sample of Ecuadorian students.'

The examples above show the occurrence of the passive form 'se' in Spanish abstracts, particularly when introducing the background, presenting the purpose, and describing the methodology of their research. When comparing the tendency of the linguistic realizations between English abstracts written in the field of humanities and science by North American and Ecuadorian authors, English texts in the field of science showed less frequency of use of the present tense (17%) compared to those of the humanities (29%). However, the frequent occurrence of past tense (15%), subordinates (13%) and passive voice (11%) was higher in the science texts rather than in the field of humanities, as 5%, 11% and 8%, respectively. Although the use of hedges and boosters appears to show a similar tendency of usage between the two fields (15% vs. 14%), they are different in terms of occurrence within the sub-corpora. For example, while boosters in texts written by native English authors in the field of humanities occur the 13% in the M1, this category is optional in the field of science. Moreover, whilst in the field of humanities hedges in M2 are used more often in non-native texts (29%) than native ones (18%), in science their usage is parallel (both at 12%). Additionally, result sections (M4) in both fields, humanities and science report to have the highest occurrence of using all the linguistic features throughout the English RAAs published in North American journals (53 vs. 68, respectively). Conclusion move (M5), meanwhile, issued in humanities and science indicate the least use of those linguistics categories written in non-native English texts, for instance, 13 and 21 tokens each. This outcome might be the result of several factors, for instance, different linguistic and cultural conventions, disciplinary practices and the context of publication, which may define authors' rhetorical preferences and writing style.

The data in the Figure 2 show that passive voice, clauses and modals report variability, in the sense that their frequency of occurrence in the fields of science and humanities across disciplines varies in both languages. Thus, the single most striking outcome to emerge from the data comparison was that auxiliaries and non-auxiliaries were omitted in the majority of abstracts moves across the Spanish sub-corpora in the field of science. Passive tense and subordinates, in contrast, occur similarly in English texts. Present and past tense, and hedges were the linguistics categories that commonly occur within the RA abstracts. However, interestingly, the present and past tense was observed to be used more frequently in Spanish texts than in English ones written in science, and, as well as compared to their counterparts in the field of

humanities. The analysis also showed that passive voice was more frequent in English than in Spanish RA abstracts. Closer inspection of the data indicates that boosters follow the general tendency to occur in English abstracts when authors discuss or present the results of the study. Nonetheless, the further analysis revealed that the set of article abstracts written in English had significantly more usage of the linguistic categories than the other set of abstracts written in Spanish.

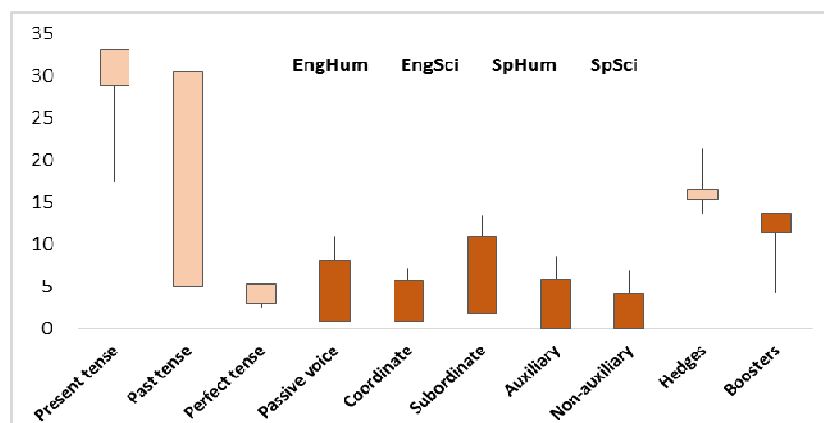


Figure 2.

Linguistic categories in English and Spanish sub-corpora in humanities and science.

The results in Table 2-4, below indicate variation in the textual organization of moves in English and Spanish RAAs in the four disciplines, namely, education, sociology, electronics and agronomy. Although English and Spanish abstracts in the disciplines above mentioned disciplines had fewer occurrences of the linguistic categories throughout the sub-corpora, present tense with the 25% of incidence was the feature that usually occurred in RA abstracts across disciplines. Nonetheless, this feature with 37% of commonness tends to be used more in English sociology abstracts authored by native academics when announcing the essence of the research, especially in moves M2, M4 and M5. Hedges and boosters were frequently used in education texts written in English by non-native speakers of the language, particularly when discussing the research findings. Indeed, they accounted for 16% and 17% of frequency each. The most interesting finding of this cross-disciplinary analysis is that perfect tense, passive voice, clauses, and modal verbs did not occur in agronomy RAAs written in Spanish; however, these features occurred in the majority of English texts written by native and non-native speakers of the language (see Table 2). Moreover, unlike in the sociology abstracts written in Spanish, boosters occurred in the rest of the disciplines such as in the education, electronics and agronomy texts in both English and Spanish sub-corpora. Moreover, a further analysis revealed that there was not any occurrence of the linguistic category in the move M5 throughout the Spanish sub-corpora. Auxiliaries in education, as well as passive voice and coordinate clauses in electronics were infrequently used in RAAs written in Spanish. Furthermore, education and agronomy abstracts written in English and issued by NNE writers had the most frequent use of the linguistics features throughout the corpora, that is, 110 and 113 times of occurrence. On the other hand, whilst in the education

and sociology texts, past tense was partially used to focus on the new findings, in science texts, this feature was deployed when describing the methods (M2) and discussing the results (M4). What is more, RA abstracts written in English by non-native academic authors used more modal verbs than their native counterparts.

The data analysis so far also showed that passive structures were sometimes used in native and non-native English abstracts to reference established knowledge, as well as to emphasize new knowledge. Subordinate clauses occurred in the agronomy text written in English to enhance that the writers' ideas depend on or are the result of a previous conceptual schema, which often indicated the authors' agreement or disagreement with previous findings as a way of justifying the investigation and setting the new knowledge. Perfect tense tends to be the least frequent linguistic element, either in English or Spanish texts. In contrast, hedges, after present tense, were the second most frequent feature used in both fields across the four disciplines. Within the Spanish sub-corpus, there is a clear statistical result revealing that auxiliaries, non-auxiliaries, passive tense and perfect tense were the least frequent micro-categories used throughout the texts written in the field of humanities and science as well.

| Features | English sub-corpora | | | | | | | |
|----------------------|---------------------|------|-------|------|--------|------|-------|------|
| | Educa | | Slogy | | Electr | | Agron | |
| | NES | NNES | NES | NNES | NES | NNES | NES | NNES |
| Present tense | 20 | 23 | 32 | 23 | 19 | 16 | 19 | 15 |
| Past tense | 5 | 5 | 3 | 4 | 12 | 10 | 17 | 19 |
| Perfect tense | 1 | 6 | 3 | - | 2 | 3 | 5 | - |
| Passive voice | 8 | 10 | 4 | 5 | 13 | 8 | 8 | 14 |
| Coordinate | 5 | 5 | 3 | 6 | 4 | 1 | 6 | - |
| Subordinate | 5 | 9 | 11 | 12 | 10 | 15 | 12 | 16 |
| Auxiliary | 3 | 9 | 5 | 3 | 13 | 5 | 8 | 3 |
| Non-auxiliary | 1 | 6 | 1 | 6 | - | 7 | 12 | 8 |
| Hedges | 14 | 18 | 10 | 10 | 13 | 9 | 16 | 16 |
| Boosters | 9 | 19 | 14 | 4 | 8 | 11 | 10 | 13 |

Table 2. Linguistic features in English sub-corpora across disciplines

| | Spanish sub-corpora | | | |
|----------------------|---------------------|-------|--------|-------|
| | Educa | Slogy | Electr | Agron |
| | SPaS | SPaS | SPaS | SPaS |
| Present tense | 23 | 20 | 17 | 21 |
| Past tense | 8 | 6 | 9 | 26 |
| Perfect tense | 3 | - | 6 | - |
| Passive voice | 6 | - | 1 | - |
| Coordinate | 5 | 5 | 1 | - |
| Subordinate | 6 | 5 | 2 | - |
| Auxiliary | 1 | - | - | - |
| Non-auxiliary | - | 3 | - | - |
| Hedges | 19 | 11 | 8 | 11 |
| Boosters | 6 | - | 7 | 6 |

Table 3. Linguistic features in Spanish sub-corpora across disciplines

| | Present tense | Past tense | Perfect tense | Passive voice | Coordinate | Subordinate | Auxiliary | Non-aux. | Hedges | Boosters |
|-------------|---------------|------------|---------------|---------------|------------|-------------|-----------|----------|--------|----------|
| Occr | 248 | 124 | 29 | 77 | 41 | 103 | 50 | 44 | 155 | 107 |
| % | 25 | 13 | 3 | 8 | 4 | 11 | 5 | 4 | 16 | 11 |

Table 4. Linguistic features in English and Spanish sub-corpora across disciplines

4 Conclusions

The analysis of RA abstracts focuses on how the lexico-grammatical features, for instance, verb tenses, clauses, modals and stance markers are distributed in the texts, particularly in the fields of humanities and social science. In this analysis, similarities and differences were discovered, which mainly derive from or depend on the social context in which the RA abstracts are constructed, used and published. Moreover, different discourse conventions practised within academic communities influence the style of academic writing within disciplines, either in national or international academic settings. The adoption of the international generic formats and forms, particularly in scholarly writing is inevitable. This is because experts or novice writers, in order to get international recognition for publishing in well-structured journals, necessarily have to follow or adapt their writing style and rhetorical strategies to the discourse conventions in which they wish to publish their research.

In general, therefore, it seems that RAAs in the three contexts communicate the gist of the entire paper by adapting their own practices and displaying the writing style of those linguistic and cultural conventions practised in each discourse community. As a result of that lingua-cultural divergence, native and non-native abstracts significantly vary in the textual organization of the abstract moves. Nonetheless, English articles use more hedges than Spanish ones to communicate the new knowledge in such a way that readers and the scientific discourse communities, on the one hand, have space for refutation, and, on the other hand, find the focus of the argument to debate. The research outcomes strengthen the idea that the written discourse is socially produced within academic communities, in effect, the meaning and interpretation of the texts greatly depend on the social practices and conventions. This, in fact, is because in accordance with Hyland and Salager-Meyer (2008:300), “knowledge is not a privileged representation of reality but a conversation between members of academic communities who have some agreement on the ground rules for negotiating what counts as plausible”. That is to say, the ways academics write allow readers to know how the information of those texts is or was constructed.

Although the differences between reader–writer relationship, in terms of content and rhetorical structures could be attributed to the variety of conventions and practices within the scientific discourse communities, based on the literature reviewed above, there is no evidence that proves whether such textual organization merely respond to those discourse conventions. Nonetheless, Mur-Dueñas (2011) argues that when addressing papers to an international readership, particularly in English, academic writers have the necessity to adjust their writing to meeting the prevalent conventions and readers’ expectations in the new broader cultural context. Yet, this requisite does not mean they put aside their own rhetorical and writing preference in their native language (2011:3076). In effect, authors adjust their research articles to the prevailing disciplinary convention in which they wish to publish, but their article

abstracts, in some cases infrequently follow a conventional style. Consequently, more broadly, research is also needed to determine if the textual variation responds to the authors' preference or discourse conventions.

From the results, it is quite unclear to what extent native or non-native RAAs written in English adopt particular discourse conventions to communicate the scope of the complete article. However, it seems that non-native article abstracts follow an English conventional style to transmit the information towards an international or national audience; nevertheless, when presenting the RAAs written in Spanish, they tend to maintain their own local discourse community practices. This difference could be the reason why the textual organization of the English translated version, in some cases, significantly varied from their original RAAs. Thus, according to Hyland (2000), the discourse practices and conventions define what disciplines are somehow academically recognized and codified across scientific discourse communities. The textual organization then is influenced by particular discourse practices, which follow disciplinary conventions. That is, the ways in which writers present their arguments, seems to signal their discourse-preferred practices, but also show their professional competence in those disciplinary-approved conventions (Hyland and Salager-Meyer, 2008). In order to promote the essence of the research study writers must convince readers of the effective selection of rhetorical and linguistic elements; otherwise, the information content will remain vague and unclear, and, as a result, the entire paper loses opportunities of indexing and citation.

References

- Al-Khasawneh, F. M. 2017. A genre analysis of research article abstracts written by native and non-native speakers of English. *Journal of Applied Linguistics and Language Research*, 4(1): 1–13.
- Alonso, F. 2014. Evidential and epistemic devices in English and Spanish medical, computing and legal scientific abstracts: A contrastive study. In Bondi, M., Lorés-Sanz R. (eds.), *Abstracts in academic discourse: variation and change*. Bern, Switzerland: Peter Lang AG. 21–42.
- Behnam, B., Golpour F. 2014. A genre analysis of English and Iranian research articles abstracts in applied linguistics and mathematics. *International Journal of Applied Linguistics & English Literature*, 3(5): 173–179.
- Bhatia, V. K. 1993. *Analyzing genre: Language use in professional settings*. Longman Publishing: New York, NY, USA.
- Bordet, G. 2014. Influence of collocational variations on making the PhD abstracts an effective “Would-be Insider” self-promotional Tool. In Marina Bondi & Lorés-Sanz Roza (eds.), *Abstracts in academic discourse: variation and change*. Bern, Switzerland: Peter Lang AG. 130–160.
- Çakır, H. 2015. A Contrastive Study of the Rhetorical Structure of Turkish and English Research Article Abstracts. In Zeyrek-Çigdem, D., Simsek S., Atas U., Rehbein J. (eds.), *Ankara Papers in Turkish and Turkic Linguistics*. Wiesbaden: Harrassowitz Verlag. 367–378.
- Can, S., Erkan K., Qin J. 2016. Structure of moves in research article abstracts in applied linguistics. *Publications*, 4(23): 1–16.
- Connor, U. 2004. Intercultural rhetoric research: beyond texts. *Journal of English for Academic Purposes*, 3: 291–304. doi:10.1016/j.jeap.2004.07.003.
- Divasson, L., León I. 2006. Textual and language flaws: problems for Spanish doctors in producing abstracts in English. *Ibérica*, 11: 61–79.
- Doró, K. 2013. The rhetoric structure of research article abstracts in English studies journals. *Prague Journal of English Studies*, 2(1): 119–139.

- Dos Santos, M. 1996. The textual organization of research paper abstracts in applied linguistics. *Text & Talk: An Interdisciplinary Journal of Language, Discourse & Communication Studies*, 16(4): 481–499.
- Ebrahimi, S., Chan S. 2015. Research article abstracts in applied linguistics and Economics: Functional analysis of the grammatical subject. *Australian Journal of Linguistics*, 35(4): 1–18. doi: 10.1080/07268602.2015.1070660.
- Ebrahimi, S., Motlagh H. 2017. A cross-disciplinary and linguistic study of context frames in research article abstracts. *Research in English Language Pedagogy*, 5(1), 81–95.
- Fallatah, W. 2016. Features of Saudi English research articles abstracts. *Arab World English Journal*, 7(2): 368–379. <https://www.researchgate.net/publication/305615690>.
- Fan, Y., Song Y. 2017. A Contrastive analysis of English abstracts for agricultural science and technology by English and Chinese writers. In Sun, L., Hale L., Fan Q., Zhang J. (eds), *Proceedings of the sixth Northeast Asia international symposium on language, literature and translation: Datong, China*. Marietta, Georgia, USA: The American Scholars Press. 276–280.
- Farzannia, S., Farnia M. 2017. Genre-based analysis of English and Persian research article abstracts in mining engineering journals. *Beyond Words: A Journal on Applied Linguistics and Language Education*, 5(1): 1–13.
- Ghasempour, B., Farnia M. 2017. Contrastive move analysis: Persian and English research articles abstracts in law discipline. *The Journal of Teaching English for Specific and Academic Purposes*, 5(4): 739–753.
- Hu, G., Cao F. 2011. Hedging and boosting in abstracts of applied linguistics articles: A comparative study of English -and Chinese- medium journals. *Journal of pragmatics*, 43: 2795–2809.
- Huckin, T. 2001. Abstracting from Abstracts. In Hewings, M. (ed.). *Academic Writing in Context: Implications and Applications*. Birmingham: The University of Birmingham Press. 93–103.
- Hyland, K. 2000. *Disciplinary Discourses. Social Interaction in Academic Writing*. London: Longman.
- Hyland, K. 2005. Stance and engagement: a model of interaction in academic discourse. *Discourse Studies*, 7(2): 173–192. doi: 10.1177/1461445605050365
- Hyland, K., Salager-Meyer F. 2008. Scientific Writing. *Annual Review of Information Science and Technology*, 42(1): 297–338. doi: 10.1002/aris.2008.1440420114
- Hyland, K., Tse P. 2005. Hooking the Reader: A corpus study of evaluative that in abstracts. *English for Specific purposes*, 24: 123–139.
- Işık-Taş, E. 2018. Authorial identity in Turkish language and English language research articles in Sociology: The role of publication context in academic writers' discourse choices. *English for Specific Purposes*, 49: 26–38.
- Kafes, H. 2012. Cultural traces on the rhetorical rrganization of research article abstracts. *International Journal on New Trends in Education and Their Implications*, 3 (3): 207–220.
- Lau, H. 2004. The structure of academic journal abstracts written by Taiwanese PhD students. *Taiwan Journal of TESOL*, 1(1): 1–25.
- Lee, K. E. 2017. A genre analysis of English and Korean research paper abstracts in electronic resource management in library and information science. *NIDA Journal of Language and Communication*, 22(30): 106–128.
- Lim, J. M. 2006. Method sections of management research articles: A pedagogically motivated qualitative study. *English for Specific Purposes*, 25(3): 282–309.
- Lorés, R. 2004. On RA abstracts: From rhetorical structure to thematic organization. *English for Specific Purposes*, 23 (3): 280–302.
- Lorés, R. 2014. Lost (and Gained) in translation: A contrastive (English/Spanish) analysis of rhetorical and lexicogrammatical patterns in sociology RA abstracts. In Bondi, M., Lorés-Sanz R. (eds.). *Abstracts in academic discourse: variation and change*. Bern, Switzerland: Peter Lang AG, International Academic Publishers. 84–109.
- Lorés, R. 2016. When the local becomes international: the lexicogrammar of rhetorical moves in English and Spanish sociology abstracts. *Languages in Contrast*, 16(1): 133–158. doi 10.1075/lic.16.1.06lor
- Loutayf, M. S. 2017. Analysis of abstracts in English: A study of abstracts written by EFL writers in Argentina. *Argentinian Journal of Applied Linguistics*, 5(1): 15–36.
- Martín-Martín, P. 2003. A genre analysis of English and Spanish research paper abstracts in experimental social sciences. *English for Specific Purposes*, 22: 25–43.

- Martín-Martín, P. 2005. *The rhetoric of the abstract in English and Spanish scientific discourse: A cross-cultural genre-analytic approach*. Bern, Switzerland: Peter Lang.
- Mur-Dueñas, P. 2011. An intercultural analysis of metadiscourse features in research articles written in English and in Spanish. *Journal of Pragmatics*, 43: 3068–3079, doi:10.1016/j.pragma.2011.05.002.
- Poveda, P. 2007. A contrastive analysis of heading in English and Spanish architecture project descriptions. *RESLA*, 20: 139–158.
- Ren, H., Li Y. 2011. A comparison study on the rhetorical moves of abstracts in published research articles and master's foreign-language theses. *English Language Teaching*, 4(1): 162–166. doi: 10.5539/elt.v4n1p162.
- Salager-Meyer, F. 1992. A text-type and move analysis study of verb tense and modality distribution in medical English abstracts. *English for specific purposes*, 11(2): 93–113.
- Samraj, B. (2002). Disciplinary variation in abstracts: The case of wildlife behavior and conservation biology. In Flowerdew, J. (ed.). *Academic discourse*. New York: Longman. 40–56
- Sayfour, N. 2010. *SFL and ESP genre analysis of English research articles in Iranian and English medical journals: A contrastive study (Doctoral Thesis)*. Tehran, Iran: University of Tehran, Faculty of Foreign Languages and Literatures, doi: 10.13140/RG.2.1.4084.7446
- Suntara, W., Usaha S. 2013. Research article abstracts in two related disciplines: Rhetorical variation between linguistics and applied Linguistics. *English Language Teaching*, 6(2): 84–99.
- Swales, J. M. 2004. *Research genres: Explorations and applications*. Cambridge, UK: Cambridge University Press.
- Swales, J. M. 1990. *Genre analysis: English in academic and research setting*. Cambridge: Cambridge University Press.
- Swales, J. M., Feak C. B. 2012. *Academic writing for graduate students: Essential tasks and skill (Third edition)*. Michigan: Michigan Press.
- Swales, J. M., Feak C. B. 2009. *Abstracts and the writing of abstracts*. Michigan: The University of Michigan Press.
- Tovar, R. 2016. Manuscript in press: *Rhetorical structures of English research article abstracts in Ecuadorian and American journal*.
- Tseng, F. 2011. Analyses of move structure and verb tense of research article abstracts in applied linguistics journals. *International Journal of English Linguistics*, 1(2): 27–39. doi:10.5539/ijel.v1n2p27
- Van Bonn, S., Swales J. M. 2007. English and French journal abstracts in the language sciences: Three exploratory studies. *Journal of English for Academic Purposes*, 6(2): 93–108.
- Wallwork, A. 2016. *English for Academic Research: Writing exercises*. New York: Springer.
- Yang, R. Y., Allison D. 2003. Research articles in applied linguistics: Moving from results to conclusions. *English for Specific Purposes*, 22(4), 365–385.