BESZÉDKUTATÁS 2012

BUDAPEST

BESZÉDKUTATÁS 2012

Szerkesztette:
Gósy Mária

MTA Nyelvtudományi Intézet
Kempelen Farkas Beszédkutató Laboratórium
Budapest, 2012

A szerkesztőbizottság elnöke: Gósy Mária
A szerkesztőbizottság tagjai: Horváth Viktória
Mády Katalin
Markó Alexandra
Marton Klára
Olaszy Gábor
Siptár Péter
Vago, Robert

A korpusz szerkesztésében részt vettek: Bóna Judit
Gósy Mária
Imre Angéla
Markó Alexandra
Váradi Viola

Az egyes tanulmányokat szakterületi kompetenciával rendelkező szakemberek lektorálták.

Technikai szerkesztés: Markó Alexandra

A kötet kiadását az MTA Könyv- és Folyóiratkiadó Bizottsága támogatta.

ISSN 1218-8727

© MTA Nyelvtudományi Intézet

Felelős kiadó: Kenesei István igazgató

Készült: Regiszter Kiadó és Nyomda Kft.

TARTALOM

Markó Alexandra: Az irreguláris zönge szerepe a magánhangzók határának jelölésében V(#)V kapcsolatokban 5

Gráczi Tekla Etelka: Obstruenskapcsolatok a spontán beszédben a zöngésségi oppozíció függvényében 30

Bóna Judit: A rövid-hosszú magánhangzók realizációi idősek spontán beszédében 43

Váradi Viola: Beszédhangtörlés a magyar spontán beszédben 58

Abari Kálmán – Olaszy Gábor: A formánsmenetek rendszere CVC kapcsolatok magánhangzóiban a C képzési helyének függvényében 70

Tóth Andrea – Beke András: Suttogott beszédhangok automatikus osztályozása felolvasott beszédben 94

Gósy Mária: Sorozatmegakadások mintázata a spontán beszédben 107

Dér Csilla Ilona: Beszélőváltások során használt diskurzusjelölők a magyar spontán beszédben 132

Schirm Anita: A *hát* diskurzusjelölő a televíziós vitaműsorokban 144

Németh Zsuzsanna: Az ismétlés és a csere interakciós funkciói magyar nyelvű spontán társalgásokban 154

Nemesi Attila László: Implikatúra és arculatmunka televíziós beszélgetőműsorokban 168

Beke András – Gyarmathy Dorottya: Az ismétlések automatikus osztályozása spontán beszédben 186

Ferenc Bunta – Michelle Norton: The production of the American English low front vowel by bilingual Spanish- and English-speaking children and their monolingual peers 200

Neuberger Tilda: Virtuális mondatok gyermekek spontán beszédében 217

Szabó Ágnes – Tóth Andrea: A beszélőváltások sajátosságai óvodáskorú gyermekek társalgásában 234

Menyhárt Krisztina: A beszéd temporális jellemzői 60 évvel ezelőtti gyermek beszélőknél 246

Beke András – Gósy Mária – Horváth Viktória: Gyakorisági vizsgálatok spontán beszédben 260

Tóth Bálint – Németh Géza – Olaszy Gábor: Beszédkorpusz tervezése magyar nyelvű, rejtett Markov-modell alapú szövegfelolvasóhoz 278

A pedagógusjelöltek Kazinczy-versenye 296
Merényi Hajnalka: A pedagógusjelöltek Kazinczy Ferencről elnevezett szép magyar beszéd versenyének 39. országos döntője 296
Latinovits Zoltán: Monológ a költészetről (részlet) 298
Bóna Judit: A 2011-es döntő kötelező szövegének értő-értető előadása .. 298

„Nyelvbotlás"-korpusz 301
1. Téves szó 302
2. Grammatikai hiba 304
3. Kontamináció 305
4. Téves kezdés 306
5. „Nyelvem hegyén van" jelenség 307
6. Perszeveráció 307
7. Anticipáció 308
8. Metatézis 309
9. Egyszerű nyelvbotlás 310
10. Több típusba sorolható jelenségek 311

A magyar és a nemzetközi tudományos élet eseményei 315
17th International Congress of Phonetic Sciences • *Gráczi Tekla Etelka* 315
INTERSPEECH 2001 • *Beke András* 316
Beszédkutatás 2011 • *Deme Andrea* 316

A kötet szerzői 321

AZ IRREGULÁRIS ZÖNGE SZEREPE A MAGÁNHANGZÓK HATÁRÁNAK JELÖLÉSÉBEN V(#)V KAPCSOLATOKBAN

Markó Alexandra

Bevezetés

Két magánhangzó kapcsolatát (pontosabban a magánhangzók közötti mássalhangzóhiányt, űrt) hiátusnak nevezzük, amennyiben ezek a magánhangzók két közvetlenül szomszédos szótagmagot alkotnak (Nádasdy–Siptár 1994: 171; Siptár 2002a: 85). „Ezt a helyzetet a nyelvek általában kevéssé kedvelik, és rendszerint »mindent elkövetnek« annak érdekében, hogy lehetőleg ne forduljon elő bennük hiátus, vagy ha valamilyen okból mégis előfordul, akkor legalább a fonetikai megvalósítás szintjén feloldódjék, eltűnjön" (Siptár 2002a: 85). A magánhangzó-kapcsolatok megszüntetésére háromféle mód kínálkozik: 1. az első vagy a második magánhangzó törlődése (pl. *barna* + *ít* > *barnít*, *kocsi* + *On* > *kocsin*); 2. az egyik magánhangzó félmagánhangzóvá alakulása (pl. *kaleidoszkóp*); 3. a hiátustöltés (pl. *dió*) (Siptár 2002a).

Hiátustöltésre a mai magyarban a *j* mássalhangzó szolgál, ennek megjelenése a kapcsolatot alkotó magánhangzók minőségével hozható összefüggésbe. Mindig kitöltődik a hiátus, ha a magánhangzó-kapcsolat egyik tagja *i/í* (*ki*[j]*áltás*, *ki*[j]*állítás*, *si*[j]*et*, *nő*[j]*i*), és *é* esetében is gyakran (*po*[j]*én* vs. *melléáll* – a pontos kritériumokat lásd pl. Siptár 2002b). Egyéb helyeken nem jellemző a hiátus kitöltődése, bár bizonyos szavaknál egyéni ejtési sajátosságoktól függ a felszíni realizáció (pl. *tea:* [tɛɔ] vagy [tɛjɔ]). Ha mindkét magánhangzó alsó és/vagy kerek, nem megy végbe hiátustöltés (pl. *fáraó*, *neon*, *műút*). A hiátustöltés jelensége nem függ tehát a morfémahatártól.

A magánhangzó-kapcsolódások normatív (köznyelvi) szempontú leírását Elekfi László (1992) végezte el – a kiejtési kézikönyv előmunkálataként. A morfémán belül megjelenő és a morfémahatárra eső, két szótagú magánhangzó-kapcsolatok esetében „**simán** megyünk át egyik magánhangzóból a másikba" (Elekfi 1992: 71; E. L. kiemelése). Ha azonban „a két szomszédos magánhangzó nyílásfoka közt igen nagy a különbség, vagy ha egyikük *i* v. *í* hang, akkor a kettő közt egy szón belül (gyakori és közismert összetett szóban is) *j*-féle kapcsolóhang hallatszik" (uo.). Megjegyezzük, hogy itt a sok *i* + V, V + *i* példa mellett csak három más magánhangzót tartalmazó kapcsolatot említ a szerző: egy-egy *éa*, *éá*, *éo* hiátust tartalmazót. Nem világos tehát, hogy a nagy nyílásfokbeli különbség mire értendő.

A fenti magánhangzó-kapcsolatok esetében is előfordulhat, de Elekfi különösen két azonos magánhangzó találkozásakor tartja gyakorinak a „**hangszalagzárral** való elválasztás"-t, amely szerinte „a gondos, választékos kiejtés sajátsága, de sokszor modoros, **túl pontoskodó**" (1992: 72; E. L. kiemelése). Így folytatja: „Csak összetett szó vagy szókapcsolat **taghatárán** és ott is inkább akkor van helyén, **ha nagyon gondosan tagolni akarjuk** a ritkább, nem mindennapi szót vagy hangsúlyozni a kapcsolat egyik tagját" (uo.; E. L. kiemelése). A hangszalagzár megjelenése tehát túlnyomórészt pragmatikai-szövegtani eredetű, a kapcsolatot alkotó magánhangzók minősége kevésbé befolyásolja, legfeljebb azonosságuk lehet motiváló tényező. Azonos magánhangzók kapcsolódása esetén az összeolvadás elkerülésére Elekfi más megoldásokat javasol (ha a második tag nem hangsúlyos): hangerő- vagy hangmagasságbeli váltást.

A VV kapcsolat egyik magánhangzójának törlődését köznyelvi és bizalmas használatban Elekfi elfogadhatónak tartja a következő szavak esetében: *odaad* [odɔd], *kétségbeesés* [ketʃ:egbeʃe:ʃ], *tudniillik* [tudnil:ik], *vákuum* [va:kum] (1992: 74). *Szótagösszevonás*ként említi a szerző az egyik magánhangzó félmagánhangzóvá válását („melyben az *i*-t sokan mássalhangzószerűen ejtik" 1992: 74). Erre igen kevés példát hoz (főként olyanokat, amelyek kerülendők).

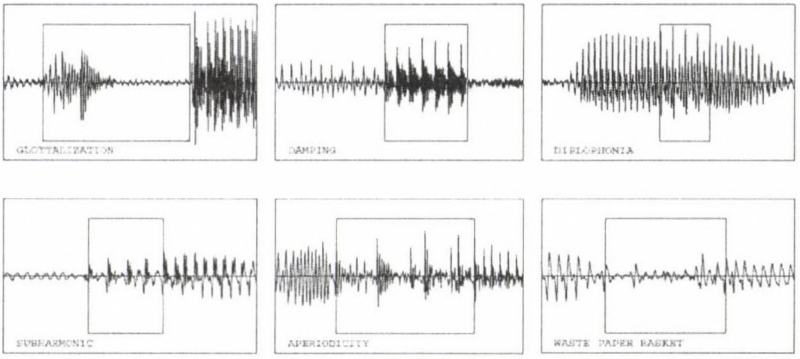
A fonológiai szabályok fonetikai megvalósulását szó belseji hiátusok esetében vizsgálta Menyhárt (2006) felolvasott mondatokon. Összefüggést mutatott ki az artikulációs tempó és a magánhangzók között megvalósuló koartikulációs folyamatok jellemzői között. Leggyakrabban hangátmenet nélkül kapcsolódtak a magánhangzók (45,6%); hiátustöltést 34,6%-ban adatolt a szerző. Ez a koartikulációs jelenség nemcsak *i* és *é*, hanem *ü, ű* környezetében is megvalósult, illetőleg *eá, ae* és *őá* kapcsolatokban is megfigyelhető volt egy-egy beszélő ejtési sajátosságaként. További 13,2%-ban hangátmenettel realizálódott a magánhangzó-kapcsolat, és néhány esetben törlődött az egyik magánhangzó (3,5%), illetve az egyik vokális *j*-ként realizálódott (3,1%). Úgynevezett hangszalagzárral realizált VV kapcsolatot Menyhárt nem adatolt.

A jelen tanulmányban a magánhangzó-kapcsolatok magyar beszédbeli megvalósulását egy korábban kevésbé alkalmazott megközelítésből elemezzük. Vizsgálatunkban (a VV kapcsolódások realizációinak általános jellemzése mellett) arra a jelenségre összpontosítunk, amikor a hiátust nem megszüntetni, hanem éppen ellenkezőleg, jelölni kívánja a beszélő. Ennek egyik módja a fent említett hangszalagzár, hipotézisünk szerint azonban általánosságban a zöngeminőség megváltozása, irregulárissá válása idézheti elő a magánhangzók határának jelöltségét a percepció számára. Kutatásunkban nem érvényesítünk normatív szempontokat, a jelenségek objektív leírására törekszünk.

A (modális) zöngét a szakirodalom a hangszalagok kváziperiodikus rezgéseként határozza meg (vö. pl. Gósy 2004). A zöngeképzés egyes esetekben azonban (szándékosan vagy a beszélő akaratától függetlenül) eltérhet ettől, és

a fonáció irregulárisba (glottalizáció, creaky voice) válthat. A glottalizáció artikulációs, akusztikai és perceptuális sajátosságaira, valamint a terminológia sokszínűségére vonatkozóan kitűnő magyar nyelvű összefoglalás olvasható Bőhm és Ujváry (2008) tanulmányának bevezetőjében.

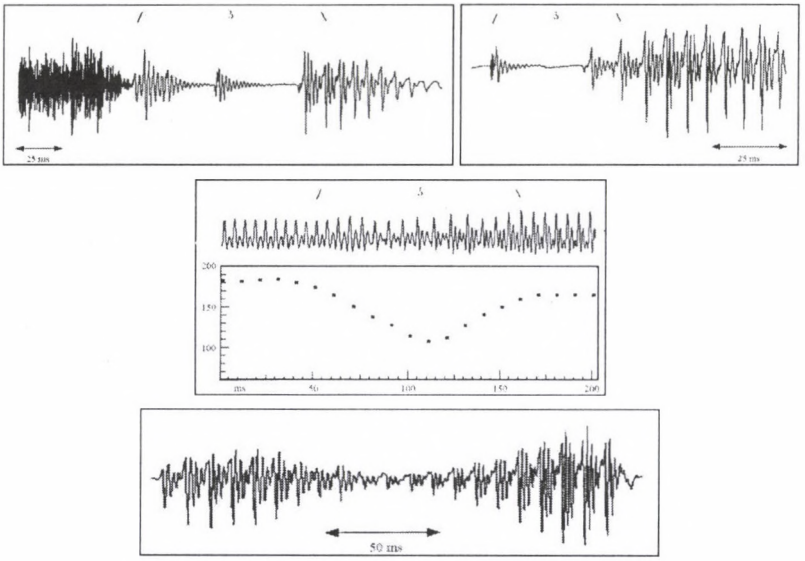
A zöngeminőségnek irregulárisként minősítése nem mindig egyértelmű, hiszen nincsenek olyan, a nemzetközi szakirodalomban elfogadott küszöbértékek (például a jitter és/vagy a shimmer tekintetében), amely fölött glottalizáltnak minősül a hangképzés. Mindemellett az irregularitás is többféle formában realizálódhat. Batliner és munkatársai (1993) például négy beszélő mintegy 30 percnyi spontán és olvasott anyagában hatféle (5 kategória + 1 vegyes) irreguláris zöngetípust különítettek el (1. ábra), hat szempont alapján. A szempontok a következők voltak: 1. az érintett periódusok száma; a 2. rezgés csillapodásának formája; 3. a környezethez viszonyított amplitúdó; 4. az amplitúdó varianciája az irreguláris szakaszon belül; 5. a környezethez viszonyított F0; 6. az F0 varianciája az irreguláris szakaszon belül. A szerzők a kontextus- és/vagy beszélőfüggőnek ítélt megvalósulásokat kizárták a csoportosításból.



1. ábra
Az irreguláris zönge megvalósulásának főbb típusai Batliner és munkatársai (1993: 179) kutatásában

Dilley és munkatársai (1996) 5 beszélő felolvasásában, rádióhírekben vizsgálták a szó eleji magánhangzókat az irreguláris zöngeminőség szempontjából. Kutatásukban négyféle realizációt határoztak meg (2. ábra). Az irregularitás leggyakrabban egymást szabálytalan időközökben követő periódusokként jelent meg (bal felső regisztrátum), amelyek gégezárhanggal is kombinálódhatnak (jobb felső regisztrátum). Az alaphangmagasság hirtelen le- és fellépése a gégezárhang hatását kelti a percepcióban (középső regisztrátum), de az F_0

hirtelen csökkenése vagy növekedése önmagában szintén glottalizációnak minősül. Végül Houde és Hillenbrand (1994) alapján „virtuális glottalizáció"-nak nevezték azokat a megvalósulásokat, amelyekben az amplitúdó csökkenése és növekedése okozza a glottalizációra jellemző perceptuális élményt.



2. ábra
Példák az irreguláris zönge megvalósulásának főbb típusaira
Dilley és munkatársai (1996: 429–430) kutatásában

A korábbi időszakban az irreguláris zönge vagy glottalizáció jelenségét marginálisnak tekintették a (különösen a hazai) fonetikai szakirodalomban, illetve olyan, ritkán jelentkező zavaró tényezőnek, amely a vizsgálatból kizárandó adatok egy részének hátterében áll, vagy lehetetlenné teszi az automatikus címkézést (pl. Tóth–Kocsor 2003). Az utóbbi években azonban a beszédkutatók tudományos érdeklődése egyre nagyobb mértékben fókuszál az irreguláris zöngeminőségre, elsősorban abból az okból, hogy több kutatás igazolta ennek a jelenségnek a kommunikatív funkcióit (Redi–Shattuck-Hufnagel 2001).

A glottalizáció multifunkcionális jelenség. Egyes nyelvekben fonológiai kontrasztot fejez ki – az ilyen nyelvek nagyobb részében szonoránsokat különböztet meg (például a Mexikóban beszélt mazatékban magánhangzókat, néhány észak-amerikai indián nyelvben pedig nazálisokat); ritkábban (mint például az afrikai hauszában) pedig obstruenseket (Ladefoged–Maddieson

1996; Gordon–Ladefoged 2001). Az angol több nyelvjárásában allofónokat különít el a glottalizáció a szótag végi /t/ és /p/ esetében (Pierrehumbert–Talkin 1991, idézi Redi–Shattuck-Hufnagel 2001).

A glottalizáció érzelemkifejező szerepét több nyelven vizsgálták, illetve kísérelték meg felhasználni az érzelmek automatikus felismerésében (például Batliner et al. 2007; Gobl–Ní Chasaide 2003).

A glottalizáció gyakorisága nagymértékben beszélőfüggő: van olyan beszélő, aki alig, és olyan is, aki igen gyakran képez irreguláris zöngét (Henton–Bladon 1988; Dilley et al. 1996; Redi–Shattuck-Hufnagel 2001; Slifka 2006; a magyarra: Markó 2005; Bőhm–Ujváry 2008). Emiatt ez a zöngeminőség a humán beszélőfelismerésben is fontos szerephez jut (Bőhm–Shattuck-Hufnagel 2007). A glottalizáció szociokulturális szerepét is kimutatták egy fiatal amerikai nőkkel végzett vizsgálatban (Yuasa 2010), illetőleg igazolták konverzációs funkcióját is az angolban, ahol a *yeah* diskurzusjelölő irreguláris és modális zöngével történő megvalósításához a beszélő (és a hallgató) különböző szándékokat társít (Grivičić–Nilep 2004).

A glottalizáció frázis/megnyilatkozás végi határjelző szerepét több nemzetközi kutatás igazolta. Henton és Bladon (1988) mondat végi helyzetben mutatta ki a brit angol presztízsváltozatában (RP = Received Pronunciation); illetve az amerikai angolban az irreguláris hangszalagrezgés szintén a mondat végét jelző akusztikai kulcsok egyike (Slifka 2006). Svéd olvasott beszéd frázishatárain is megjelenik az irreguláris zönge (Fant–Kruckenberg 1989), akárcsak a finnben, a csehben és a „szerbhorvát"-ban (Lehiste 1965, idézi Gordon–Ladefoged 2001). Társalgási egység (turn) lezárását is jelezheti a glottalizáció (Redi–Shattuck-Hufnagel 2001).

A magyar beszédet illetően inkább tudománytörténeti szempontból érdekes, hogy már a *Nyelvművelő kézikönyv*nek a hanglejtésbeli hibákról szóló részében is olvashatunk arról, hogy szakasz/mondatvégen glottalizáció jelentkezhet (ott a szerző ezt kellemetlennek, természetellenesnek minősíti, vö. Elekfi 1983: 774). A kísérleti kutatások is kimutatták, hogy a glottalizáció gyakran jelentkezik mondat- vagy közlésvégen – mind olvasott, mind spontán beszédben (Bőhm–Ujváry 2008; Markó 2009, 2010, 2011).

Szóhatáron, két magánhangzó találkozásánál jelentkezik glottalizáció az angolban (Gimson 1980); illetve a magánhangzóval kezdődő szavak esetében a németben és az angolban is gyakori a szó elején (Dilley et al. 1996). Pierrehumbert és Talkin (1992) eredményei szerint a szó eleji glottalizáció akkor a legjellemzőbb, ha a szó intonációs frázist kezdő helyzetben áll.

Egy korábbi, a mondat végi glottalizációt vizsgáló kutatásunkban elemeztük az irreguláris zöngének más pozíciókhoz köthető előfordulásait is. Ennek alapján megállapítottuk, hogy a magyarban a magánhangzós szókezdet (önmagában) nem indukál irreguláris zöngeképzést, ugyanakkor szóhatárra eső magánhangzó-kapcsolódások esetében – a mondatvégeknél kisebb arányban ugyan, de – gyakran jelentkezik glottalizáció (vö. Markó 2011). Ezek az

eredmények motiválták a jelen vizsgálatot, amelyben különböző helyzetekben (morfémán belül, morféma-, illetve összetételi határon és szóhatáron) megjelenő magánhangzó-kapcsolatok vizsgálatát végezzük el, a glottalizációnak a vokálisok elkülönítésében játszott szerepére összpontosítva.

Anyag, módszer, kísérleti személyek

A vizsgálatban 19 különálló mondat felolvasását kértük az adatközlőktől. Az anyagban 9 egyszerű bővített és 10 összetett mondat szerepelt, közülük 16 kijelentő, 2 kiegészítendő, 1 pedig eldöntendő kérdő mondat volt. A mondatok összesen 222 szóból álltak, a szavak száma mondatonként 4–24 között szóródott, az átlagos szószám 11,7 volt. Néhány példa: *Egyszerűen nem értem, miért lenne ez ideális. Mi az oka annak, hogy bezárták a fáraós kiállítást? Néha annyira unta a hosszú utazásokat, hogy elhatározta, felmond, még ha ezzel alááss a is a további karrierjét. Innen a Deák térre egy óra alatt se érsz oda.*

A mondatokban összesen 97 magánhangzó-kapcsolat jelent meg, mondatonként átlagosan 5,1 darab; a mondatonkénti előfordulás 1 és 9 között szóródott. A magánhangzó-kapcsolatok pozíció szerinti megoszlása, valamint típusonként néhány példa az 1. táblázatban olvasható.

A mondatokat magyar szakos egyetemisták és ugyanilyen szakirányú felsőfokú végzettséggel rendelkező felnőttek olvasták fel. A 10 adatközlő életkora 21 és 51 év között szóródott (átlagéletkoruk 30 év), a nemek szerinti eloszlás egyenlő volt. Artikulációjuk és hallásuk ép, és zöngeképzésük sem mutatott patológiás eltérést. A felolvasásokat stúdiókörülmények között rögzítettük. Az adatközlőknek volt lehetőségük előzetesen átolvasni a mondatokat, a felvételt akkor indítottuk, amikor jelezték, hogy felkészültek. Ennek ellenére előfordultak kivitelezési bizonytalanságok, javítatlan félreolvasások, nyelvbotlások, így beszélőnként 95-96 magánhangzó-kapcsolat realizálódott a leírtaknak megfelelően. Végül összesen 959 előfordulást elemeztünk.

A felvételeket a Praat program (5.1 verzió; Boersma–Weenink 2009) segítségével címkéztük fel: különböző szinteken annotáltuk a mondatokat, a magánhangzó-kapcsolatokat, illetve a megvalósulásuk módját.

A magánhangzó-kapcsolatokat az alábbi szempontok szerint elemeztük:

1. a kapcsolat helyzete: morfémán belül, morfémahatáron, összetételi határon és szóhatáron;

2. a kapcsolatot alkotó magánhangzók fonetikai minősége;

3. a magánhangzók határának realizációja: sima átmenet, hiátustöltés, az egyik magánhangzó törlődése, glottalizáció, szünet, valamint ezek kombinációja (pl. hiátustöltés + glottalizáció, szünet + glottalizáció).

1. táblázat: A vizsgált magánhangzó-kapcsolatok pozíciók szerint, példákkal

Pozíció	Előfordulás (db)	Magánhangzó-kapcsolatok	Példák
Morféma belsejében	13	*aó, au, eá, ee, eo, ie, ió, oá, uá, üá*	*fáraó, kalauz, ideális, teendő, neon, karrier, operáció, oázis, aktuálpolitikai, nüánsznyit*
Morfémahatáron	13	*ai, ei, ia, ié, iu, óa, őí, űe*	*kritikai, ismereteit, udvarias, Danié, miután, állandóan, szerkesztőinek, egyszerűen*
Összetételi határon	31	*aa, áa, áá, ae, áo, aő, áu, ea, eá, eí, iá, ie, ió, óe, ói, óu, óú, őa, őá, őe, őú, őü, ua, úe, űe, űú*	*faarccal, ráadás, hozzáállásán, hazaengedték, hozzáolvas, pályaőrnek, ráunt, helyreállító, beírni, kiállítást, úriember, faliórát, adóellenőrtől, folyóirat, lóugrásban, hajóútról, előadás, főúr*
Szóhatáron	40	*a#a, a#e, a#i, a#u, e#a, e#e, e#é, e#ö, e#ü, i#a, i#á, i#e, i#é, i#i, ó#a, ó#á, ó#e, ő#é, ó#o, ő#o, ú#é, ú#u, ű#i*	*oka annak, ha ezzel, belga író, annyira unta, ebbe a könyvbe, lenne ez, eszébe ötlött, mellette ülő, szemközti apartmanban, engedi át, szeptemberi előadáson, korábbi ismereteit, oktató állandóan, helyreállító operáció, lévő oázist, hosszú utazásokat, című irodalmi*

A glottalizált realizációk elemzése a korábbi kutatások módszertanához igazodva (pl. Dilley et al. 1996, Bőhm–Ujváry 2008) vizuális és auditív információk együttes figyelembevételével történt. A Praatban megjelenítettük a hullámformát és a hangszínképet, szükség esetén a dallam- és az intenzitásgörbét is (a címkesorral együtt természetesen); illetőleg folyamatosan és többször ellenőrizve hallgattuk le a hanganyagokat. Akusztikai szempontból glottalizáltnak tartottuk az adott beszédhangrészletet, ha

1. az alapperiódusok időtartama vagy amplitúdója hirtelen jelentősen megváltozott (idetartozónak vettük a hangszalagzár/gégezárhang jelenségét is, vö. Dilley et al. 1996); vagy ha

2. hirtelen a beszélő normál/szokásos hangterjedelme alá csökkent az alapfrekvencia.

Mindemellett percepciós szempontként tekintetbe vettük, hogy a szegmentum hangszínezete jól hallhatóan érdes, rekedtes. Akkor címkéztük glottali-

záltnak a beszédhangrészletet, ha az akusztikai lenyomaton látható, **és** auditív úton észlelhető volt az irregularitás.

Fontos megjegyeznünk, hogy (mint azt a Bevezetésben láttuk) a glottalizáció többfunkciós jelenség. Az a tény tehát, hogy a magánhangzó-kapcsolat irreguláris zöngével realizálódott, nem feltétlenül jelenti azt, hogy az adott szegmentumon a glottalizációnak az a szerepe, hogy a két magánhangzó közötti határt jelölje. Az adatok téves értelmezésének elkerülése érdekében ezért azokat az előfordulásokat, amelyekben a teljes V(#)V szekvencia glottalizált volt, vagy az egyik magánhangzó és annak környezete (a megelőző/követő egy vagy több szótag) irreguláris zöngével realizálódott, nem glottalizáltként vettük tekintetbe – ez az összes megvalósult V(#)V kapcsolat 11,2%-át érintette. Ezeket a realizációkat a nekik megfelelő (sima átmenet, hiátustöltés stb.) típusokba soroltuk be. Megjegyezzük, hogy ez a 10% körüli arány megfelel az átlagos glottalizációs gyakoriságnak (vö. Bőhm–Ujváry 2008, Markó 2011), valamint hogy az ilyen megvalósulások elsősorban a mondatvégekhez közeledve jelentek meg.

Összességében tehát csak a következő realizációkat vettük számba a glottalizált típusban, mert ezek azok, amelyek esetében az irreguláris periódusok helyzete nagy biztonsággal összefüggésbe hozható a magánhangzók differenciálásának szándékával:

– V_1 vége (ritkábban V_1 egészében) glottalizált volt, V_2 modális zöngével realizálódott;

– V_1 modális zöngével realizálódott, V_2 eleje (ritkábban V_2 egészében) glottalizált volt;

– V_1 vége és V_2 eleje glottalizált volt (beleértve azt is, hogy V_2 gégezárhanggal indul, rövid szünetet követően).

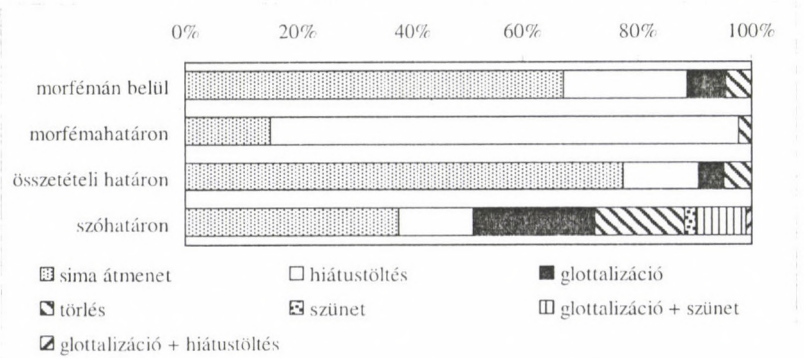
Vizsgáltuk az összefüggést a realizáció típusa és a V(#)V kapcsolat pozíciója, illetve a típus és a magánhangzók minősége között. Elemeztük a glottalizáció megjelenési formáit, valamint az egyéni kiejtési sajátosságok hatását. A statisztikai elemzésekhez (független mintás *t*-próba és korrelációelemzés) az SPSS 15.0 verzióját alkalmaztuk.

Eredmények

1. A magánhangzó-kapcsolat pozíciója és a részt vevő magánhangzók minősége

A magánhangzó-kapcsolatok morfémán belül, összetételi határon és szóhatáron – várhatóan – leggyakrabban sima átmenettel realizálódnak (3. ábra). Morfémahatáron a hiátustöltés volt a leggyakoribb megvalósulási mód, ami azzal függ össze, hogy ebben a pozícióban igen gyakoriak a V + *i* és *i* + V kapcsolatok a mindennapi nyelvhasználatban, valamint ennek megfelelően a jelen korpuszban is. Szende (1973) adatai alapján a leggyakoribb (szó belseji) VV kapcsolatok a következők: *ia, ai, iá, ie, ió, ei,* és csak hetedikként áll a sorban az *i*-t nem tartalmazó *óa.*

A realizációk szóhatáron a legváltozatosabbak, mivel itt a – helyzetből adódóan – szünet (illetve szünettel kombinálódó glottalizáció) is megjelent. A glottalizáció is ebben a helyzetben a leggyakoribb, de morféma belsejében és összetételi határon is találtunk ilyen példákat.



3. ábra
A realizációtípusok megoszlása a különböző szintű VV kapcsolatokban

A következőkben az egyes szinteken megjelenő V(#)V kapcsolatokat részletesebben is megvizsgáljuk, elemezve a részt vevő magánhangzók minőségét is a realizáció típusával összefüggésben.

a) Morfémán belüli magánhangzó-kapcsolatok

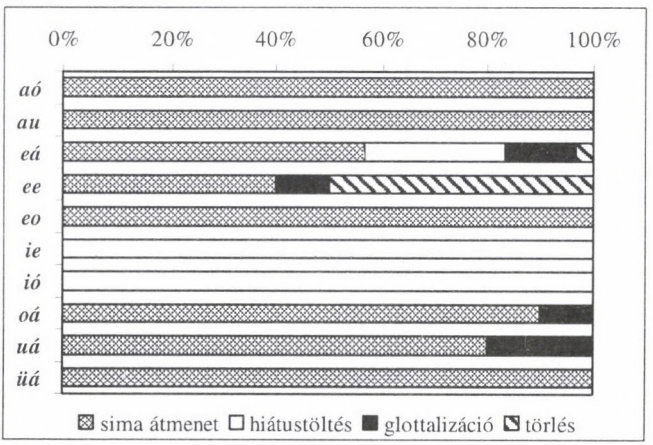
Morfémán belül 130 magánhangzó-kapcsolatot adatoltunk. Ezek közül a sima átmenettel megvalósultak aránya 66,9% volt, hiátustöltő *j* 21,5%-ukban jelent meg, glottalizáció 6,9%-ban. A kapcsolatok 4,6%-ából törlődött az egyik magánhangzó.

A magánhangzó-minőség természetesen egyértelműen befolyásolta a hiátustöltő megjelenését (4. ábra): az *i*-t tartalmazó kapcsolatok az elvártaknak megfelelően mindig így realizálódtak (*operáció, karrier*). A fonológiai szabálynak ellentmond azonban, hogy az *eá* esetében 26,7%-ban ugyancsak adatoltunk hiátustöltést: a ***Deák*** szót 6, az *ideális* és az *óceáni* szót 1 beszélő ejtette így.

Az *aó, au, eo, üá* kapcsolatok mindig sima átmenettel realizálódtak (pl. *fáraó, kalauz, neon, nüánsz*), vagyis a nagy nyílásfokbeli különbség sem idézett elő hiátustöltést. Törlésként vettük számításba (bár a konkrét példában fonetikai szempontból inkább egybeolvadásról beszélhetünk), hogy a ***teendőkről*** szót a beszélők fele egy hosszú [ɛ:]-vel ejtette. (Megjegyzendő, hogy két magánhangzóként vettük számításba azokat az előfordulásokat, ahol a beszélő az alaphangmagasság modulációjával különítette el a hangzókat, mivel ezáltal különálló szegmentumokként voltak észlelhetők, annak ellenére, hogy a

hullámformán és a hangszínképen nem volt látható változás.) Továbbá 1 beszélő az *óceáni* szót [o:tsa:ni]-ként valósította meg.

Glottalizációt négy VV kapcsolat esetében adatoltunk: *eá, ee, oá, uá.* Kétkét (különböző) beszélő irreguláris zöngével különítette el az *ideális*, az *oboás* és az *aktuálpolitikai* morfémán belül összekapcsolódó magánhangzóit. Egy-egy glottalizált realizációt találtunk az *óceáni*, a *Deák* és a *teendőkről* szavak esetében (különböző beszélőknél). E néhány adat alapján nem látunk tendenciát abban, hogy mely összekapcsolódó magánhangzó-minőségekre jellemző(bb) a glottalizált átmenet, egyéni kiejtési sajátosságról lehet szó.



4. ábra
A morfémán belüli magánhangzó-kapcsolatok realizációtípusai

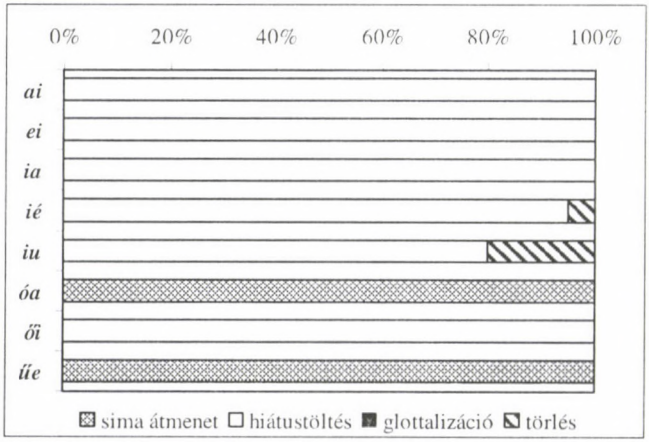
b) A morfémahatáron megjelenő magánhangzó-kapcsolatok

Ebben a pozícióban is megvalósult mind a 130 magánhangzó-kapcsolat. Itt a nagyszámú, *i*-t tartalmazó kapcsolat következtében a hiátustöltés aránya kimagasló: 82,3%. Sima átmenettel a kapcsolatok 15,4%-a valósult meg, az egyik magánhangzó törlődését 2,3%-ban adatoltuk. Glottalizációra nem találtunk példát ebben a kategóriában.

Az *ai, ei, ia, ői* kapcsolatok (természetesen) minden esetben *j*-s átmenettel valósultak meg (pl. *kritikai, nőügyei, udvarias, szerkesztőinek*). Az *ié* és az *iu* esetében törlést is adatoltunk (5. ábra). Mivel az *ié*-t tartalmazó egyik szó a *miért* volt, nem meglepő, hogy előfordult [me:rt] változat is, inkább az a meglepő, hogy csak egyetlen beszélőnél – ez nyilvánvalóan a laboratóriumi beszédhelyzettel indokolható. Ennek fényében azonban váratlannak mondható, hogy a *miután* esetében két beszélőnél is adatoltunk [muta:n] ejtésmódot, ami

a mindennapi beszédben talán nem annyira elterjedt, mint a *miért* lenizált alakja.

Az *i*-t nem tartalmazó kapcsolatok (*állandóan, egyszerűen*) minden beszélő ejtésében sima átmenettel valósultak meg (annak ellenére, hogy a hétköznapi spontán nyelvhasználatban egyre gyakrabban tapasztalható ezeknek hiátustöltővel való megvalósítása).



5. ábra
A morfémahatáron megjelenő magánhangzó-kapcsolatok realizációtípusai

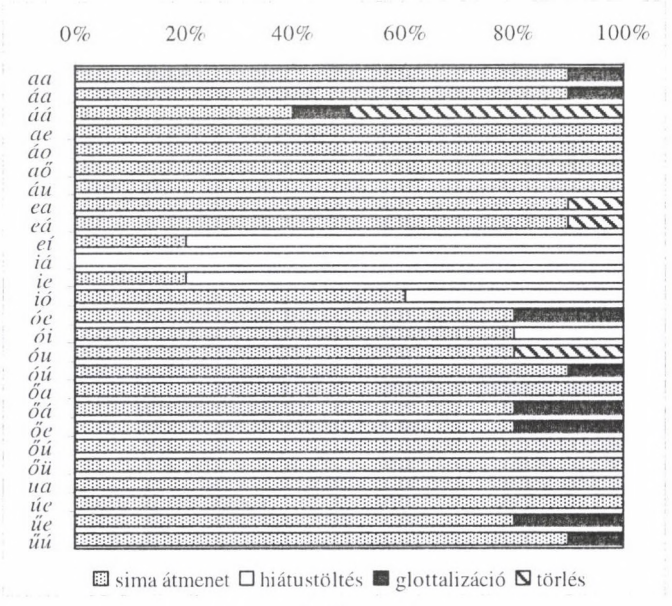
c) Az összetételi határon megjelenő magánhangzó-kapcsolatok

Összetételi határon 304 VV kapcsolatot vizsgáltunk, ebből 77,3% a sima átmenetek aránya, és 13,2%-ban találtunk hiátustöltést. A fennmaradó közel 10%-on a törlés (4,9%) és a glottalizáció (4,6%) egyenlő arányban oszlik meg.

A sima átmenet egy kivételével minden magánhangzó-kapcsolat esetében megjelent (6. ábra), és kizárólagos volt a következőkben: *ae, áo, aő, áu, őa, őú, őü, ua, úe* (pl. *haz**ae**ngedték, hozz**áo**lvas, pály**aő**rnek, r**áu**nt, elő-ad(ni/áson), f**őú**r, n**őü**gyeit, kap**ua**ljban, búcs**úe**stre*).

A korábban látottakkal ellentétben az *i/í*-t tartalmazó kapcsolatok közül csak az *iá* (*ki**á**llításon*) realizálódott minden beszélő ejtésében hiátustöltővel. Az *eí, ie, ió, ói* kapcsolatokban különböző arányban jelent meg a *j*. A ***beí**rni* szót 8 beszélő ejtette hiátustöltővel. Az *ie* kapcsolatot tartalmazó *úri**e**mber* és *ki**e**gészíti* szavakat 9, illetve 7 beszélő artikulálta *j*-vel, így ez a magánhangzó-kapcsolat is 80%-ban mutatott hiátustöltést. Az *i*-nek az *ó*-val alkotott kapcsolatai jóval ritkábban realizálódtak hiátustöltővel (a *fali**ó**rát* 4, a *folyó-irat* 2 beszélő ejtésében). Mindez arra utal, hogy a hiátustöltés jelensége – a

fonológiai szabály megfogalmazásától eltérően – nem teljesen független a kapcsolat másik tagjának minőségétől, valamint a morfémahatár jellegétől.



6. ábra
Az összetételi határon megjelenő magánhangzó-kapcsolatok realizációtípusai

Ebben a csoportban törlés (egybeolvadás) az *áá* kapcsolatra volt a legjellemzőbb (50%), hozzá kell azonban tennünk, hogy ez a kapcsolat két különböző szóban jelentős eltéréssel realizálódott. A *hozz**áá**llásán* szót 8, az *al**áá**ssa* szót azonban csak 2 beszélő ejtette egyetlen *á*-val.

A *l**óu**grásban* VV kapcsolatából 2 beszélő artikulációjában törlődött az *u*, ugyanakkor több beszélőnél tapasztaltuk az *u* félhangzóssá redukálódását. Tekintettel arra, hogy effajta diftongizálódás (azzal a megkötéssel, hogy az *ó* hosszú maradt) csak ebben a szóban volt megfigyelhető, nem alkottunk rá külön kategóriát – ezeket a példákat a sima átmenetek közé soroltuk. Megemlítését mégis szükségesnek tartjuk, mert feltehető, hogy az *u* törlődése a diftongizációs folyamat végállomása, azaz a kettő között fokozati különbség van ($\acute{o}^{u} > o$).

Az *ea, eá* esetében az *e* törlődését egyéni ejtési sajátosságnak tekinthetjük, mivel ezek ugyanannak a beszélőnek a felvételén voltak megfigyelhetők a *bel**ea**d(ta/ott)* és *helyre**á**llító* szavakban.

Glottalizációt 4 kapcsolat (*óe, őá, őe, űe*) esetében adatoltunk 20%-os, 5 továbbinál (*aa, áa, áá, óu, űű*) pedig 10%-os előfordulási arányban. Ez utóbbi közül négy ugyannak a beszélőnek a produkciójában jelentkezett. Két-két beszélőnél az alábbi szavakban fordult elő glottalizáció: *aláássa, adóellenőrtől, előállt, esőemberre, betűejtés.*

d) A szóhatáron megjelenő magánhangzó-kapcsolatok realizációtípusai

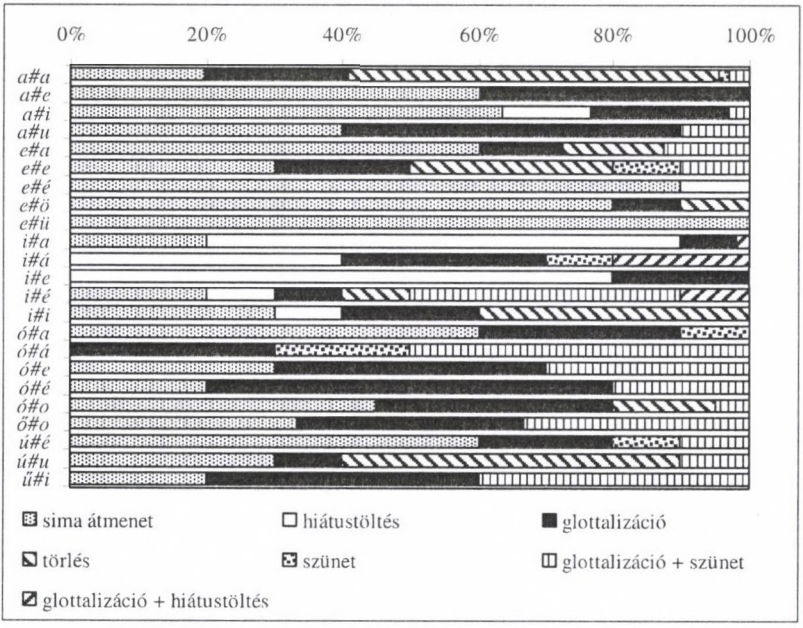
Ebben a pozícióban 395 V#V volt adatolható. Ezek között szintén a sima átmenet volt a leggyakoribb, de már korántsem akkora arányban, mint a morféma belsejében és az összetételi határon megjelenő VV kapcsolatok esetében. A sima átmenetek 37,5%-os részesedéséhez viszonyítva szintén gyakoriak voltak a glottalizált realizációk, együttesen 31,1%-ot tettek ki. Ezen belül a csak glottalizációval megvalósult VV határok 21,3%-ot képviseltek, a glottalizáció szünettel kombinálódott az esetek 8,9%-ában, és hiátustöltésnél is előfordult 1,0%-ban. Törlést 15,7%-ban dokumentáltunk, hiátustöltés 13,7%-ban jelent meg, szünet (önmagában) 2,0%-ban.

A sima átmenet kizárólagosságát csak az *e#ü* (*mellette **ülő***) kapcsolat esetében tapasztaltuk (7. ábra), 90%-os gyakoriságot láttunk az *e#é* (*se **érsz***) esetében. Néhány, *i*-t tartalmazó kapcsolaton kívül még az *ó#á* (*oktató **állandóan***) esetében nem fordult elő egyetlen sima átmenetes realizáció sem.

A hiátustöltés ebben a pozícióban az *i* kapcsolataiban is jelentősen ritkábban jelent meg, mint a szó belseji helyzetekben. 100%-os arányt egyik magánhangzó-kapcsolat esetében sem találtunk. Bár az *i#a* kapcsolatot majdnem minden példában a legtöbb beszélő hiátustöltővel ejtette (*mi **az** 9, aki **a** 8, intézi **a** 8, kiegészíti **a** 10*), volt olyan előfordulása is, amely sosem tartalmazott *j*-s átmenetet (*szemközti **a**partmanban*). Látható, hogy a gyakran hiátustöltővel ejtett kapcsolatok második tagja határozott névelő volt, ami alapján feltehető, hogy szóhatáron ez motiváló tényező a hiátus kitöltődése számára. 80%-os arányban jelent meg hiátustöltés az *i#e* kapcsolatot tartalmazó *szeptemberi **e**lőadáson* jelzős szerkezetben is, ugyanakkor csak 4 beszélő ejtette *j*-vel az *i#á* (*engedi **át***) kapcsolatot. Az *a#i/í* kapcsolat három példából csak egyben valósult meg hiátustöltővel néhány adatközlőnél (*aláássa **i**s* 4, de *haza **i**dőben* és *belga **í**ró* 0). Egy-egy beszélő ejtette hiátustöltővel a *művészeti **é**s, korábbi **i**smereteit* kapcsolatokat, és nem találtunk ilyen realizációt a *című **i**rodalmi* szóhatáron. Az *é* kapcsolatai közül csak a *se **é**rsz* valósult meg hiátustöltővel, és csak 1 esetben. A hiátustöltés összesen 4 *i#V* kapcsolatban kombinálódott glottalizációval.

A törlés ebben a fonetikai helyzetben volt a leggyakoribb, feltehetően azért, mert azonos magánhangzók kapcsolatára a legjellemzőbb, ami szóhatáron fordul elő a legnagyobb valószínűséggel. Az *aa* kapcsolat minden szóhatáron legalább egyszer, összesen 54,5%-ban olvadt egybe (az *óra **a**latt* 8, az *unta **a**, lista **a**, arra **a**z* 6, a *beleadta **a**z* 5, az *oka **a**nnak* 4, a *néha **a**nnyira* 1 beszélőnél). Az *ú#u* (*hosszú **u**tazásokat*) 50,0%-ában, az *i#i* (*korábbi **i**smereit*) 40%-ában, az *e#e* (*lenne **e**z* 5, *térre **e**gy* 1 beszélőnél) 30%-ában, az *ó#o*

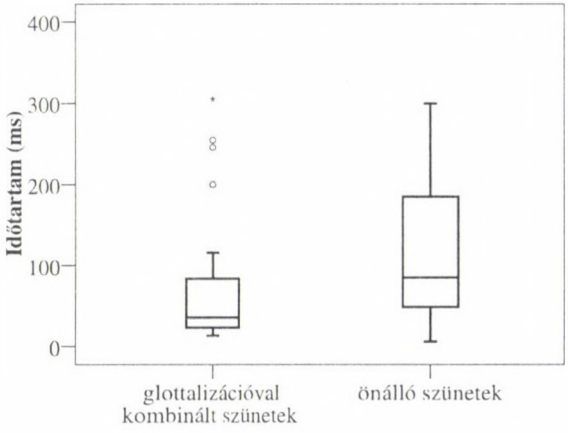
(*helyreállító operáció* 2, *jó oktató* 1 beszélőnél) 15%- ában realizálódott egy magánhangzó. Az eltérő magánhangzó-minőségek esetében az *e#a* 15%-ában (az *ebbe a* esetében 5-ször, a *rendezője az*-nál 1-szer, az *esőemberre a* és a *be a* ejtésében egyszer sem), valamint az *i#é* 10%-ában (*művészeti és* 1 beszélőnél) törlődött a V#V kapcsolat első tagja.



7. ábra
A szóhatáron megjelenő magánhangzó-kapcsolatok realizációtípusai

Szünet önmagában az esetek 2,0%-ában (8 előfordulás) jelent meg, glottalizációval együtt 8,9%-ban (35 előfordulás), azaz 10% feletti a szünettel realizált szóhatárok aránya. A két szünettípus időtartamai is jelentősen különböznek. Az önállóan megjelenő szünetek átlagos időtartama 118 ms, átlagos eltérése 79 ms; míg a glottalizációval kombinálódott szünetek esetében 68 ms az átlag, és 52 ms az átlagos eltérés. A két adathalmaz közötti jelentős különbségek a 8. ábrán is jól látszanak, ahogyan az is, hogy a glottalizációval kombinálódott szünetek közül a 200 ms-os és az annál hosszabbak extrém értékűnek számítanak. Ugyan az önállóan (glottalizáció nélkül) megjelenő szünetek kis száma miatt a statisztikai elemzés eredménye természetesen fenntartásokkal kezelendő, az extrém értékek kihagyásával elvégzett *t*-próba

eredménye szintén szignifikáns eltérést mutat [$t(37) = -3{,}665$; $p = 0{,}001$]. A szünetek időtartama és az irreguláris zönge funkciója között tehát összefüggés látszik, amely úgy értelmezhető, hogy a glottalizáció a rövid szünet tagoló hatását erősíti. Ebből az következik, hogy azok a glottalizált V#V kapcsolatok, ahol a két magánhangzó között extrém hosszú szünet jelent meg, a további elemzésben külön kezelendők, mert ezeknél a glottalizáció funkciója eltér a többi megvalósulásban feltételezett magánhangzó-elkülönítő és – egyértelműsítő szereptől.



8. ábra
A szünetek időtartama (szóródás és medián)

A kizárólagosan szünettel elválasztott szókapcsolatok között 2 beszélőnél szerepelt az *oktató állandóan* és a *térre egy*; illetve egy-egy beszélőnél a következők: *unta a; engedi át; szó[.] ami; hosszú élménybeszámolót.* A magánhangzó-minőségek tekintetében nem látszik tendencia, s bár a frázishatárok jelölése némiképp befolyásolja a szünet megjelenését, ez sem egyértelmű motivációs tényező.

A glottalizáció kimagasló gyakorisággal jelent meg szóhatáron, közel a kapcsolatok harmadánál adatoltuk magában, illetve szünettel vagy hiátustöltővel kombinálódva. A legnagyobb arányban glottalizált szóhatárok frázishatáron jelentek meg, 8-8 beszélőnél: *a jó oktató állandóan hozzáolvas; a helyreállító operáció értelmetlen lesz, a Műút című irodalmi, művészeti és kritikai folyóirat.* 70%-ban realizálódott irreguláris zöngével a következő, szintén frázishatárra eső kapcsolat: *a belga író ebbe a könyvbe.* 66,7%: *az utcán lévő, oázist ábrázoló neonok;* 60%: *a Deák térre egy óra alatt; művészeti és kritikai.* A *néha annyira* és az *annyira unta* esetében a 60%-os arányú glottalizált

megvalósulás a hangsúlyozással függhet össze (lásd alább). Ezek között változó arányban (1–5 beszélő) szerepeltek szünettel kombinálódó estek is.

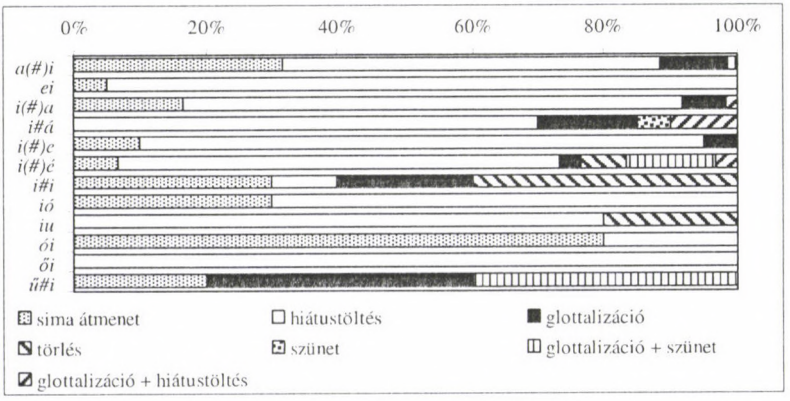
Ebben a pozícióban mindössze két olyan szókapcsolat volt, amelyet egy adatközlő sem valósított meg glottalizációval: *se érsz* és *mellette ülő*. Ezekben az esetekben feltehetően a klitikumos kapcsolat miatt nem volt szükség a határ jelölésére.

A fenti eredmények alapján a szavak által alkotott szerkezet grammatikai (és ezzel összefüggésben prozódiai) minősége befolyásolja a szóhatárra eső magánhangzó-kapcsolat megvalósulását. Ezért a többi példát szintén megvizsgáltuk ennek a tényezőnek a figyelembevételével. A frázishatár nagymértékben előidézte a zönge irregulárissá válását: a beszélők legalább fele glottalizáltan valósította meg azokat a V#V kapcsolatokat, amelyeknél intonációs frázis határát jelölték. Megjegyezzük, hogy nem minden beszélő tagolta azonos módon a mondatokat, például *A belga író ebbe a könyvbe beleadott apait-anyait.* mondatot két és három intonációs frázisra is tagolták az adatközlők. A *Néha annyira unta a hosszú utazásokat...* kezdetű mondat erős emocionális hangsúlyai miatt ez a néhány szó akár három-négy frázisban is megvalósult. Ezzel szemben az is elmondható, hogy az egyes tagmondatokat nem feltétlenül valósították meg különálló intonációs fráziskent a beszélők, például *Az első szó, ami Kádárról eszébe ötlött, a „betűejtés" volt.* mondat első vesszője láttán a beszélők egy kivételével nem álltak meg, és csak hárman jeleztek határt glottalizációval.

Néhányszor klitikumos egységek belsejében, a szavak prozódiai összetartozása ellenére is adatoltunk glottalizációt. Az *engedi át* és a *belga író* kapcsolatokban 5, illetve 4 beszélő váltott zöngeminőséget (a *szemközti apartmanban* és a *haza időben* esetében ugyanez 3-3 beszélőnél történt meg), feltehetően azért, hogy elkerülje a hiátustöltéssel történő realizációt, ahogyan a többi *i*-t tartalmazó kapcsolat túlnyomórészt megvalósult (lásd fent, a szóhatári hiátustöltésről írtakat). Máshol a beszélők feltehetően a törlést (egybeolvadást) kívánták elkerülni, ez a legvalószínűbb magyarázat az azonos (vagy csak időtartamban eltérő) magánhangzók kapcsolatában: *jó oktató, helyreállító operáció, oka annak* (4-4 glottalizált realizáció). A törléseknél láttunk példákat *e#a* kapcsolatokra is, így e két alsó nyelvállású magánhangzó határán jelentkező irreguláris zönge is lehet a törlés elkerülésére szolgáló stratégia, például *ha ezzel* (4 beszélőnél). Mindezek – Elekfi (1992) megállapításával egyezően – a gondos ejtés szándékára utalnak.

Vizsgáljuk most meg a realizációkat elsősorban a magánhangzó-minőségekre összpontosítva. A jobb átláthatóság kedvéért a fonológiai szabály alapján különválasztottuk az *i*-t tartalmazó kapcsolatokat (9. ábra). A hiátustöltés szó belsejében jellemzően megvalósul, bár az igekötő + ige kapcsolat esetében jóval ritkábban. Szóhatáron az összekapcsolódó szavak grammatikai-szemantikai viszonya, valamint feltehetőleg a szókapcsolat gyakorisága befolyásolja a hiátustöltés érvényesülését. A törlés az azonos magánhangzók találkozása-

kor a leggyakoribb, de *miért* > *mért* és a *miután* > *mután* esetében is megjelent, feltehetően a grammatikalizációs folyamathoz tartozó alaki redukció következményeként (Dér 2008). A glottalizáció megjelenését ezekben a helyzetekben a hiátustöltés elkerülésére vonatkozó stratégiaként értelmezhetjük.



9. ábra
Az *i* + V, V + *i* kapcsolatok realizációtípusai

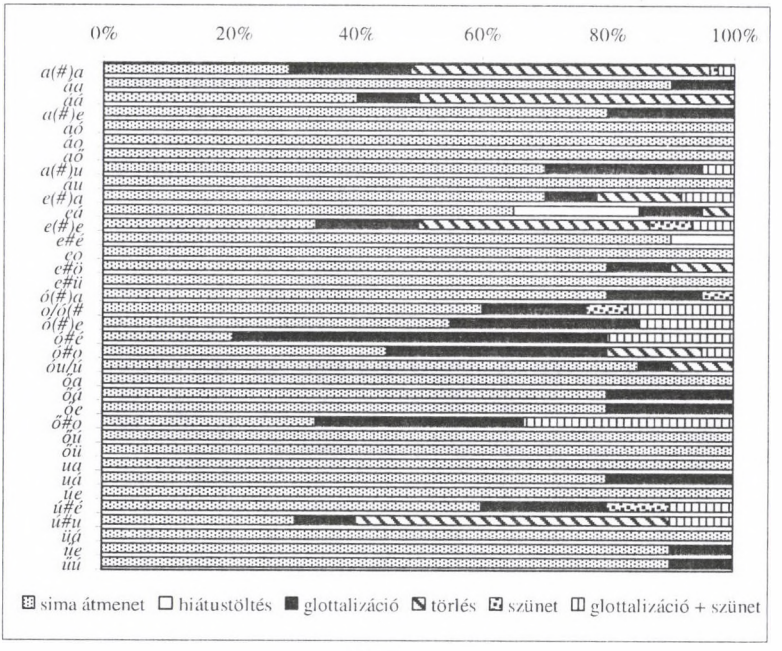
Az *i*-t nem tartalmazó kapcsolatok esetében (10. ábra) várhatóan a sima átmenetes realizációk dominanciáját figyelhetjük meg. Az *é* utótaggal összefüggésbe hozható hiátustöltés illeszkedik a fonológiai szabályhoz. Meglepő azonban, hogy az *eá* (*ideális*) esetében tapasztalt *j*-s átmenet még gyakoribb, amit [ha van ilyen tendencia, l. Elekfi (1992)] a nagy nyílásfokbeli különbséggel sem magyarázhatunk a konkrét esetben. Ugyancsak megfelel az elvárásoknak, hogy a törlés jellemzően azonos (vagy csak időtartamukban eltérő) magánhangzók kapcsolatában jelentkezik, ugyanakkor az is látható, hogy ezen kapcsolatok mindegyike megjelenik glottalizáltan is, vélhetőleg a törlés elkerülése céljából. A glottalizáció a legnagyobb arányban az *ó#é* kapcsolatban jelentkezett, ahol a frázishatár erős motivációs erejét láttuk (*operáció értelmetlen*).

2. A glottalizáció megjelenési formái

Összesen 150 glottalizált V(#)V kapcsolatot adatoltunk. A 200 ms-os és hosszabb szünetekkel kombinálódott glottalizáció esetében a hosszú szünet önmagában is alkalmas a tagoló funkció betöltésére, ezért ezeket a példányokat kizártuk a további elemzésből. Egyébiránt mind a négy ilyen előfordulásra az volt a jellemző, hogy a szünet után a magánhangzó gégezárhanggal (kemény zöngeindítással) kezdődött.

A 146 elemzett magánhangzó-kapcsolat közül 9 realizálódott morfémán belül, 15 összetételi határon, 122 pedig szóhatáron. Ezeknek a realizációtípusonkénti megoszlását mutatja be a 11. ábra, ahol a rövid szünettel kombi-

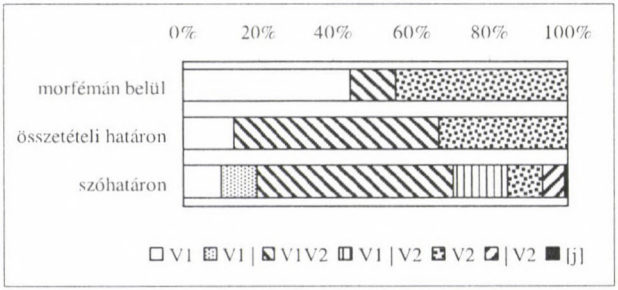
nálódott megvalósulásokat is ábrázoltuk. A morféma belsejében irreguláris zöngével realizálódott VV kapcsolatok esetében általában csak az egyik magánhangzó volt (részben vagy egészben) glottalizált (4-4 esetben, ami egyenként 44,4%-os előfordulási arányt jelent), és mindössze 1 (11,1%) olyan esetet adatoltunk, ahol a magánhangzók közötti átmenet realizálódott irreguláris zöngével. Összetételi határon a glottalizált átmenet volt a jellemzőbb (53,3%), gyakoriságban ezt követte a második magánhangzó elejének (ritkábban egészének) irreguláris volta (33,3%). Olyan eset, amelyben kizárólag az első magánhangzó vége volt glottalizált, csak 13,3%-ban fordult elő.



10. ábra
Az *i*-t nem tartalmazó VV kapcsolatok realizációtípusai

A szünettel kombinálódott glottalizáció természetesen szóhatáron jelent meg, illetve itt adatoltunk egyetlen (0,8%) olyan realizációt is, amelyben a hiátustöltő volt glottalizált (*engedi*[j]*át*). Az esetek mintegy felében itt is a magánhangzók határán jelentkezett irreguláris zönge (50,8%). Ha ehhez hozzávesszük azokat az előfordulásokat is, amelyekben a magánhangzókat szünet választotta el (miközben V_1 vége és V_2 eleje glottalizált volt), ez az arány 64,8%-ra emelkedik. Csak az első magánhangzó volt glottalizált (részben,

ritkábban egészben) 19,7%-ban. Ezen belül egyenlő arányban (9,8-9,8%) oszlottak meg azok az esetek, amikor szünet követte a V_1-et, és amikor nem. 14,8%- ban csak a V_2 (eleje, ritkábban egésze) volt glottalizált, ezen belül szünet nélkül realizálódott 9,0%, és szünet előzte meg a második magánhangzót 5,7%-ban.



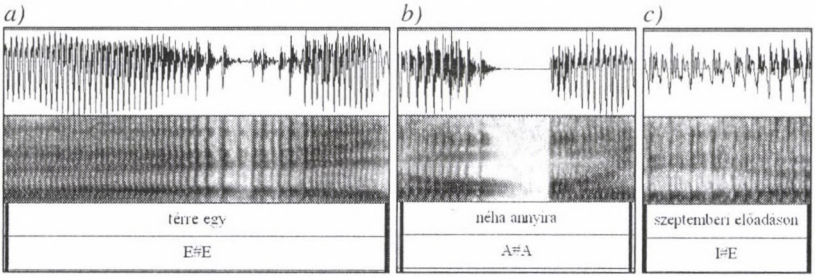
11. ábra
A glottalizáció megjelenési formái a különböző szinteken
(V1 = csak V_1 vége glottalizált; V1| = csak V_1 vége glottalizált, utána szünettel;
V1V2 = a magánhangzók közötti átmenet glottalizált; V1|V2 = az átmenet glottalizált, közben szünettel; V2 = csak V_2 eleje glottalizált; |V2 = csak V_2 eleje glottalizált, előtte szünettel; [j] = a hiátustöltő glottalizált)

Megvizsgáltuk azt is, hogy az irregularitás milyen paraméterekben mutatkozik meg. Az esetek felében a zönge periódusideje ingadozott, 26,7%-ban az amplitúdó, 20,5%-ban mindkét paraméter. Az amplitúdóingadozásnak egy sajátos formája a diplofónia, amely 6-szor jelentkezett. Gégezárhangot (a hosszú szünetes realizációkon túl) további 4 esetben (2,7%) adatoltunk, és minden esetben szünet előzte meg. A 12. ábrán néhány konkrét megvalósulást mutatunk be. A 12. *a)* ábrán egymást rendszertelenül követő és változó amplitúdójú periódusokat látunk a két azonos magánhangzó határán. A 12. *b)* ábrán V_1 végén megritkulnak a periódusok, majd a zöngeképzés is megszűnik, miközben a toldalékcső artikulációs konfigurációja még egy ideig fennmarad, majd szünet után artikulálja a beszélő a következő magánhangzót. A 12. *c)* ábrán diplofónia látható.

3. A beszélők közötti variancia

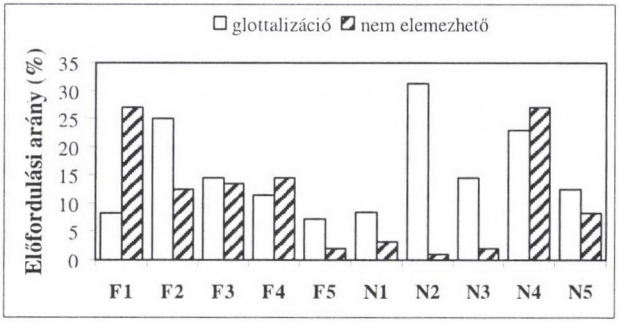
Ahogy láttuk, a V(#)V kapcsolatok 11,2%-ában úgy is megjelent irreguláris zönge, hogy ezt valószínűsíthetően nem a magánhangzók elkülönítésének szándéka motiválta, hanem valamilyen más kommunikációs jelzés vagy pusztán egyéni ejtési sajátosság. Ezt a megállapítást arra alapozhatjuk, hogy ilyenkor vagy az egész magánhangzó-kapcsolat glottalizált volt (ezért nem

lehet az irreguláris zöngének differenciáló szerepe), vagy az irreguláris zöngeképzés kiterjedt a környező beszédhangokra, illetve akár több szótagnyi szekvenciákra is. Ezek többnyire szakasz/mondatvégeken jelentkeztek.



12. ábra
Példák a zönge irreguláris megvalósulására

Az egyes adatközlőkre jellemző glottalizációs gyakoriságot az ilyen V(#)V kapcsolatok számbavételével jellemeztük, és beszélőnként összevetettük az irreguláris zönge e két funkcióban való megjelenési arányát. A magánhangzókat elkülönítő (a 13. ábrán „glottalizáció") és az egyéb (például a mondatvég jelzésére irányuló – a 13. ábrán „nem elemezhető"-ként jelölt) szerepben megjelent irreguláris zönge gyakorisága beszélőnként jelentősen eltér. Az adatközlők fele (F2, F5, N1, N2, N3) jóval nagyobb arányban (legalább 2-szer gyakrabban) glottalizált az összekapcsolódó magánhangzókat elkülönítő szándékkal, mint más okból. Szélsőséges értéket képvisel ebből a szempontból a 2-es számmal jelölt női adatközlő, aki 30-szor olyan gyakran alkalmazza az irreguláris zöngét az egymást követő magánhangzók elkülönítésére, mint más célra.

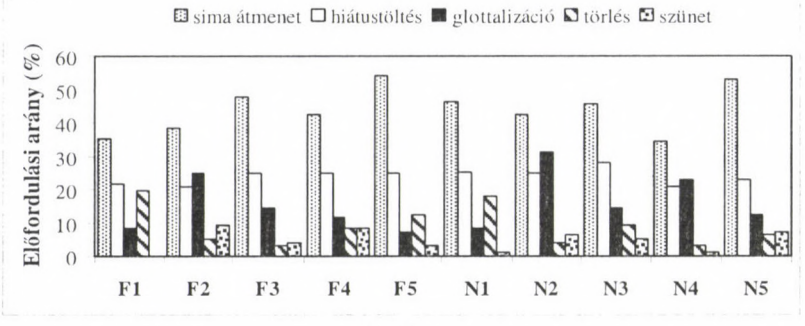


13. ábra
A glottalizációs gyakoriság beszélőnként (F = férfi, N = nő)

Négy beszélő (F3, F4, N4, N5) hasonló arányban alkalmazta a glottalizációt mindkét funkcióban. Noha az 1-es számmal jelölt férfi adatközlő zöngeképzését nagymértékben jellemzi az irregularitás, ő csak a glottalizált V(#)V kapcsolatok egynegyedében differenciálta a hangzókat ezen a módon, nála a zönge irreguláris megvalósulása inkább egyéni beszédsajátosságnak tekinthető.

Megvizsgáltuk azt is, hogy az irregularitás hogyan, milyen paraméterekben jelentkezik az egyes beszélők zöngeképzésében, ebben a tekintetben azonban nem találtunk összefüggést. A glottalizáció megvalósulási módjai tehát – a gyakorisággal ellentétben – nem beszélőfüggőek.

A V(#)V kapcsolatok realizációjának megoszlását is összevetettük a beszélők tekintetében. Természetesen mindenkinek az artikulációjában a sima átmenet dominált, 34,4–54,2%-ban. A hiátustöltés értelemszerűen – mivel fonológiai szabály vezérli – szinte azonos mértékben jelent meg, bár kismértékben jelentkeztek az egyéni ejtési sajátosságok is (20,8–28,1%). A beszélők közötti variancia a legnagyobb mértékben a glottalizációt tekintve érvényesült, ez 7,3–31,3%-os arányban fordult elő. Ugyancsak jelentős különbségek mutatkoztak a törlés gyakoriságában: 3,1–19,8%. A szünet minden beszélőnél viszonylag ritka volt (0,0–9,4%), ami abból is adódik, hogy a VV kapcsolatok nagyobb része szó belsejében realizálódott.



14. ábra
A realizációtípusok aránya beszélőnként (F = férfi, N = nő)

Láttuk, hogy a glottalizáció néhány esetben elkerülő stratégiaként mutatkozott, ezért megvizsgáltuk, hogy van-e összefüggés e jelenség és más realizációtípusok beszélőnkénti gyakorisága között. A Pearson-próba szignifikáns, erős negatív korrelációt ($r = -0{,}708$, $p = 0{,}022$) mutatott ki a glottalizáció és a törlés között, ami igazolja ezt a hipotézist. A hiátustöltés esetében nem találtunk összefüggést (az *i*-t tartalmazó kapcsolatokra korlátozva az adathalmazt).

Következtetések

Tanulmányunkban a magánhangzó-kapcsolatok megvalósulását vizsgáltuk magyar beszédben, különös tekintettel a glottalizációnak az egymást követő magánhangzók elkülönítésében játszott szerepére. Ezekben az esetekben a hiátust nem megszüntetni, hanem valamilyen okból (pl. frázishatár) jelölni kívánja a beszélő. Adataink szerint ennek csak az egyik módja a szakirodalomból ismert gégezárhang (kemény zöngeindítás), általában szünettel kombinálódva.

Vizsgáltuk az összefüggést a realizáció típusa és a VV kapcsolat pozíciója, illetve a típus és a magánhangzók minősége között. Elemeztük a glottalizáció megjelenési formáit, valamint az egyéni kiejtési sajátosságok hatását.

A glottalizáció szóhatáron volt a leggyakoribb, továbbá néhányszor (10% alatti gyakorisággal) megjelent morfémán belül és összetételi határon is. A legnagyobb arányban glottalizált szóhatárok frázishatáron valósultak meg (hatással volt az irreguláris zönge megjelenésére a hangsúlyozás is). A VV kapcsolatok glottalizált realizációnak másik fő motivációja valamely fonológiai-artikulációs automatizmus (a hiátustöltés vagy a törlés) elkerülése volt. A törlés esetében ezt statisztikailag is igazoltuk.

A glottalizáció az esetek jelentős többségében a két magánhangzó határán jelent meg, és főként a periódusidő extrém ingadozásában jelentkezett.

A beszélők között nagyfokú variancia mutatkozott meg a különböző funkciókban megjelenő glottalizáció gyakoriságában. Felmerül a kérdés, hogy a beszélők tudatosan alkalmazzák-e a határjelzésnek ezt a módját. Az eredmények alapján feltételezhetünk valamiféle szándékoltságot a tagolás tekintetében, a határjelölés módjának megválasztásában (hogy az adott helyen éppen irreguláris zöngével él a beszélő) azonban kevéssé tételezhető fel a tudatosság. A glottalizáció megjelenése ezekben a pozíciókban feltehetően abból adódik, hogy az artikuláció során a beszélő fel kívánja függeszteni a zöngeképzést annak érdekében, hogy az egymást követő szegmentumokat megfelelően elhatárolja egymástól, de ez túl sok energiát, időt venne igénybe, ezért a hangszalagok mozgása nem jut el a glottális zár állapotába, csak megközelíti azt. Ezt a magyarázatot támasztja alá Peter Ladefogednek az az elmélete, amely szerint a fonációs típusok egy kontinuum mentén helyezkednek el, nincsenek közöttük diszkrét határpontok (Ladefoged 1971, idézi Gordon–Ladefoged 2001). Eszerint a kontinuum két végpontja a nyitott és a zárt hangszalag-pozíció. Az előbbiben képződő zöngétlen fonációs típustól – a hangszalagok fokozatos záródásával – a levegős (breathy) hangképzésen keresztül jutunk el a modális zöngéig, majd az irreguláris (creaky) zöngén át végül a glottális zárig (15. ábra).

A kapott eredmények igazolják, hogy a glottalizáció a magyar beszédben több okból sem tekinthető marginálisnak. Viszonylag gyakran jelenik meg, és ebből adódóan befolyásolja mind a szegmentális, mind a szupraszegmentális fonetikai elemzések adatait. Másrészt egyértelmű határjelölő funkciója van, a

V#V kapcsolatok esetében éppúgy, mint a mondat/megnyilatkozásvég jelzésében.

15. ábra
A fonációtípusok kontinuuma (Gordon–Ladefoged 2001: 384)
(A két végpont a legnyíltabb és a legzártabb állás; a legnyíltabbhoz tartozik a zöngétlen hangok képzése, innen fokozatosan záródnak a hangszalagok, ami levegős zöngeképzéshez vezet. A hangszalagok mintegy középső állása a modális zöngét idézi elő, ebből a zárt állás felé elmozdulva jelenik meg a glottalizáció. Végül a teljes glottális zárhoz jutunk.)

Vizsgálatunkban felolvasott mondatokban megvalósított magánhangzó-kapcsolatokat elemeztünk. Egy már megkezdett kutatás előzetes eredményei alapján azt is kimondhatjuk, hogy a glottalizáció a spontán beszédben ugyancsak megjelenik magánhangzó-elkülönítő szerepben. A magyar beszédre vonatkozó vizsgálatok folytatása a zöngeminőség kommunikációs funkcióiról való ismereteink bővítését célozza.

Irodalom

Batliner, Anton – Burger, Susi – Johne, Birgit – Kiessling, Andreas 1993. MÜSLI: A classification scheme for laryngealizations. In: *Working Papers, Prosody Workshop*. Lund, Schweden, 176–179.

Batliner, Anton – Steidl, Stefan – Nöth, Elmar 2007. Laryngealizations and emotions: How many Babushkas? In Schröder, Marc – Batliner, Anton – d'Alessandro, Christophe (eds.): *Proceedings of the International Workshop on Paralinguistic Speech (ParaLing'07, Saarbrücken 03.08.2007)*. DFKI, Saarbrücken, 17–22. (http://www5.informatik.uni-erlangen.de/Forschung/Publikationen/2007/Batliner07-LAE.pdf)

Boersma, Paul – Weenink, David 2009. *Praat: doing phonetics by computer (Version 5.1)*. http://www.fon.hum.uva.nl/praat/download_win.html

Bőhm, Tamás – Shattuck-Hufnagel, Stefanie 2007. Listeners recognize speakers' habitual utterance final voice quality. In Schröder, Marc – Batliner, Anton – d'Alessandro, Christophe (eds.): *Proceedings of the International Workshop on Paralinguistic Speech (ParaLing'07, Saarbrücken 03.08.2007)*. Saarbrücken, 29–34. (http://www.bohm.hu/publications/BohmShattuckHufnagelParaling2007.pdf)

Bőhm Tamás – Ujváry István 2008. Az irreguláris fonáció mint egyéni hangjellemző a magyar beszédben. *Beszédkutatás 2008*. 108–120.

Dér Csilla Ilona 2008. *Grammatikalizáció*. Nyelvtudományi Értekezések 158. Akadémiai Kiadó, Budapest.

Dilley, Laura – Shattuck-Hufnagel, Stefanie – Ostendorf, Mari 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24. 423–444.

E[lekfi] L[ászló] 1983. Hanglejtés. In Grétsy László – Kovalovszky Miklós (főszerk.): *Nyelvművelő kézikönyv. Első kötet. A–K.* Második, változatlan kiadás. Akadémiai Kiadó, Budapest, 755–784.

Elekfi László 1992. *A magyar hangkapcsolódások fonetikai és fonológiai szabályai.* Linguistica. Series A. Studia et dissertationes 10. MTA Nyelvtudományi Intézet, Budapest.

Fant, Gunnar – Kruckenberg, Anita 1989. Preliminaries to the study of Swedish prose reading and reading style. *Speech Transmission Laboratory Quarterly Progress and Status Report* 30/2. Royal Institute of Technology, Stockholm, 1–80. (http://www.speech.kth.se/prod/publications/files/qpsr/1989/1989_30_2_001-080.pdf)

Gimson, Alfred Charles 1980. *An introduction to the pronunciation of English.* 3rd edition. Edward Arnold, London.

Grivičić, Tamara – Nilep, Chad 2004. When phonation matters: The use and function of *yeah* and creaky voice. *Colorado Research in Linguistics* 17/1. 1–11. http://www.colorado.edu/ling/CRIL/Volume17_Issue1/paper_GRIVICIC_NILEP.pdf

Gobl, Christer – Ní Chasaide, Ailbhe 2003. The role of voice quality in communicating emotion, mood and attitude. *Speech Communication 40.* 189–212.

Gordon, Matthew – Ladefoged, Peter 2001. Phonation types: a cross-linguistic overview. *Journal of Phonetics* 29. 383–406.

Gósy Mária 2004. *Fonetika, a beszéd tudománya.* Osiris Kiadó, Budapest.

Henton, Caroline – Bladon, Anthony 1988. Creak as a sociophonetic marker. In Hyman, Larry M. – Li, Charles N. (eds.): *Language, speech and mind. Studies in honour of Victoria A. Fromkin.* Routledge, London–New York, 3–29.

Houde, Robert A. – Hillenbrand, James M. 1994 The role of voice pitch in the perception of glottal stops. *Journal of the Acoustical Society of America* 95. (S1) 2872.

Ladefoged, Peter 1971. *Preliminaries to linguistic phonetics.* University of Chicago, Chicago.

Ladefoged, Peter – Maddieson, Ian 1996. *The sounds of the world's languages.* Blackwell Publishing, Oxford.

Markó Alexandra 2005. *A spontán beszéd néhány szupraszegmentális jellegzetessége. Monologikus és dialogikus szövegek összevetése, valamint a hümmögés vizsgálata.* PhD-értekezés. ELTE, Budapest.

Markó Alexandra 2009. Stigmatizált hanglejtésforma a spontán beszédben. *Beszédkutatás 2009.* 88–106.

Markó Alexandra 2010. A prozódia szerepe a spontán beszéd tagolásában. *Beszédkutatás 2010.* 82–99.

Markó Alexandra 2011. A glottalizáció határjelző szerepe a felolvasásban. *Beszédkutatás 2011.* 31–45.

Menyhárt Krisztina 2006. Koartikulációs folyamatok két magánhangzó kapcsolatában. *Beszédkutatás 2006.* 44–56.

Nádasdy Ádám – Siptár Péter 1994. A magánhangzók. In Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 2. Fonológia.* Akadémiai Kiadó, Budapest, 42–182.

Pierrehumbert, Janet – Talkin, David 1992. Lenition of /h/ and glottal stop. In Doherty, Gerard J. – Ladd, D. Robert (eds.): *Papers in laboratory phonology II: Gesture, segment, prosody.* Cambridge University Press, Cambridge, 90–117.

Redi, Laura – Shattuck-Hufnagel, Stefanie 2001. Variation in the realization of glottalization in normal speakers. *Journal of Phonetics* 29. 407–429.

Siptár Péter 2002a. Hiátus. In Hunyadi László (szerk.): *Kísérleti fonetika – laboratóriumi fonológia 2002*. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen, 85–97.

Slifka, Janet 2006. Some physiological correlates to regular and irregular phonation at the end of an utterance. *Journal of Voice* 20/2. 171–186.

Szende Tamás 1973. *Spontán beszédanyag gyakoriségi mutatói*. Nyelvtudományi Értekezések 81. Akadémiai Kiadó, Budapest.

Tóth László – Kocsor András 2003. A Magyar Telefonbeszéd-adatbázis (MTBA) kézi feldolgozásának tapasztalatai. *Beszédkutatás 2003*. 134–146.

Yuasa, Ikuko Patricia 2010. Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women. *American Speech* 85/3. 315–337.

A tanulmány a Bolyai János kutatási ösztöndíj támogatásával készült.

OBSTRUENSKAPCSOLATOK A SPONTÁN BESZÉDBEN A ZÖNGÉSSÉGI OPPOZÍCIÓ FÜGGVÉNYÉBEN

Gráczi Tekla Etelka

Bevezetés

A mássalhangzó-kapcsolatok vizsgálatával több tanulmány foglalkozott már a magyar nyelvben más-más szempontokat alapul véve. A jelen tanulmányban a zöngésség aspektusából elemezzük az obstruenscsoportokat spontán beszédben.

Az obstruensek vagy zörejhangok azon mássalhangzók, amelyek képzése során jelentősen megemelkedik az akadály mögött a nyomás, a külsőtől, akadály előttitől jóval magasabb értékre (pl. Clark–Yallop 1990). Ezek az explozívák, spiránsok és affrikáták, mivel ezeket zár és/vagy rés segítségével hozzuk létre, amelyen semennyi, vagy csak kevés levegő távozik. Az akadály mögötti nyomás növekedésének egyértelmű következménye, hogy a hangszalagok feletti, szupraglottális nyomás is megemelkedik, hiszen a képzési hely és a glottis közötti térben nincs olyan további pont, amely rést vagy zárat hozna létre. A szupraglottális nyomás megnövekedése miatt a zöngeképzéshez szükséges transzglottális nyomáskülönbség a hangszalagok megrezegtetéséhez szükséges érték alá csökkenhet, ami a zöngeképzés leállásához vezethet (pl. Stevens 1998, Shadle 1999). Mindez azt jelenti, hogy a zöngés fonémák megvalósulása részben vagy teljesen zöngétlenedhet. A zönge fenntartásának lehetősége eltér az egyes képzési módok és képzési helyek között. A spiránsok esetében egy további a zöngeképzéssel versengő aerodinamikai mozzanat, hogy a konszonáns minél intenzívebb képzéséhez nagyobb légáramra van szükség, aminek a zöngétlen képzés kedvez a folyamatos nagyobb mértékű levegőáramlás lehetővé tételével (Stevens et al. 1992). A képzési hely is szerepet játszik az egyes konszonánsok alatti zönge fenntarthatóságában. Explozívák esetében minél elülsőbb szervek képzik az akadályt, annál nagyobb az intraorális térfogat és a nyomás által tágítható felület, ezáltal lassabb nyomásnövekedés válik lehetővé (vö. pl. Ohala 1997). Spiránsok esetében a posztalveoláris és a labiodentális mássalhangzók esetében is magasabb zöngearányt adatoltak, mint az alveolárisok esetében (Jesus–Shadle 2007). Ebben fontos szerepet játszhat például az akadályt képező szervek felületének kiterjedése és keménysége is (Shadle 1999).

További meghatározó tényező a fonetikai zöngésségben a közlésbeli helyzet (Westbury–Keating 1986), a beszélők közötti eltérések (pl. van Alphen–Smith 2004), mint például az anatómiai különbségek vagy aktív kompenzáci-

ós gesztusok különböző alkalmazása. Ellentmondó eredmények születtek továbbá a hangkörnyezet hatásáról is (vö. Jesus–Shadle 2003).

Két konszonáns találkozásakor azok realizációja nem vezethető le egyértelműen a részt vevő mássalhangzók intervokális megvalósulásaiból, hiszen például bizonyos artikulációs gesztusok átfedhetnek, például bilabiális és alveoláris képzés esetében nem verseng a két beszédhang a képzőszervért, így az artikulációs gesztusok részben párhuzamosan is megjelenhetnek. Ez a különbség áll fenn például a /sl/ és a /sm/ kapcsolatok között, mivel az elsőben mindkét konszonáns képzésében az apex vesz részt, ejtésük nem fedhet át, míg a második esetben a nazálishoz az ajkak záródása/zártsága kis időben megjelenhet átfedő gesztusként (Hoole 1999). További eltérés lehet például az intervokális helyzettől a mássalhangzó-kapcsolatokban, hogy ejtés könnyítése érdekében, avagy a képzési helyek közötti váltás átmenetének következtében megjelenhet svá a két mássalhangzó között (magyarra: Gósy 2006), ami mint magánhangzós közbeékelődés érdekes hatással lehet a zönge alakulására.

A magyar nyelvben az intervokális és szó eleji, ill. szóvégi explozívák zöngétlenedésében a képzési hely és a szóbeli helyzet is szerepet játszik (Gósy–Ringen 2009). Szókezdő helyzetben ritkán, belsejiben valamivel gyakrabban, szóvégen pedig többször adatoltak zöngétlenedést (átl. 70–74%-nyi zöngearánnyal). A mássalhangzó zöngés részének arányára a képzési hely is hatással volt (Gósy–Ringen 2009, Gráczi 2011).

A magyar spiránsok esetében számos vizsgálat készült. A /v/ megvalósulásainak zöngésségére a hangkörnyezet (intervokális, szókezdő, szóvégi, mássalhangzó-kapcsolatbeli helyzet és környező mássalhangzó) jelentős hatással volt (Bárkányi–Kiss 2006, Bőhm–Olaszy 2007). Abszolút szóvégi és zöngétlen konszonáns körüli helyzetben zöngétlen réshangként, obstruens után általában réshangként, egyéb helyzetekben (némely esetben a beszélő függvényében) approximánsként való megjelenést mutatott leggyakrabban. A hátsóbb képzési helyű konszonánsok intervokális környezetben magasabb zöngearányt mutattak (a zöngések és a zöngétlenek is), mint abszolút szóvégen (Bárkányi–Kiss 2009; Gráczi–Bárkányi megj. alatt). Az egyes képzési helyek esetében magyarban is eltérő mértékű zöngétlenedési tendenciát találtak (Gráczi 2008b).

Spontán beszédben a zöngésségi oppozíció szempontjából a /z/ és a /s/ realizációit (Gráczi 2008a) és a bilabiális, alveoláris és veláris explozívák VOT-értékét (Gráczi et al. 2009) elemezték. A zöngés alveoláris spiráns realizációjának fonetikai zöngéssége függött a hangkörnyezettől (megelőző/követő magánhangzó, zöngés obstruens, szonoráns), a zöngésségi hasonulás következtében létrejött [s] allofón pedig nem különbözött a /s/C kapcsolatoktól, amelyekben a C zöngétlen obstruens volt (Gráczi 2008a). Statisztikailag elkülönültek a zöngekezdési idő alapján az explozívapárok, zöngétlenedés pedig ritkábban fordult elő, mint a spiránsok esetében (Gráczi et al. 2009).

A magyar nyelvben az obstruenskapcsolatokban csak fonológiai zöngésség tekintetében azonos mássalhangzók vehetnek részt, kivéve ha a második fonéma /v/ vagy az első /h/. Ez a szóátvételek ejtésében is megjelenik, két morféma találkozása során pedig regresszív iteratív módon zöngésségi hasonulás megy végbe ismét a C/v/ és /h/C kapcsolatok kivételével. A folyamat posztlexikális és nem tartják tempófüggőnek (ti. beszédstílusfüggőnek, vö. pl. Siptár 2001). A szabály rövid közbeékelt szünet esetében érvényesül (< 55 ms), míg hosszabb (> 314 ms) esetében már nem – Gósy 1998, 2002). A zöngésségi hasonulás eredményeként létrejövő beszédhangok esetében a tanulmányok általában nem teljes oppozícióneutralizációt találtak (pl. Jansen–Toft 2002; Jansen 2004; Gow–Im 2004). Olaszy (2007) például az explozívák esetében kimutatott időtartambeli eltérést a zöngésségi hasonulás eredményeképp megjelent és az alapvariáns időtartama között, míg a spiránsok esetében csak a [f] viselkedett hasonlóan. Gósy (1998) spontán beszédben a lehetséges zöngésségi hasonulási helyek 29%-ban adatolta annak elmaradását. A szóhatáron bekövetkezett szünettartás 77,2%-ban okozta a szabály elmaradását. Bóna és munkatársai (2008) az elvárt regresszív helyett progresszív hasonulást, félzöngés beszédhangot és elmaradt hasonulást is adatolt. Markó és munkatársai (2009) sem olvasott sem spontán beszédben nem adatoltak olyan realizációt, amely szóhatáron megjelent szünet ellenére bekövetkezett volna a zöngésítés irányában, míg zöngétlenítést igen, amely azonban, mint említik, lehet szakaszvégi/szünet előtti zöngétlenedés (vö. Szigetvári 1998) is. Ez a magyar sztenderdnek nem jellemző vonása, de megjelenhet (vö. pl. Kassai 1982). Markó és munkatársai (2009) elemezték a félzöngés beszédhangokat tartalmazó megvalósulásokat is. Megjelenési gyakoriságuk és hogy a CC kapcsolat mely tagja (avagy mindkettő) valósult meg félzöngésen a zöngésségi hasonulás típusától (zöngésítés vagy zöngétlenedés), a beszélőtől és a beszédstílustól is függött. Néhány esetben a zöngésségi hasonulás részleges bekövetkeztét, más esetekben azonban a zöngésség és az obstruensképzés ellentmondásának aerodinamikai magyarázatát tételezték fel.

Ezek alapján számos kérdés felmerülhet. A jelen tanulmányban arra a kérdésre kerestük a választ, hogy a spontán beszéd során az obstruenskapcsolatok [C(#)C] milyen zöngésségi jellemzőkkel rendelkeznek. Egyrészt hogyan alakul az azonos zöngésségű mássalhangzó-találkozások fonetikai zöngéssége, másrészt ehhez képest esetleg mennyiben tér el a zöngésségi hasonulásból származó realizációk viselkedése. Feltételeztük, hogy a mássalhangzókapcsolatokban is megjelenik zöngétlenedés, de a zöngésségi hasonulásból származó zöngésségi eredmények teljes oppozícióneutralizálódást mutatnak szemben a korábbi vizsgálatok időtartamadatainak (részleges) gradualitásával. A /v/ esetében a spontán beszédben is a korábbi eredményeket vártuk, miszerint mássalhangzós környezetben első tagként megtartja a zöngéjét, míg követő hangzóként elveszti a megelőző konszonáns fonológiai zöngésségétől függetlenül.

Kísérleti személyek, anyag és módszer

A kérdések megválaszolásához a BEA felvételek (Gósy 2008) öt női adatközlőjének (22–54 év; átl. 36 év – 1. táblázat) tanulmányairól, munkájáról készült interjújában elemeztük a zöngésségi oppozíciót alkotó obstruensek realizációit. Az adatközlők ép hallók és ép beszédűek voltak.

1. táblázat: A beszélők életkora és átlagos artikulációs tempója

Beszélő	Életkor	Artikulációs tempó (hang/s)
B1	44	11,5±2,8
B2	22	14,3±3,9
B3	29	11,3±2,5
B4	32	12,8±2,8
B5	54	13,5±3,6

A felvételeket hangszinten annotáltuk (Praat 5.3; Boersma–Weenink 2011). Az elemzésben azon kéttagú C(#)C kapcsolatokat vizsgáltuk, melyekben mindkét konszonáns obstruens volt. Az elemzés során kihagytuk azokat a kapcsolatokat, amelyek zajosak voltak, egyszerre beszélés történt, avagy nevetett az adatközlő. Adatoltuk, hogy a mássalhangzó mely fonéma realizációja, milyen hangsorbeli helyzetben áll (szó eleji, szóvégi, abszolút szó eleji, abszolút szóvégi, avagy intervokális konszonánskapcsolat tagja), az időtartamot és a zöngés rész időtartamát. Azokban az esetekben, ahol a kapcsolat nem volt elemeire bontható [például ha elmaradt az első mássalhangzó felpattanása, vagy a megvalósult beszédhang egy mássalhangzót eredményezett, esetleg a zárszakasz maradt el réshang és felpattanó zárhang találkozásakor), egy közös egységként elemeztük. Ezeket C_{12}-ként hivatkozzuk a tanulmányban (pl. elmaradt a [t] zárszakasza: /ɔ**zt**a:n/ → [ɔ**s͡t**a:n], elmaradt a /k/-megvalósulás felpattanása: /ma:ʃodi**kb**ɔ/ → [ma:ʃodi**g͡b**ɔ]) A továbbiakban C_1-gyel a kapcsolat realizációjának első mássalhangzóját, C_2-vel a másodikat jelöljük. Azokat a kapcsolatokat, ahol a képzési hely és a képzési mód azonos volt, tehát egy hosszú mássalhangzót vártunk, C_1 ~ C_2-vel jelöltük (pl. /se:**p** **b**ør/ → [se:**b**:ør], /inka:**b** **b**u:tor/ → [iŋka:**b**:u:tor]). Adatoltuk továbbá, hogy a két mássalhangzó szóhatáron állt-e (C#C avagy CC). Csak azokat a kéttagú mássalhangzó-kapcsolatokat elemeztük, amelyek között nem jelent meg semmilyen időtartamú szünet. A vizsgálatba az azonos és eltérő fonológiai zöngésségű kapcsolatokból is csak azokat vontuk be, amelyek során nem volt várható sorozatos hasonulás (vö. Gósy 2004).

Az akusztikai elemzéseket manuálisan végeztük az oszcillogram és a spektrogram elemzésével, illetve auditív ellenőrzéssel a Praat 5.3 szoftverben. A hanghatár megállapításánál a mássalhangzó indulását a rá jellemző zörejszerkezet megjelenésétől mértük, avagy az előző konszonáns akusztikai szerkezetének végétől mértük. A mássalhangzó végét a követő magánhangzó második formánsának indulásában határoztuk meg követő magánhangzó ese-

tében. A zönge lecsengésének és megindulásának az abszolút zönge lecsengését és megindulását tekintettük.

A zöngésségi oppozícióban részt vevő mássalhangzók zöngés részének arányát elemeztük a konszonánskapcsolat típusa és az abban elfoglalt helye alapján. Összesen 633 (elemezhető) fonémarealizációt adatoltunk, amelyben a konszonánsok eloszlása természetesen nem volt egyenlő (vö. Szende 1973, Gósy 2004) Elemeztük, hogy milyen zöngésségi kategóriába sorolhatóak az egyes realizációk. A határokat Bóna és munkatársai (2008) és Markó és munkatársai (2009) vizsgálatához hasonlóan állapítottuk meg: 20–80% zöngés részarány között félzöngésnek tekintettük a konszonáns realizációját, ez alatt zöngétlennek, felette pedig zöngésnek.

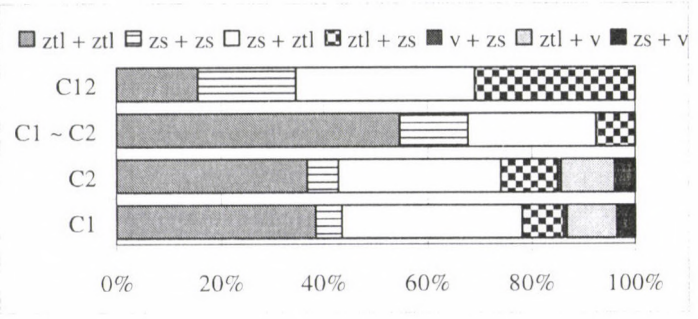
A statisztikai elemzésekhez (egytényezős varianciaanalízis) az SPSS 19.0 szoftvert alkalmaztuk. Csak azokon az eseteken végeztünk statisztikai próbát, amelyek legalább 10 előfordulást mutattak.

Eredmények

A kizárt esetek (zaj, egyszerre beszélés stb.) miatt nem a beszédre jellemző gyakorisági mutatót, hanem a jelen elemzéshez felhasznált adatmennyiséget és -eloszlást közlünk. A felvételeken összesen 633, a fenti feltételeknek megfelelő konszonánst adatoltunk. Ezek CC kapcsolat típus szerinti megoszlását az 1. ábra szemlélteti. Az eloszlásba azokat a konszonánsokat tüntetjük fel, amelyek megjelentek az ejtésben. Amennyiben az ejtés során az egyik mássalhangzó nem jelent meg, nem szerepel a táblázatban. Azokban az esetekben, amikor az első mássalhangzó explozíva, a követő pedig ugyancsak zárképzéssel kezdődik, a zárfeloldás elmaradásakor a megelőző magánhangzó formánsmenete és ennek percepciós hatása alapján döntöttünk arról, hogy valóban csak a zárfeloldás nem volt adatolható az akusztikai lenyomaton, avagy a mássalhangzó nem valósult meg. Hat esetben maradt el zöngésségi párral rendelkező konszonáns szóhatáron (pl. /vɔlɔho**ɟ** **t**ɔnita:ʃ/ → [vɔlɔho **t**ɔnita:ʃ], /ɛ**ɟ** **s**ɔkmunka:ʃ/ → [ɛ **s**ɔkmunka:ʃ]), három esetben szó belsejében (pl. /kɛ**zd**em/ → [kɛ**z**ɛm]).

A leggyakrabban két zöngétlen mássalhangzó találkozását adatolhattuk. Az összes elemzett C_1 közül 38,3% (118 db), C_2 közül 36,7% (90 db), C_{12} közül 15,4% (4 db), az azonos képzési helyű és módú konszonánsok találkozásának (C_1 ~ C_2) pedig 54,7%-át (29 db) tette ki. A zöngés + zöngés találkozás jóval ritkábban fordult elő (C_1: 5,2%; C_2: 6,1%, C_{12} = 19,2%, C_1 ~ C_2: 13,2%). Zöngésségi hasonulás okozta zöngésedés hasonlóan ritkán volt várható: C_1: 7,8%; C_2: 10,6%, C_{12} = 16,7%, C_1 ~ C_2: 10,0%; míg zöngétlenedés a zöngétlen + zöngétlen kapcsolat gyakoriságához hasonló arányban fordulhatott elő (C_1: 34,7%; C_2: 31,4%; C_{12} = 30,8%; C_1 ~ C_2: 7,5%). A /v/-vel alkotott kapcsolatokat külön elemeztük sajátos viselkedése miatt. Zöngés obstruens előtt kétszer fordult elő, zöngétlen előtti megvalósulást nem adatoltunk. 30 /v/-t megelőző zöngétlen és 11 zöngés obstruens fordult elő. A /v/-t zöngétlen

obstruens után 27, zöngés után 10 esetben elemeztük. A vizsgált konszonánsok többnyire nem azonos arányban fordultak elő szóhatáron (2. táblázat).



1. ábra
A CC kapcsolatok típusainak gyakorisága (zs = zöngés, ztl = zöngétlen)

2. táblázat: Az elemzett mássalhangzók CC kapcsolatának megjelenésének aránya (%) szóhatáron

	C_1	C_2	$C_1 \sim C_2$	C_{12}
ztl + ztl	55,9	44,4	89,7	25,0
zs + zs	31,3	53,3	85,7	40,0
zs + ztl	37,4	41,6	7,7	33,3
ztl + zs	45,8	50,0	100,0	75,0
/v/ + zs	0,0	0,0	–	–
ztl + /v/	80,0	84,0	–	–
zs + /v/	45,5	–	–	–

A mássalhangzó-kapcsolatok **első tagjában** zöngétlen mássalhangzók találkozásakor 16,6±21,5%, zöngétlenedést feltételező zöngésségi hasonulási helyzetben pedig 23,6±24,1% zöngés részt adatoltunk. A zöngésségi kategóriákat elemezve a jelen esetben a zöngétlen + zöngétlen kapcsolatban a mássalhangzók 20,3%-a, a zöngés + zöngétlen kapcsolatokban pedig 35,6%-a volt félzöngés, míg további 5,6%-ban zöngés (3. ábra). A zöngés megvalósulással várt esetekben a következő adatokat kaptuk. A zöngés + zöngés kapcsolatokban 75,0±30,7% a zöngétlen + zöngés kapcsolatokban pedig 72,3±43,2% volt a zöngés rész aránya (2. ábra). A szóródás tehát a zöngésnek várt realizációknál a hasonulásból létrejövő konszonánsok esetében mintegy másfélszer nagyobb. Az azonosan zöngés C-k kapcsolatában 43,8%-ban fordult elő félzöngés, a többi 56,2%-ban pedig zöngés megvalósulás. A zöngésségi hasonulás következtében várt zöngés C_1-k esetében 8,3%-ban jelent meg zöngétlen megvalósulás, a zöngés (50,0%) és félzöngés realizációk (41,7%)

pedig hasonló gyakoriságot mutattak (3. ábra). A zöngés rész aránya alapján a C_1-k különböznek [ANOVA: $F(3,264) = 53{,}295$, $p < 0{,}001$], de ez alapvetően a zöngésnek és a zöngétlennek várt megvalósulások eltérésének következménye a Tukey post hoc próba alapján, míg az azonos zöngésségűnek várt esetek nem mutattak szignifikáns eltérést (3. táblázat).

A kapcsolatok **második tagja** zöngétlen + zöngétlen találkozáskor 18,3±31,3%-ban, a zöngés + zöngétlenek esetében pedig 14,8±27,6%-ban voltak zöngések (2. ábra). Ez az első esetben az előfordulások 85,6%-ában, a másodikban pedig 88,3%-ában zöngétlen, 2,2%-ban és 2,6%-ban félzöngés, 12,2% és 9,1%-ban zöngés megvalósulást jelent (3. ábra). Két zöngés mássalhangzó találkozásakor a zöngés rész aránya a C_2-ban 72,3±43,2%, míg zöngétlen + zöngés esetében 66,7±42,0% (2. ábra). Mindkét esetben közel 27%-ban fordult elő zöngétlen megvalósulás, a hasonulást feltételező esetben valamivel gyakoribb volt a félzöngés realizáció (6,7% és 15,4%) (3. ábra). Mindkét esetben az adatok körülbelül kétharmada volt zöngés (66,6% és 57,7%). Az ANOVA alapján a C_1-ekhez hasonló eredményt kaptunk, a csoportok zöngéssége szignifikánsan eltér [$F(3, 207) = 28{,}352$; $p < 0{,}001$], de ez ismét az eltérő zöngésségűnek várt realizációk különbsége (3. táblázat).

A **C_1 ~ C_2 mássalhangzó-kapcsolatok** esetében két zöngétlen konszonáns találkozásakor átlagosan a realizálódó beszédhang 30,9±36,4%-ban, zöngés és zöngétlen kapcsolatban pedig 10,1±5,3%-ban tartalmaz zöngét. Zöngés és zöngés kapcsolatban időtartamának 84,0±27,5%-ban, zöngétlen + zöngés kapcsolatban pedig 47,7±24,1%-ban volt zöngés. Két zöngétlen konszonáns találkozásakor 69%-ban adatoltunk zöngétlen és 10,2%-ban félzöngés megvalósulást, míg hasonulás következtében a beszédhangok 100%-a volt zöngétlen. Zöngésségi hasonulás következtében hasonló eredményt kaptunk. Míg két zöngés mássalhangzó találkozásakor csak az adatok 71,4%-a volt zöngés, 28,6%-a pedig félzöngés, zöngésségi hasonulás következtében a megvalósult beszédhangok 75%-a a zöngés kategóriába volt sorolható, míg 25%-a lényegében zöngétlen volt. Az eltérés az átlagok és az eloszlások között az azonos és a különböző zöngésségű mássalhangzók találkozásában nagyobb mértékű, mint a két külön beszédhangként megvalósuló esetekben akár az első, akár a második mássalhangzó esetében. A kapcsolat szóhatáron való átívelésének hatása részben magyarázhatja az eltéréseket. A zöngés + zöngétlen C_1 ~ C_2 kapcsolatokban ritka volt a szóhatár átlépése, leggyakrabban /d/ találkozott a múlt idő jelével. Ebben a kapcsolattípusban nem végeztünk statisztikai próbát a nagyon alacsony elemszám miatt.

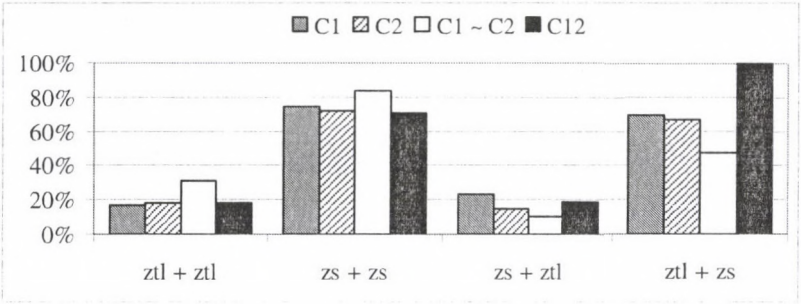
Azokban a megvalósulásokban, ahol két beszédhang megjelenését várjuk, de nem lehetett azokat az akusztikai szerkezet alapján külön elemezni (C_{12}) a zöngétlen mássalhangzók találkozásakor hangszalagrezgés az időtartamnak átlagosan a 18,4±9,9%-ában, zöngés + zöngétlen találkozás esetében hasonlóan 18,6±19,6%-ban volt jelen. A szóródás kétszer nagyobb a második esetben. Az előbbi kapcsolatokban az esetek felében adatolhattunk félzöngés és

zöngétlen megvalósulást is, míg az utóbbiban gyakoribb volt (77,2%) a zöngétlen realizáció. Ez az esetek 60%-ában jelentett félzöngés, 40,0%-ában zöngés megvalósulást. Két zöngés konszonáns találkozásakor 70,8±29,0%-ban, míg zöngésségi hasonulás következtében a teljes időtartamban volt zöngés a realizáció. Az ANOVA alapján ismét statisztikailag jelentős eltérést adatoltunk [$F(3,52) = 9,698$; $p < 0,001$], itt azonban a kapcsolatokat egyesével összevetve eltérő eredményt kaptunk a korábbiaktól. A zöngétlen + zöngés mássalhangzó találkozásából létrejött beszédhangok fonetikai zöngéssége nem tért el egyik más kapcsolatban létrejött beszédhangtól sem. Ebben az esetben azonban csak nagyon kevés adattal dolgozhattunk ennél a csoportnál (4 előfordulás), így az elemzést a többi, nagyobb számú csoport szempontjából érdemes csak figyelembe venni. Azok a korábbiakhoz hasonlóan azt mutatták, hogy a zöngésnek és a zöngétlennek várt realizációk eltérnek, a zöngésségi hasonulás következtében létrejött konszonáns azonban nem mutatott szignifikáns különbséget az azonos zöngésségű fonémák kapcsolatából származóktól (3. táblázat).

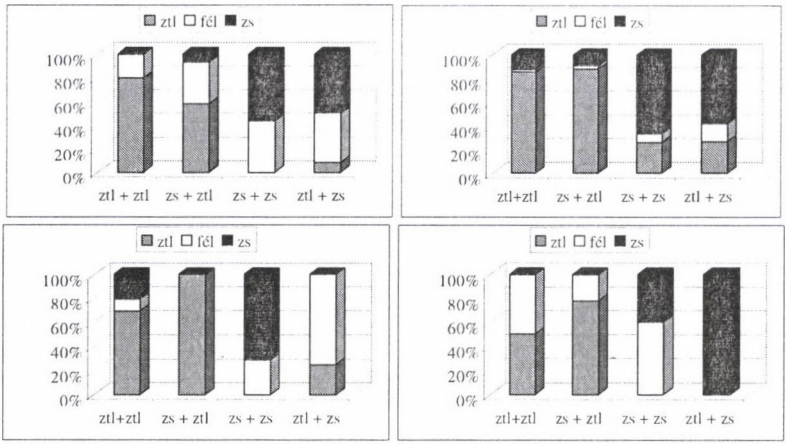
A zöngés labiodentális spiráns CC kapcsolatok első tagjaként minden megvalósulásában teljesen zöngés volt, míg a követő mássalhangzó átlagosan időtartamának 54,1%-ában, míg ha a /v/ a második konszonáns volt, a C_1 átlagosan 94,3%-ában volt zöngés, a /v/ megvalósulása pedig ismét teljes idejében. Zöngétlen konszonánssal alkotott kapcsolatban a megelőző mássalhangzó átlagosan 12,2±9,3%-ban, a /v/ realizációja pedig 70,4±38,2%-ban volt zöngés. A C_1 tehát a várt módon zöngétlen maradt, a labiodentális spiráns pedig részben zöngétlenedett a korábbi nem spontán anyagokhoz hasonló viselkedést mutatva.

Az azonos típusú mássalhangzó-kapcsolatok egyes tagjait (C_1, C_2, $C_1 \sim C_2$ és C_{12}) összevetve nem adatoltunk szignifikáns eltérést [ANOVA: $F(3,240) = 2,144$, $p = 0,095$; $F(3,42) = 0,209$, $p = 0,890$; $F(3,205) = 2,545$, $p = 0,057$; $F(3,61) = 2,605$, $p = 0,060$].

Ezek alapján ugyan előfordul, hogy a zöngésségi hasonulás elmarad, progresszív, vagy zöngétlenedés történik, de alapvetően az első konszonánsok realizációi zöngésség szempontjából nem választhatóak el az alaprealizációjában azonos zöngésségűnek várt beszédhangoktól. Elemeztük azokat a mássalhangzó-kapcsolatokat, amelyek megvalósulása két önálló beszédhangként jelent meg, és mindkét beszédhang zöngésségi oppozíciót alkotó fonéma megvalósulása volt. A 4. ábra ezek zöngésségi kategóriájának megoszlását mutatja. A két zöngétlen fonéma és a zöngés + zöngétlen találkozásából megvalósult hangkapcsolatokban nem adatoltunk zöngés realizációt, és zöngétlen-félzöngés megvalósulást is csak kevesebb mint 3%-ban. Az első tag félzöngéssége kétszer gyakoribb volt a hasonulási helyzetben, 31,3% a 15,7%-hoz képest, de minden esetben zöngétlen konszonáns követte. Zöngés-zöngétlen realizáció csak a hasonulási helyzetben jelent meg, összesen 3,1%-ban.



2. ábra
Az elemzett mássalhangzók zöngés részének aránya (átlag) a fonémakapcsolat típusa és az abbéli helyzete alapján



3. ábra
Az elemzett mássalhangzók zöngésségi kategóriája a mássalhangzó-kapcsolat típusa és az abbeli helyzet függvényében.
Bal fent: C_1, jobb fent: C_2, bal lent: C_1 ~ C_2, jobb lent: C_{12}

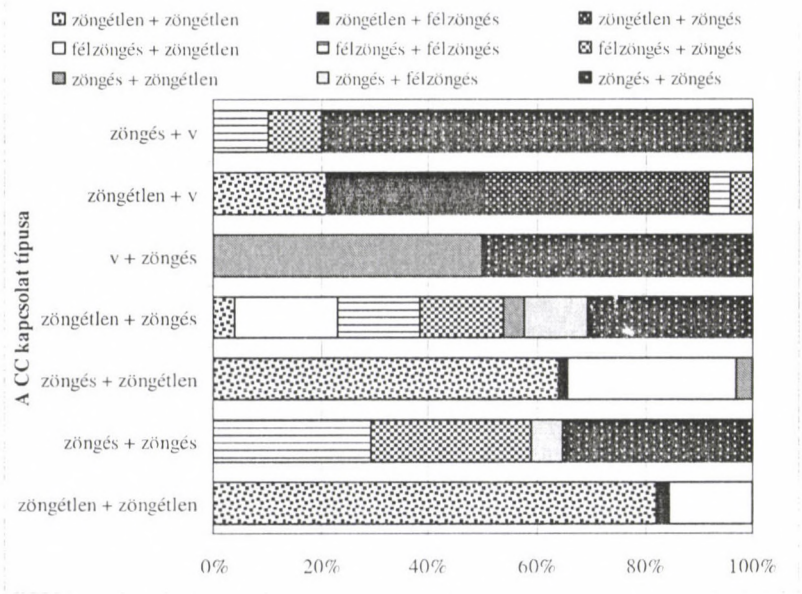
Hasonlóan alakult a zöngés + zöngés és zöngétlen + zöngés kapcsolatok hangpárjainak eloszlása. Az első konszonáns többnyire zöngés vagy félzöngés megvalósulást mutatott mindkét esetben, de ezt kizárólag a hasonulási helyzetben követte zöngétlen mássalhangzó (22,7%). Ugyancsak csak a hasonulási helyzetben következett be két zöngétlen beszédhang megjelenése (4,5%).

A /v/-s kapcsolatokban a labiodentális spiráns csak mássalhangzó utáni helyzetben vesztett zöngéjéből, azonban a követő zöngés konszonáns gyakran

zöngétlenedett. Zöngétlen mássalhangzó után a labiodentális gyakorta zöngétlen megvalósulással jelent meg, míg zöngés után általában zöngésen, és ezekben a kapcsolatokban a C_1 is legfeljebb részben lett zöngétlen (22,2%).

3. táblázat: A Tukey post hoc próba eredményei (*p*-érték) a mássalhangzók egyes kapcsolattípusban történt realizációinak fonetikai zöngésségére

	C_1	C_2	$C_1 \sim C_2$
ztl+ztl vs. zs+zs	< 0,001	< 0,001	< 0,001
ztl+ztl vs. zs+ztl	0,134	0,900	1,000
ztl+ztl vs. ztl+zs	< 0,001	< 0,001	< 0;001
zs+zs vs. zs+ztl	< 0,001	< 0,001	< 0,001
zs+zs vs. ztl+zs	0,904	0,951	0,036
zs+ztl vs. ztl+zs	< 0,001	< 0,001	< 0,001



4. ábra
A beszédhangpárok zöngésségi kategóriáinak eloszlása

Következtetések

A jelen tanulmányban az obstruensek CC kapcsolatokbeli megvalósulásait elemeztük spontán beszédben. A spontán beszéd során az egyes fonémák megjelenése nem azonos arányú (Szende 1979; Gósy 2004), ezért a tanul-

mányban nem elemeztük a kapcsolatokat a részt vevő tagok képzési módja és helye szerint.

A mássalhangzó-kapcsolatokban megjelenő obstruensek realizációi zöngés részük aránya alapján statisztikailag nem különböztek azokban az esetekben, amikor az adott zöngésség fonetikai megvalósulása zöngésségi hasonulás vagy alaprealizáció következtében volt várható. Ugyanakkor azokat a megvalósulásokat elemezve, ahol a mássalhangzók két külön beszédhangban jelentek meg, az eredmények részben eltérő tendenciát mutatnak a hasonulásból és az azonos zöngésségű fonémák kapcsolatából eredő megvalósulásokban. Ennek feltehetően az oka, hogy a zöngésségi hasonulás néhány esetben elmaradt, néhány esetben esetleg csak részlegesen ment végbe, míg a zöngés obstruensek CC kapcsolatbeli teljes zöngétlenedése ritka volt, és hasonlóan a zöngétlenek részben vagy teljesen zöngés megjelenése is. Tehát a félzöngés és „progresszív" hasonulások egy részében várhatóan nem csak a zöngésség fenntartásának nehézsége játszik szerepet. Ennek a kérdésnek a pontos megválaszolásához azonban további elemzéseket kell végezni nagyobb adatmennyiségen és több beszélő bevonásával, amelyek a CC kapcsolatok további jellemzőinek mint független változók bevonásával választ adhatnak.

A labiodentális zöngés spiránst fonológiai és fonetikai sajátosságai miatt részben a többi obstruenstől elkülönítve elemeztük. A spontánbeszédbeli megvalósulások a korábbi olvasott beszéden végzett elemzések eredményeihez igazodni látszanak, vagyis a zöngésség megtartása függött a mássalhangzó-kapcsolatbeli helyzettől és a másik résztvevő konszonánstól.

Irodalom

van Alphen, Petra M. – Smits, Roel 2004. Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: the role of prevoicing. *Journal of Phonetics* 32/4. 455–491.

Bárkányi Zsuzsanna – Kiss Zoltán 2006. A fonetikai célpontok összeférhetetlenségéről: a magyar *v* fonológiájának fonetikai alapú megközelítése. In Kálmán László (szerk.): *KB 120: A titkos kötet. Nyelvészeti tanulmányok Bánréti Zoltán és Komlósy András tiszteletére.* MTA Nyelvtudományi Intézet–Tinta Könyvkiadó, Budapest, 263–275.

Bárkányi Zsuzsanna – Kiss Zoltán 2009. Word-final fricative contrasts in Hungarian. A phonetic approach. Előadás. *Budapest Phonology Circle.* 2009. november 5. http://budling.nytud.hu/~cash/papers/buphoc09-slide.pdf

Boersma, Paul – Weenink, David 2011. *Praat: doing phonetics by computer* (Version 5.3). http://www.fon.hum.uva.nl/praat/download_win.html

Bóna, Judit – Gráczi, Tekla Etelka – Markó, Alexandra 2008. Coarticulation rules and speaking style dependency. Rudolph Sock – Susanne Fuchs – Yves Laprie (eds.): *Proceedings of Internationatl Seminar fo Speech Production 2008.* Strasbourg, 245–247.

Bőhm Tamás – Olaszy Gábor 2007. A magyar [v] hang szerkezetének és zörejességének fonetikai vizsgálata. *Beszédkutatás 2007.* 19–35.

Clark, John – Yallop, Colin 1995. *An introduction to phonetics and phonology.* Blackwell Publishers, Oxford.

Gósy Mária 1998. A zöngésségi folyamat a (spontán) beszédben. *Beszédkutatás '98.* 1–20.

Gósy Mária 2002. Temporal coding of voicing assimilation in speech production. *Acta Linguistica Hungarica* 49/3–4. 257–276.

Gósy Mária 2004. *Fonetika, a beszéd tudománya.* Osiris Kiadó, Budapest.

Gósy Mária 2006. A semleges magánhangzó nyelvi funkciói. *Beszédkutatás 2006.* 8–22.

Gósy Mária 2008. Magyar spontánbeszéd-adatbázis – BEA. *Beszédkutatás 2008.* 194–207.

Gósy Mária – Ringen, Cathrine O. 2009. Everything you always wanted to know about VOT in Hungarian. Előadás. *International Conference on the Structure of Hungarian 2009.* Budapest, 2009. szeptember 1. http://icsh9. unideb.hu/pph/handout/Ringen_Gosy_handout.pdf

Gow, David W. – Im, Aaron M. 2004. A cross-linguistic examination of assimilation context effects. *Journal of Memory and Language* 51. 279–296.

Gráczi Tekla Etelka 2008a. Alveoláris spiránsok akusztikai fonetikai vizsgálata. *Beszédkutatás 2008.* 33–51.

Gráczi Tekla Etelka 2008b. „Z"-öngétlenedés: beszélőfüggő paraméter? In Váradi Tamás (szerk.): *II. Alkalmazott Nyelvészeri Doktorandusz Konferencia.* MTA Nyelvtudományi Intézet, Budapest. 34–43.

Gráczi Tekla Etelka 2011. Explozívák a zöngésségi oppozíció függvényében. In Váradi Tamás (szerk.): *V. Alkalmazott Nyelvészeti Doktoranduszkonferencia.* MTA Nyelvtudományi Intézet, Budapest, 51–66.

Gráczi Tekla Etelka – Markó Alexandra – Beke András 2009. Zöngekezdési idő a spontán beszédben. Előadás. *Beszédkutatás 2009.* Budapest, 2009. október 17.

Gráczi Tekla Etelka – Bárkányi Zsuzsanna (megj. alatt). Voiced fricatives in Hungarian in utterance-final position. In: *Proceedings of Grammar and Context – New Approaches to the Uralic Languages III.*

Hoole, Philip 1999. Coarticulatory investigations of the devoicing gesture. In Hardcastle, William J. – Hewlett, Nigel (eds.) *Coarticulation. Theories, data and techniques.* Cambridge University Press, Cambridge, 107–121.

Jansen, Wouter 2004. *Laryngeal contrast and phonetic voicing: A laboratory phonology approach to English, Hungarian, and Dutch.* Groningen Dissertations in Linguistics 47. Rijksuniversiteit Groningen, Groningen.

Jansen, Wouter – Toft, Zoe 2002. On sounds that like to be pairs (after all): and acoustic investigation of Hungarian voicing assimilation. *SOAS Working Papers in Linguistics* 12. 19–52.

Jesus, Louis M. T. – Shadle, Christine H. 2003. Temporal and devoicing analysis of European Portuguese fricatives. In Solé, Maria-Josep – Recasens, Daniel – Romero, Joaqín (eds.): *Proceedings of the 15th International Congress of Phonetic Sciences. Barcelona 3-9 August 2003.* Universitat Autonoma de Barcelona, Barcelona, 779–782.

Kassai Ilona 1982. A magyar beszédhangok időtartamviszonyai. In Bolla Kálmán (szerk.): *Fejezetek a magyar leíró hangtanból.* Akadémiai Kiadó, Budapest, 115–154.

Markó Alexandra – Gráczi Tekla Etelka – Bóna Judit 2009. Zöngésségi hasonulás a spontán beszédben és a felolvasásban. Esettanulmányok. *Beszédkutatás 2009.* 5–27.
Ohala, John J. 1997. Aerodynamics of phonology. In: *Proceedings of 4th Seoul International Conference on Linguistics.* Seoul, 92–97.
Olaszy Gábor 2007. *Mássalhangzó-kapcsolódások a magyar beszédben.* Tinta Könyvkiadó, Budapest.
Shadle, Christine H. 1999. The aerodynamics of speech. In Hardcastle, William J. – Laver, John (eds.) *The handbook of phonetic sciences.* Blackwell Publishers, Oxford, 33–64.
Siptár Péter 2001. A mássalhangzók. In Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 2. Fonológia.* Akadémiai Kiadó, Budapest, 183–272.
Stevens, Kenneth N. 1998. *Acoustic phonetics.* The MIT Press, Cambridge, Massachusetts–London.
Stevens, Kenneth N. – Blumstein, Sheila E. – Glicksman, Laura – Burton, Martha – Kurowski, Kathleen 1992. Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters. *Journal of the Acoustical Society of America* 91/5. 2979–3000.
Szende Tamás 1973. *Spontán beszédanyag gyakorisági mutatói.* Nyelvtudományi Értekezések 81. Akadémiai Kiadó, Budapest.
Szigetvári, Péter 1998. Voice assimilation in Hungarian: the hitches. In Varga, László (eds.): *The Even Yearbook 3. ELTE SEAS Working Papers in Linguistics.* Eötvös Loránd University, Budapest.
Westbury, John R. – Keating, Patricia 1986. On the naturalness of stop consonant voicing. *Journal of Linguistics* 22. 145–166.

A kutatás a 78315 számú OTKA támogatásával készült.

A RÖVID-HOSSZÚ MAGÁNHANGZÓK REALIZÁCIÓI IDŐSEK SPONTÁN BESZÉDÉBEN

Bóna Judit

Bevezetés

A magyar magánhangzórendszerben fonológiailag hét rövid-hosszú magánhangzópár található (Nádasdy–Siptár 2001); fonetikailag azonban csak öt olyan magánhangzópár van, amelynek tagjai mindössze egy képzésjegyben, az időtartamban különböznek (Gósy 2004). Ezek a fonetikai párokba rendezhető, rövid és hosszú magánhangzók fonológiai funkciójuknak megfelelően megkülönböztetik a szavak jelentését: pl. *kint ~ kínt; koma ~ kóma; növel ~ nővel; furat ~ fúrat; tüzet ~ tűzet.* A nyelvileg rövid és a nyelvileg hosszú magánhangzók fizikai időtartamában lehet átfedés; a fonológiailag rövid–hosszú párok realizációi spektrálisan is különböznek. Nádasdy és Siptár (2001) fonológusok hallási benyomásuk alapján megállapítják, hogy a köznyelvi ejtésben olyan időtartambeli ingadozások tapasztalhatók, amelyek a rövid-hosszú oppozíció különbségének a csökkenését mutatják. Különösen igaznak vélik ezt a felső nyelvállású párokra; véleményük szerint valódi fonetikai és fonológiai rövid–hosszú szembenállás csak a középső nyelvállású magánhangzóknál, az *o/ó*[1] és az *ö/ő* párnál tapasztalható. Ugyanakkor a szerzők a magyar magánhangzók fonológiai szempontú elemzéséhez a kvantitatív megközelítést tartják adekvátnak, azaz elsősorban – az említett rövidülés tendencia ellenére – az időtartamban érzik a nyelvileg rövid és a nyelvileg hosszú magánhangzók közötti legfőbb különbséget.

A magánhangzók időtartamát fonetikailag számos tényező befolyásolja a különböző nyelvekben. Meghatározó például maga a beszédhang minősége, spektrális jellemzői, a hangsúly, az alaphangmagasság (Engstrand–Krull 1994; Kovács 2002); a hangidőtartamokat befolyásolhatja a közlés teljes hossza, a beszédstílus (spontán beszéd vagy felolvasás), a beszéd- és artikulációs tempó, a beszédhelyzet, a szógyakoriság, a beszélő egyéni tulajdonságai stb. (Crystal–House 1982; Bell et al. 2003; Gósy 2004). További tényezők még a magánhangzók időtartamának alakulásában a szomszédos beszédhangok, a magánhangzó helyzete a szótagban, a szótag típusa, a szó helyzete a mondatban, a magánhangzót megelőző és követő szótagok száma a mondatban stb. (van Santen 1992).

1. A fonémákat, illetve esetenként a beszédhangokat IPA-jelek helyett dőlt betűkkel jelenítjük meg.

A magyar magánhangzók időtartamára vonatkozóan számos kutatás ismeretes. Az első vizsgálatokat több mint száz évvel ezelőtt, 1909-ben végezte Gombocz és Meyer; később például Hegedűs Lajos (1941), Magdics Klára (1966), Szende Tamás (1976), Kassai Ilona (1979; 1982; 1991), Bolla Kálmán (1995), Kovács Magdolna (2002), Olaszy Gábor (2006), Bata Sarolta (2007), Mády Katalin (2008), Gósy Mária és Beke András (2010) elemezték különböző szempontok szerint a magánhangzókat. A kutatások eredményei nem összehasonlíthatók, mivel különböző módszerrel, különböző beszédstílusú anyagon (szavak felolvasása, mondatok felolvasása, spontán beszéd stb.) készültek.

Több fonetikai vizsgálatban elemezték a nyelvileg rövid és nyelvileg hosszú magánhangzók közötti időtartam-különbségeket; spontán beszédanyagra öt kutatás ismeretes. Hegedűs Lajos 1941-ben 2000 beszédhang időtartamát mérte meg. Kassai Ilona 1979-ben egy spontán rádióbeszélgetésben elemezte a hangidőtartamokat, három beszélőtől 921 db olyan magánhangzót elemzett amely fonetikailag is egy rövid-hosszú magánhangzópár tagja volt. A rövid magánhangzók időtartamát 40–100 ms között, a hosszúakét 60–145 ms között mérte. Olaszy Gábor (2006) egyetlen beszélő felolvasását elemezve azt találta, hogy bár van átfedés a nyelvileg rövid és hosszú hangok között, a nyelvileg hosszú magánhangzók fizikai időtartama átlagosan hosszabb, mint a rövid magánhangzók átlagos tartama; a rövid/hosszú arány 1 : 1,4. Olaszy (2006) egy rövid spontán rádióriportot is elemzett, amelyben négy beszélő szólalt meg; és megállapította, hogy a spontán beszédben adatolt rövid magánhangzók az *ö* kivételével kissé rövidebbek a felolvasásban mért értékeknél. Bata Sarolta (2007) négy adatközlő spontán beszédében vizsgálta a magánhangzók időtartamait (az anyagban 559 olyan magánhangzó szerepelt, amely valamely fonetikai magánhangzópár egy tagja volt). Az ő eredményei is megerősítik azt, hogy a hosszú magánhangzók átlagos időtartama hosszabb, mint a rövid magánhangzóké. Az említett kutatások objektív mérési eredményeket közöltek, és igazolták, hogy van különbség a nyelvileg rövid és a nyelvileg hosszú magánhangzók fizikai időtartamában, de azt nem elemezték, hogy statisztikailag van-e különbség a rövid és a hosszú hangok között.

Az eddigi legátfogóbb, legtöbb magánhangzót érintő és statisztikai vizsgálatot is tartalmazó kutatást a magyar nyelvre vonatkozóan Gósy Mária és Beke András végezték 2010-ben. Kutatásukban 10 fiatal beszélő spontán beszédében 5258 magánhangzó időtartamát elemezték, ebből 4362 fonológiailag rövid és 896 fonológiailag hosszú hangét. Az adataik egyrészt megerősítették a korábbi kutatások eredményeit, azaz igazolták a fizikai időtartam különbségét a fonológiailag rövid-hosszú magánhangzópárok mindegyikénél; másrészt további fontos megállapításokat tettek arra vonatkozóan, hogy milyen paraméterek befolyásolják még az időtartam-különbségeket. Eszerint a legfontosabb tényező a hangidőtartamok alakulásában a fonológiai időtartam, ezután következik a nyelv függőleges mozgása, majd kevésbé meghatározó a

nyelv vízszintes mozgása és az ajakműködés, végül a legkevésbé hat az első szótagi pozíció (Gósy–Beke 2010). Az objektív, akusztikai fonetikai vizsgálatok tehát nem igazolták statisztikailag a fonológusok azon benyomását, hogy a hosszú hangok rövidülnek. Ez egyes szavakban igaz lehet, de nagyszámú spontán beszédbeli adat szerint nem érvényesül ez a feltételezés.

Az első szótagi pozíció a hangsúlyos helyzettel áll összefüggésben. A magyar nyelv szabályai szerint, ha van hangsúly, ez az első szótagra esik (Gósy 2004). Az esetek döntő többségében így van ez a spontán beszédben is, de a pozíció nem jelenti azt, hogy valóban megjelenik a hangsúly az első szótagon. A korábbi, akár több évtizeddel ezelőtti szakirodalomban ellentmondó adatok olvashatók a hangsúlyos helyzet és a magánhangzók időtartamának összefüggéseire vonatkozóan. Fónagy Iván és Baráth Judit azt találták (1966), hogy a hangsúlyos szótagban általában hosszabb a magánhangzó, mint a hangsúlytalanban. Hasonló eredményre jutott Magdics Klára (1966), Olaszy Gábor (1994) és Kassai Ilona is (1979). Ezzel szemben Kovács Magdolna kísérletének (2000) tanúsága szerint a rövid magánhangzók időtartama a hangsúlyos szótagokban átlagosan 10%-kal volt rövidebb a szóbelsei és a szóvégi hangsúlytalan magánhangzókhoz képest. Szintén Kovács kutatásaiban, nagyobb elemszámú korpusz vizsgálata azt mutatta, hogy a hangsúlyos és a hangsúlytalan magánhangzókat nem állítja szembe a fizikai időtartamuk (Kovács 2002). Gósy és Beke (2010) hasonló eredményeiket azzal indokolják, hogy a spontán beszéd hangsúlyviszonyait főként a tételhangsúlyok, az érzelmi hangsúlyok és az aktuális beszédtervezési sajátosságok határozzák meg. A magánhangzó-időtartamokat befolyásoló tényezők még a hangkörnyezet (vö. Kovács 2000; 2002), a hangsor hossza (a hosszabb hangsorokban rövidebbek a magánhangzók, vö. Kassai 1993), a hangnak a hangsorban elfoglalt helye, az artikulációs tempó (a gyors beszédben csökken a különbség a rövid és a hosszú magánhangzók időtartama között, vö. Kassai 1993) stb.

Az 1. táblázatban a magyar rövid-hosszú oppozícióban álló magánhangzók időtartamára kapott szakirodalmi adatok olvashatók.

A nyelvileg rövid-hosszú magánhangzópárok esetén felmerül a kérdés, hogy a rövid és a hosszú magánhangzópárok célkonfigurációja azonos-e, avagy nem csak az időtartamban különböznek (Kovács 2002; Mády 2008).

Kovács Magdolna 430 hangsúlyos pozícióban álló magánhangzó formánsszerkezeti adatait vizsgálta, és azt találta, hogy a rövid–hosszú magánhangzópárok spektrális szembenállásának mértéke függ a nyelvállás fokától. Az eltérés nagysága szerint a magánhangzópárok sorrendje a következő: *o/ó* > *ö/ő* > *u/ú* > *i/í* > *ü/ű* (Kovács 2002). A percepciós vizsgálatok azt mutatják, hogy a spektrális minőség másodlagos felismerési kulcs a magyar középső nyelvállású, rövid–hosszú oppozícióban álló magánhangzópárok azonosításában (Kovács 1998).

1. táblázat: Az átlagos magánhangzó-időtartamok a szakirodalomban (ms)

Mgh	Olaszy (2006) felolvasás, 1 beszélő, 480 mondat, 1913 beszédhang	Bata (2007) spontán beszéd (társalgás), 4 beszélő, 559 beszédhang	Gósy–Beke (2010) spontán beszéd (narratíva), 10 beszélő, 5258 beszédhang
i	69	70	57
í	87	–	73
o	81	69	58
ó	116	103	82
ö	78	83	63
ő	105	137	73
u	75	65	52
ú	106	90	71
ü	76	78	50
ű	101	–	62

Mády Katalin (2008) két beszélő ejtésében vizsgálta a magyar magánhangzókat kétdimenziós elektromágneses artikulográffal. Az adatközlők kétféle tempóban (gyorsan és átlagosan) olvastak fel álszavakat tartalmazó mondatokat, amelyekben szimmetrikus [p, t, k] környezetben szerepelt a 14 magánhangzó. Egy magánhangzó a női beszélő ejtésében a három mássalhangzó-környezetben és a kétféle tempóban nyolcszor, a férfi beszélő felolvasásában tízszer fordult elő. Az eredmények azt mutatták, hogy a felső nyelvállású magánhangzók esetén nincs jelentős különbség az artikulációban a rövid és a hosszú hangok között. A középső nyelvállású ajakkerítéses hangoknál *(o/ó* és *ö/ő)* azonban a nyelv helyzete egyértelműen eltér a kétféle nyelvi időtartamú hang ejtésekor; az ajakműködés pedig a hosszú hangzók képzésekor általában intenzívebb, mint a rövid hangok artikulációjakor.

A nyelvileg rövid-hosszú magánhangzók vizsgálata a magyarra vonatkozóan eddig főként a fiatal és középkorú felnőttek beszédére korlátozódott (Gombocz 1909; Magdics 1966; Kassai 1979; Kassai 1991; Kovács 1998; 2000; 2002; Olaszy 2006; Bata 2007; Gósy–Beke 2010); a gyermeknyelvi magánhangzó-időtartamokról csak sporadikus vizsgálatok vannak (vö. Gósy 1984; Bóna–Imre 2010; holland–magyar összevetésben van der Stelt et al. 2005); az időskori beszédben előforduló rövid-hosszú magánhangzókról pedig nincsenek akusztikai fonetikai adataink a magyarban. Holott a hangidőtartamokat és a spektrális szerkezetet befolyásolja az artikulációs és a beszédtempó, amely a különböző életkorokban szignifikáns különbséget mutat (Gósy 1997; Gocsál 2000; Menyhárt 2000; Bóna 2010); illetőleg az artikuláció pontossága is, ami szintén változik az életkor előrehaladtával.

Az anyanyelv-elsajátítás során a magyar gyermekek beszédhangképzése az 5-6. életévben stabilizálódik; ennek hátterében a beszédhang-differenciálás stabil működése áll. Egyedül a rövid-hosszú fonológiai oppozícióban álló be-

szédhangok fizikai időtartamának megbízható elkülönülése tehető későbbre, 6-7 éves korra a beszédpercepcióban (Gósy 2006). 5-6 éves óvodások és 9 éves kisiskolások magánhangzóinak elemzése azt mutatta, hogy amíg az óvodásoknál a rövid–hosszú oppozícióban álló magánhangzók időtartama nem mutat stabil elkülönülést minden magánhangzó esetében (csak az *o/ó* és az *u/ú* hangoknál, ami a gyakorisággal és az anyanyelv-elsajátításban korábbi megjelenéssel magyarázható), addig a 9 évesek beszédében a fizikai időtartamok egyértelműen definiálják a fonológiailag rövid–hosszú magánhangzókat. A formánsszerkezetre vonatkozóan csak egyetlen esetben mutatkoztak eltérések, az *o/ó* magánhangzópárnál az óvodások ejtésében az F_1 és az F_2 is, a kisiskolások beszédprodukciójában pedig csak az F_2 mutatott szignifikáns különbséget a rövid-hosszú hangok között (Bóna–Imre 2010).

Amíg a gyermekek ejtése az artikulációs gyakorlatlanság miatt tér el a felnőttek ejtésétől, addig az idősek artikulációja anatómiai és fiziológiai változások miatt válik pontatlanabbá (vö. Balázs 1993; Torre–Barlow 2009). Ez a beszédhangok akusztikai fonetikai sajátosságaiban is tetten érhető. A magánhangzók spektrális szerkezetének megváltozását az életkor előrehaladtával több nyelvre is igazolták. Egy amerikai angol anyanyelvűekkel végzett kutatásban kapott eredmények azt mutatták, hogy az F_1 szignifikánsan eltér az idősek és a fiatalok ejtése között minden magánhangzónál; míg az F_2-ben a hátul képzett magánhangzók esetén találtak szignifikáns eltérést (Watson–Munson 2007). Egy másik, szintén amerikai angol anyanyelvűekkel végzett kutatás az életkori változásokat a nemek, illetve a hallásállapot függvényében vizsgálta. 27 fiatal és 59 idős személy beszédében elemezték az alapfrekvenciát, az első három formánst, illetve a zöngekezdési időt (VOT). A nemek és életkorok szerint differenciált csoportok között az F_0, az F_1 és a VOT értékeiben találtak szignifikáns eltérést, de az egyes paraméterek erősen függtek a vizsgált beszédhangok minőségétől is (Torre–Barlow 2009).

A magyarra vonatkozóan *a, á, e, é, i, o, ö, u* hangok formánsértékeit elemeztük fiatal és idős (70 év feletti) női beszélők ejtésében. Az eredmények szerint az idősek kisebb frekvenciatartományban, illetőleg valamivel hátrébb képzik a magánhangzókat. Az előbbi jelenti, hogy kevésbé különítik el az ejtés során a különböző magánhangzó-minőségeket. Az *ö* és az *u* kivételével a többi vizsgált beszédhang esetében – a magánhangzó képzési jegyeitől függően – szignifikáns eltérést találtunk az egyes formánsértékek között a két életkori csoport ejtésében (Bóna 2009). Ezek az eltérések a formánsértékekben összefüggésben lehetnek az időzítéssel is. Egyrészt az ejtés pontossága befolyásolhatja az időtartamot, másrészt az időtartam hatással lehet a képzési konfigurációra is.

A magyarra vonatkozóan tehát nincsenek adataink a magánhangzó-időtartamok realizációjáról az időskori beszédben, illetőleg arról, hogy az idősek kevésbé pontos artikulációja hogyan befolyásolja a rövid-hosszú magánhangzók akusztikai sajátosságait. A jelen kutatásunk fő kérdései a követ-

kezők: 1. Hogyan realizálódnak a rövid-hosszú magánhangzópárok az idősek spontán beszédében? 2. Különböznek-e a rövid-hosszú magánhangzóik az időtartamban és a spektrális szerkezetben? 3. Miben térnek el a magánhangzóik a fiatalokéihoz képest? A tanulmányban először közlünk adatokat az idősek magánhangzóinak időtartamáról, illetve először mutatjuk be nagyszámú és spontánbeszédbeli magánhangzón a rövid-hosszú magánhangzók spektrális szerkezetének hasonlóságait és különbségeit. Hipotéziseink a következők. 1. A szignifikánsan lassabb artikulációs tempó miatt nincs akkora különbség a rövid és a hosszú hangok között az idősek beszédében, mint az átlagos artikulációs tempójú fiataloknál. 2. Az idősek pontatlanabb artikulációja miatt a spektrális sajátosságokban nincs szignifikáns különbség a rövid-hosszú hangok között.

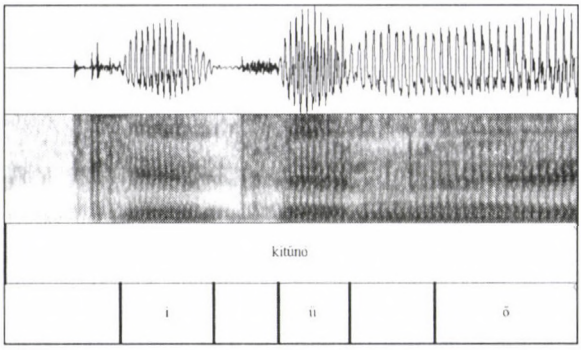
Anyag, módszer, kísérleti személyek

A kutatáshoz a BEA adatbázisból (Gósy 2008) választottuk ki nyolc idős (életkoruk 70–80 év) és nyolc fiatal (életkoruk 20–32 év) női adatközlő spontán narratíváját. Beszélőnként mintegy 5 percet; összesen 80 percnyi hanganyagot elemeztünk.

A hangfelvételekben a Praat 5.0 szoftverrel annotáltuk a vizsgálandó magánhangzókat: *i/í, o/ó, ö/ő, u/ú, ü/ű*, illetve a beszédszakaszokat (1. ábra). Kiszámítottuk az artikulációs tempót; majd elemeztük a magánhangzókat. Egy erre a célra írt szkript segítségével kiírattuk Excel-táblázatba a hangidőtartamokat, illetve az F_1 és F_2 értékeit. A szkript a magánhangzó tiszta fázisának a közepén mérte a formánsértékeket. A hangidőtartamok mérése ugyanakkor több problémát is felvet, hiszen a koartikulációs hangátmenetek megnehezítik az adott hang határának pontos kijelölését. Ezekben az esetekben egyrészt az auditív, másrészt a vizuális jelre támaszkodva jelöltük ki a szegmentum határát; amikor azonban nem lehetett egyértelműen meghatározni a hanghatárt, a hangátmenet közepén jelöltük azt ki. Az esetleges mérési hibák miatt az adatokat kézzel ellenőriztük.

A hanganyagban összesen 5520 olyan magánhangzót adatoltunk, amelynek volt nyelvileg rövid vagy nyelvileg hosszú párja. Ez 4458 rövid és 1062 hosszú magánhangzót jelentett; az időseknél 2536, a fiataloknál 2984 magánhangzót elemezhettünk. A különbség adódhat abból, hogy a fiatalok ugyanannyi idő alatt több hangot ejtettek, illetőleg a beszédükben előforduló szavak, hangok gyakoriságából is. Az egyes magánhangzók előfordulásának száma a 2. táblázatban olvasható. A legtöbb előfordulást a rövid *o* hangnál, a legkevesebbet a hosszú *ű*-nél adatoltuk mindkét életkori csoportban. A hosszú *ű* annyira ritka, hogy összesen mindössze 41-szer fordult elő a vizsgált hanganyagban, 21-szer a fiataloknál, 20-szor az időseknél.

Az adatokon statisztikai elemzést végeztünk az SPSS 13.0 szoftverrel (egytényezős varianciaanalízis és Tukey post hoc teszt) 95%-os konfidenciaszinten.



1. ábra
A *kitűnő* szó magánhangzóinak annotálása

2. táblázat: A vizsgált magánhangzók előfordulása életkoronként (a nyelvileg elvárt formát számoltuk a darabszámba, függetlenül a realizációtól)

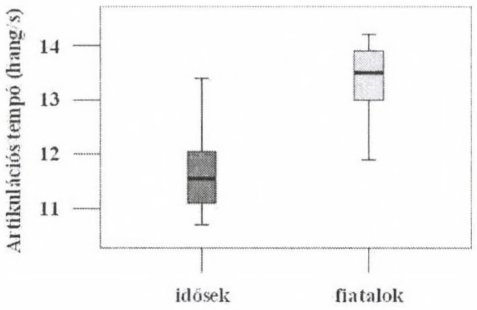
	i	í	o	ó	ö	ő	u	ú	ü	ű
Fiatalok	805	134	1062	162	218	173	205	86	118	21
Idősek	737	111	876	149	146	148	196	58	93	20

Eredmények

Az eredmények bemutatásakor először az artikulációs tempót tárgyaljuk, mivel ez az egyik legfontosabb befolyásoló tényező a magánhangzók akusztikai szerkezetének alakulásában. Az időtartamok és a formánsértékek bemutatásakor először megvizsgáljuk az egyes életkorok jellemzőit, majd összevetjük a két életkori csoportban kapott adatokat.

Az artikulációs tempó

Az artikulációs tempó adatai alátámasztották a korábbi eredményeinket (Bóna 2010). Az idősek szignifikánsan lassabb tempóval beszéltek, mint a fiatalok (2. ábra): $F(1, 14) = 17{,}678$; $p \leq 0{,}001$. Az idősek átlagos artikulációs tempója (11,7 hang/s) a fiatalokénál (13,4 hang/s) 1,7 hang/s-mal lassabb volt. Három olyan idős adatközlő volt, aki gyorsabb tempóval (12,0–13,4 hang/s) artikulált, mint a leglassabban beszélő fiatal (11,9 hang/s; a leglassabb artikulációs tempóra az időseknél 10,7 hang/s-ot számítottunk). Ez utóbbi (leglassabb artikulációjú) fiatalt leszámítva a fiatalok artikulációs tempója legalább 13,0 hang/s volt. Az időseknél nagyobb szórást adatoltunk az artikulációs tempó értékeiben (0,9 hang/s), mint a fiataloknál (0,7 hang/s).



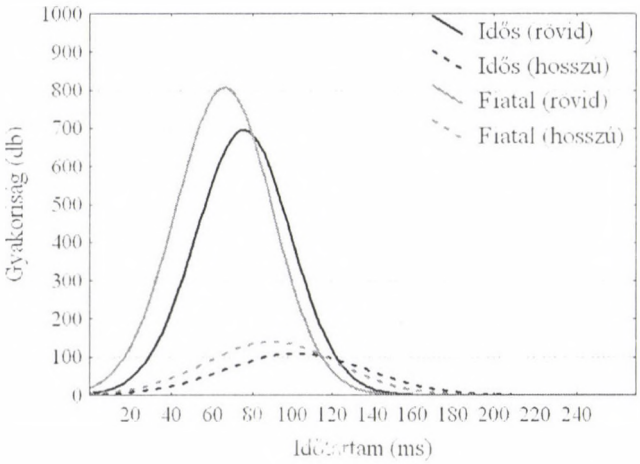
2. ábra
Az artikulációs tempó a két életkori csoportban (hang/s)

A hangidőtartamok

A hangidőtartamok eloszlása a 3. ábrán látható. Megfigyelhető, hogy nagy átfedések vannak mind a rövid-hosszú beszédhangok, mind a különböző életkorú beszélők ejtése között. Ennek ellenére azonban van különbség a rövid-hosszú hangidőtartamokban. Az idősek vizsgált rövid magánhangzóinak átlagos időtartama 75 ms (szórás: 23 ms) volt; a hosszú hangjaik átlaga 101 ms (szórás: 35 ms). A fiatalok rövid magánhangzói átlagosan 66 ms tartamúak voltak (szórás: 24 ms); a hosszú hangjaik 89 ms-osak (szórás: 33 ms). Ez utóbbi adatok gyakorlatilag megegyeznek Gósy Mária és Beke András 2010-es adataival; ők hasonló artikulációs tempójú fiatalok (nők és férfiak) beszédében vizsgálták a magánhangzókat. Az idősek átlagosan hosszabb magánhangzókat ejtettek a szignifikánsan lassabb artikulációs tempó miatt, mint a fiatalok, ugyanakkor a hangidőtartamok szórásában nem volt különbség a két életkori csoport között. Az idősek rövid hangjainak a szórása 23 ms volt, a hosszú magánhangzóiké 35 ms. A fiatalok rövid magánhangzóinak a szórása 24 ms, a hosszú hangzóiké 33 ms. Az, hogy a hosszú hangok időtartamainak szórása mindkét életkorban nagyobb volt, mint a rövid hangoké, azt mutatja, hogy ezeket a hangokat kisebb artikulációs biztonsággal ejtették a beszélők. Az összes rövid-hosszú magánhangzó időtartamára elvégzett statisztikai próba szignifikáns különbséget mutatott mind a rövid-hosszú párok, mind az életkori csoportok között: $F(3, 5517) = 327{,}193$; $p \leq 0{,}001$; a Tukey post hoc teszt szerint mind a négy magánhangzócsoport között (idősek rövid, idősek hosszú, fiatalok rövid, fiatalok hosszú hangjai) $p \leq 0{,}001$.

Elvégeztük az időtartamok elemzését hangpáronként is mindkét életkorban (2. táblázat). Gósy és Beke (2010) kutatásukban a következő sorrendet állították fel a hangidőtartamok alapján: a rövid magánhangzók esetén *ü* < *u* < *i* < *o* < *ö;* a hosszú magánhangzóknál *ű* < *í* < *ő* < *ú* < *ó*. A saját adataink ettől némileg különböznek mind az időseknél, mind a fiataloknál. A különbség okára

a hanggyakoriság, a szavak szótagszáma, a fonetikai kontextus és a spontán beszéd változó temporális viszonyai adhatnak magyarázatot; illetve az is, hogy nem azonos szavak fordultak elő a beszélőknél, a fonetikai pozíció a két kutatás anyagában nem feltétlenül volt azonos stb. Ugyanakkor több hasonlóság is megfigyelhető Gósy és Beke (2010) és a saját sorrendünk között: a rövid magánhangzók közül az *ö*, majd az *o* időtartama a leghosszabb; míg a hosszú hangoknál az *ű* a legrövidebb magánhangzó életkortól függetlenül, mindkét kutatásban. Az időseknél a rövid magánhangzók sorrendje az átlagidőtartamok alapján a következőképpen alakult: *i* = *u* = *ü* < *o* < *ö*. A hosszú magánhangzók sorrendje: *ű* < *ú* < *í* < *ó* < *ő*. A fiatalok rövid magánhangzóinak a sorrendje: *i* < *ü* < *u* < *o* < *ö*. A fiatalok hosszú magánhangzóinak a sorrendje: *ű* < *í* < *ú* < *ő* < *ó*.



3. ábra
Az összes magánhangzó időtartama a fonológiai hosszúság és az életkor függvényében

A statisztikai elemzés szerint egy kivétellel mindkét életkori csoportban szignifikáns különbség volt a magánhangzópárok rövid és hosszú hangjainak időtartama között ($p \leq 0{,}007$). A kivételt az *ü/ű* hangpár jelentette az idősek ejtésében; azt, hogy ebben az esetben nem volt szignifikáns a különbség, az *ü/ű* hangok ritka előfordulása is okozhatta.

Nem volt szignifikáns különbség az idősek és a fiatalok által ejtett hosszú *ó* hangok időtartamában, illetve a hosszú *ú* hangokéban sem; azaz az idősek és a fiatalok hasonló időtartamban ejtették ezeket a magánhangzókat. Az idősek rövid *ö* hangjainak időtartama statisztikailag nem különbözött a fiatalok *ő* idő-

tartamaitól. Az idősek hosszú *ű*-je nem mutatott szignifikáns különbséget a fiatalok *ü/ű* hangjainak időtartamától. Ennek az *ű* hang ritka előfordulása mellett a kérdéses hangot tartalmazó szavak fonológiai felépítése is lehetett az oka. Mindkét életkori csoport spontán beszédében előfordult az *egyszerűen* szó. Az idősek beszédében a hosszú *ű*-t tartalmazó szavak egy részében az *ű* gyakran rövidült: *törvényszerűség, törvényszerűen, sugárszerűen, törvényszerű, kitűnően.* Ugyanakkor mindkét életkori csoport beszédében voltak olyan szavak, amelyekben nem rövidült a hosszú *ű*. Az időseknél ilyen szavak voltak például a *szűken, tűrőképességem, eltűrni, hűségesebbek, műszaki, nyelvű;* a fiataloknál pedig olyanok fordultak elő, mint például *szellemű, sűrűn, egyszerű, műve, betűtornyozunk, működik.* Azokban a szavakban, ahol röviden ejtették az *ű*-t, a rövidülés oka valószínűleg a szóhosszúságban, a magánhangzó-kapcsolat könnyebb ejtésében és a szógyakoriságban keresendő.

2. táblázat: A fonológiailag rövid-hosszú magánhangzók átlagos időtartama és szórása életkoronként

A magánhangzó-minőség	A magánhangzók időtartama (ms)			
	Fonológiailag rövid		Fonológiailag hosszú	
	Átlag	Szórás	Átlag	Szórás
Idősek				
i/í	71	23	95	32
o/ó	78	23	105	34
ö/ő	85	23	109	38
u/ú	71	24	93	34
ü/ű	71	22	72	18
Fiatalok				
i/í	60	20	83	30
o/ó	69	26	93	36
ö/ő	71	22	92	31
u/ú	64	24	87	32
ü/ű	62	23	80	31

A spektrális szerkezet

Az egyes magánhangzók F_1-re és F_2-re kapott átlagos formánsértékeit a 3. táblázatban tüntettük fel. A formánsértékek összevetése szerint az idősek kevésbé különböztetik meg spektrálisan a rövid-hosszú magánhangzópárok tagjait, mint a fiatalok. Ez azt jelenti, hogy az idősek ejtésében kevésbé különbözik a fonetikai értelemben vett magánhangzópárok rövid és hosszú tagjának a formánsszerkezete, mint a fiatalok beszédprodukciójában. Az idősek beszédében nagyobb a variabilitása az egyes magánhangzók ejtésének, mint a fiataloknál; azaz az idősek kevésbé stabilan képzik az egyes magánhangzókat. A 4. és az 5. ábrán például megfigyelhetjük, hogy az idősek *o/ó* hangjai

sokkal szélesebb frekvenciatartományban realizálódtak, mint a fiatalokéi. Ennek a már említett artikulációs változások (Balázs 1993), a fonetikai kontextus, a tempóértékek stb. állhatnak a hátterében.

3. táblázat: A fonológiailag rövid-hosszú magánhangzók átlagos formánsértékei és (zárójelben) szórásuk életkoronként

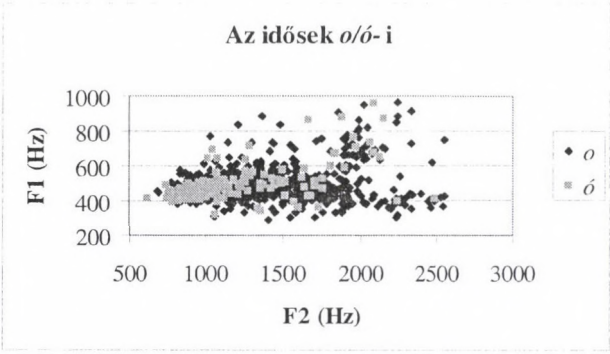
A magánhangzó-minőség	A magánhangzók formánsai (Hz)			
	Fonológiailag rövid		Fonológiailag hosszú	
	F_1	F_2	F_1	F_2
Idősek				
i/í	451 (190)	2204 (353)	409 (113)	2317 (307)
o/ó	508 (201)	1379 (468)	506 (135)	1230 (406)
ö/ő	503 (161)	1736 (253)	458 (75)	1744 (220)
u/ú	435 (156)	1392 (406)	436 (238)	1364 (493)
ü/ű	494 (250)	1863 (316)	405 (37)	1773 (143)
Fiatalok				
i/í	416 (68)	2315 (293)	397 (51)	2404 (291)
o/ó	519 (89)	1391 (287)	476 (62)	1216 (357)
ö/ő	499 (68)	1724 (167)	470 (110)	1761 (179)
u/ú	403 (52)	1276 (314)	390 (48)	1257 (335)
ü/ű	389 (43)	1682 (288)	383 (41)	1919 (211)

Az időseknél az *o/ó; ü/ű* és az *u/ú* hangpárok között nem volt spektrális különbség. Az idősek ejtésében az *i/í* hangpár tagjai mind az F_1-ben (p = 0,012); mind az F_2-ben (p = 0,003) szignifikánsan különböztek; az *ö/ő* esetén pedig (csak) az F_1 mutatott szignifikáns különbséget a fonológiailag rövid-hosszú hangok között (p = 0,002).

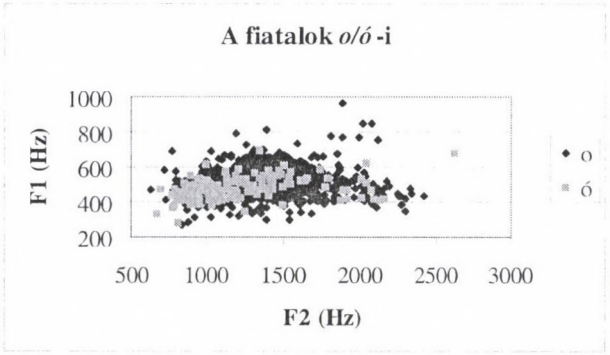
A statisztikai próba a fiataloknál csak az *u/ú* hangpár között nem mutatott szignifikáns különbséget a formánsértékeket illetően. A fiatalok *i/í* hangjai az F_2-ben különböztek szignifikánsan (p = 0,014). Ugyanők *o/ó* hangjai mindkét formánsértékben szignifikáns különbséget mutattak ($p \leq 0{,}001$). Az *ö/ő* esetén – hasonlóan az idősekhez – az F_1 mutatott szignifikáns különbséget a fonológiailag rövid-hosszú hangok között ($p \leq 0{,}033$). Az *ü/ű* esetén a fiataloknál az F_2-ben volt statisztikailag szignifikáns különbség (p = 0,001).

Összevetettük a fiatalok és az idősek magánhangzó-realizációinak formánsait is. Mindegyik magánhangzó-minőség tekintetében volt valamilyen különbség a két életkori csoport között. Az *i/í* hangpárnál szignifikáns különbség volt az F_1-értékekben a fiatalok és az idősek rövid hangjai között ($p \leq$ 0,001); a fiatalok hosszú *í* hangjainak és az idősek rövid *i*-jeinek F_1-ei között ($p \leq 0{,}001$); illetve az F_2-ben a két életkori csoport rövid hangjai között ($p \leq$ 0,001), továbbá a fiatalok hosszú és az idősek rövid hangjai között ($p \leq$

0,001). Az *o/ó* hangpár esetén az F_1 a fiatalok hosszú és az idősek rövid hangjaiban különbözött szignifikánsan ($p = 0,013$); míg az F_2 a fiatalok rövid és az idősek hosszú hangjaiban mutatott szignifikáns különbséget ($p = 0,027$).



4. ábra
Az F_1- és az F_2-értékek az idősek *o/ó* hangjaiban



5. ábra
Az F_1- és az F_2-értékek a fiatalok *o/ó* hangjaiban

Az *ö/ő* hangok ejtésekor az F_1-ben volt csak szignifikáns különbség a fiatalok rövid és az idősek hosszú ($p = 0,002$), illetve a fiatalok hosszú és az idősek rövid hangjai között ($p = 0,027$). Az *u/ú* esetén a fiatalok hosszú és az idősek rövid hangjainak F_1-eiben mutatott a statisztikai próba szignifikáns különbséget ($p = 0,032$); illetőleg az F_2-ben az idősek rövid hangjai és a fiatalok mind rövid ($p = 0,011$), mind hosszú hangjai között ($p = 0,028$). Végezetül az *ü/ű* hangpár vizsgálatakor azt találtuk, hogy a fiatalok és az idősek rö-

vid hangjai mind az F_1-ben ($p = 0{,}032$), mind az F_2-ben ($p \leq 0{,}001$) szignifikánsan különböznek.

Következtetések

Kutatásunkban a rövid-hosszú magánhangzópárok realizációit vizsgáltuk az időtartam és a spektrális minőség szempontjából két életkori csoportban. Az eredményeink szerint az idősek szignifikánsan lassabb artikulációs tempóval beszéltek, és szignifikánsan hosszabb magánhangzókat ejtettek, mint a fiatalok. Statisztikai különbséget találtunk az összes rövid és az összes hosszú magánhangzó időtartamában is mindkét életkori csoportban. A hangonkénti elemzés során a kiinduló hipotéziseink azonban csak részben igazolódtak. Az eredmények azt mutatták, hogy az idősek is megkülönböztetik a fizikai időtartam szerint a nyelvileg rövid és a nyelvileg hosszú magánhangzókat (a nagyon ritka *ü/ű* kivételével). Ugyanakkor az időseknél mért szignifikánsan lassabb artikulációs tempó hatása, ami időt biztosítana az adott magánhangzó elvárt képzési konfigurációjának létrehozására, nem mutatkozik meg a spektrális szerkezetben. Azaz önmagában a tempó nem határozza meg a hangidőtartamokat, illetve az ejtés pontosságát. Az adatok azt igazolják, hogy az életkor előrehaladtával nehezítetté váló nyelvmozgás, a renyhülő izomműködés, a zörejesebb zöngeképzés miatt a rövid és a hosszú hangok spektrális szerkezete nem különül el úgy, mint a fiataloknál.

Kutatásunk elején azt vártuk, hogy a pontatlanabb artikuláció és a köznyelvi átlagos artikulációs tempónál szignifikánsan lassabb tempó a rövid-hosszú oppozícióban álló magánhangzók fizikai különbségének semleges(ebb)é válásához vezet mind az időtartamokat, mind a spektrális minőséget tekintve. Ez az időtartamokra nem érvényesült. Az objektív adatok azt mutatják, hogy a nyelvileg rövid és a nyelvileg hosszú magánhangzók időtartamának megkülönböztetésére való törekvés olyan erős, hogy nemcsak a fiatalokra (sőt már a kilencévesekre is) igaz, hanem a hetven év feletti idősekre is. Mivel a magyarban fonológiai funkcióval bírnak a hangidőtartamok, és a különböző életkorokban az ejtés pontossága nem mindig éri el a célkonfigurációt, a rövid-hosszú hangok eltéréseit a beszélők a fizikai időtartammal tudják realizálni.

Irodalom

Balázs Boglárka 1993. Az időskori hangképzés jellemzői. *Beszédkutatás '93*. 156–165.

Bata Sarolta 2007. Időszerkezeti mintázatok a spontán beszédben. Szakdolgozat. ELTE, Budapest.

Bell, Alan – Jurafsky, Daniel – Fosler-Lussier, Eric – Girand, Cynthia – Gregory, Michelle – Gildea, Daniel 2003. Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America* 113. 1001–1024.

Boersma, Paul – David Weenink 1998. Praat: doing phonetics by computer (Version 5.0.1). http://www.fon.hum.uva.nl/praat/download_win.html.

Bolla Kálmán 1995. *Magyar fonetikai atlasz.* Nemzeti Tankönyvkiadó, Budapest.

Bóna Judit 2009. Az idős életkor tükröződése a magánhangzók ejtésében. *Beszédkutatás 2009.* 76–87.

Bóna Judit 2010. Beszédtervezési folyamatok az életkor és a beszédstílus függvényében. *Magyar Nyelvőr* 134. 332–341.

Bóna Judit – Imre Angéla 2010. A rövid-hosszú magánhangzók óvodás és kisiskolás gyermekek beszédprodukciójában. In Navracsics Judit (szerk.): *Nyelv, beszéd, írás. Pszicholingvisztikai tanulmányok I.* Segédkönyvek a nyelvészet tanulmányozásához 107. Tinta Könyvkiadó, Budapest. 49–56.

Crystal, Thomas H. – House, Arthur S. 1982. Segmental durations in connected-speech signals: Preliminary results. *Journal of the Acoustical Society of America* 72. 705–16.

Engstrand, Olle – Krull, Diana 1994. Durational correlates of quantity in Swedish, Finnish and Estonian: Cross-language evidence for a theory of adaptive dispersion. *Phonetica* 51. 80–91.

Fónagy Iván – Baráth Judit 1966. Időtartam és hangosság. *Nyelvtudományi Közlemények* 68. 115–124.

Gocsál Ákos 2000. A beszéd időviszonyai különböző életkorú személyeknél. *Beszédkutatás 2000.* 39–50.

Gombocz Zoltán 1909. A magyar hangok időtartamáról. *Nyelvtudomány* 2. 93–100.

Gombocz Zoltán – Meyer, E. A. 1909. *Zur Phonetik der ungarischen Sprache.* Edv. Berlings Buchdruckerei, Uppsala.

Gósy Mária 1984. *Hangtani és szótani vizsgálatok hároméves gyermek nyelvében. Nyelvtudományi Értekezések* 102. Akadémiai Kiadó, Budapest.

Gósy Mária 1997. A magyar beszéd tempója és a beszédmegértés. *Magyar Nyelvőr* 121. 129–139.

Gósy Mária 2004. *Fonetika, a beszéd tudománya.* Osiris Kiadó, Budapest.

Gósy Mária 2006. A beszédhangok megkülönböztetésének fejlődése. *Beszédkutatás 2006.* 147–159.

Gósy Mária – Beke András 2010. Magánhangzó-időtartamok a spontán beszédben. *Magyar Nyelvőr* 134. 140–165.

Hegedűs Lajos 1941. Elektro-akusztikai berendezések a beszéd- és nyelvjáráskutatás szolgálatában. *Magyar Nyelv* 27. 162–185.

Kassai Ilona 1979. *Időtartam és kvantitás a magyar nyelvben.* Nyelvtudományi Értekezések 102. Akadémiai Kiadó, Budapest.

Kassai Ilona 1982. A magyar beszéd időtartamviszonyai. In Bolla Kálmán (szerk.): *Fejezetek a magyar leíró hangtanból.* Akadémiai Kiadó, Budapest, 115–154.

Kassai Ilona 1991. Hol tart a magyar magánhangzók időtartamának ingadozása? Szociolingvisztikai próbainterjúk tanulságai. *Magyar Nyelvőr* 115. 69–80.

Kassai Ilona 1993. Gyorsult-e a magyar beszéd tempója az elmúlt 100-120 évben? *Beszégkutatás '93.* 62–69.

Kovács Magdolna 1998. A spektrális minőség és az időtartam szerepe a magánhangzók percepciójában. *Beszédkutatás '98.* 35–46.

Kovács Magdolna 2000. Beszédhangok kontextusfüggő időviszonyai. *Beszédkutatás 2000.* 15–25.

Kovács Magdolna 2002. *Tendenciák és szabályszerűségek a magánhangzó-időtartamok produkciójában és percepciójában.* Doktori Értekezések 16. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen.

Magdics Klára 1966. A magyar beszédhangok időtartama. *Nyelvtudományi Közlemények* 68. 125–139.

Mády Katalin 2008. Magyar magánhangzók vizsgálata elektromágneses artikulográffal normál és gyors beszédben. *Beszédkutatás 2008.* 52–66.

Menyhárt Krisztina 2000. A beszéd temporális sajátosságai kétnyelvűeknél (kisiskoláskortól időskorig). *Beszédkutatás 2000.* 51–62.

Nádasdy Ádám – Siptár Péter 2001. A magánhangzók. In Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 2. Fonológia.* Akadémiai Kiadó, Budapest. 42–94.

Olaszy Gábor 1994. Hangidőtartam-módosító kísérletek a gépi beszéd ritmusának javításához. *Beszédkutatás '94.* 140–151.

Olaszy Gábor 2006. *Hangidőtartamok és időszerkezeti elemek a magyar beszédben.* Nyelvtudományi Értekezések 155. Akadémiai Kiadó, Budapest.

van Santen, Jan P. H. 1992. Contextual effects on vowel duration. *Speech Communication* 11. 513–546.

van der Stelt, Jeannette M. – Zajdó, Krisztina – Wempe, Ton G. 2005. Exploring the acoustic vowel space in two-year-old children: Results for Dutch and Hungarian. *Speech Communication* 47/1–2. 143–159.

Szende Tamás 1976. *A beszédfolyamat alaptényezői.* Akadémiai Kiadó, Budapest.

Torre, Peter – Barlow, Jessica A. 2009. Age-related changes in acoustic characteristics of adult speech. *Journal of Communication Disorders* 42. 324–333.

Watson, Peter J. – Munson, Benjamin 2007. A comparison of vowel acoustics between older and younger adults. In *Proceedings of the 16th International Congress of Phonetic Sciences.* 561–564.

Köszönöm Beke Andrásnak az adatok kinyerésében nyújtott segítségét.
A kutatás a Bolyai János kutatási ösztöndíj támogatásával készült.

BESZÉDHANGTÖRLÉS A MAGYAR SPONTÁN BESZÉDBEN

Váradi Viola

Bevezetés

A spontán beszéd Wacha (1974) definíciója szerint élőszó, amely a gondolataink megfogalmazásával párhuzamosan, azzal egyidőben zajlik le. A mondandónk megtervezése és elhangzása (meghangosítása) szinkrón tevékenység. A spontán beszédet a nem spontán beszédfajtáktól (pl. felolvasás) a beszédprodukciós folyamatok eltérő működése különbözteti meg (Markó 2005).

A spontán megnyilatkozások különböző szempontú vizsgálata napjainkban a fonetikai kutatások középpontjába került. Gósy (1997) vizsgálatai szerint a magánhangzók jelentős része neutralizálódik, vagyis [ə]-ként realizálódik a spontán beszédben. Az alveoláris spiránsok spontán beszédbeli megvalósulásának akusztikai fonetikai vizsgálatát végezte el Gráczi (2008). Gráczi–Horváth (2010) elemezte a magyar magánhangzók első két formánsának realizációját spontán beszédben. Gósy (2000) jellegzetes különbségeket mutatott ki két zöngétlen explozíva, a [p] és a [k] zöngekezdési idejében izolált szavak kiejtése és spontán beszéd során. Bóna (2011) az idősek és a fiatalok [p, t, k] mássalhangzóinak zöngekezdési idejét összehasonlítva azt találta, hogy a bilabiális és az alveoláris zárhangok az idősek beszédében szignifikánsan hosszabbak, mint a fiataloknál, ugyanakkor a veláris mássalhangzónál ellentétesen alakultak a zöngekezdési idők. A zöngétlen alveoláris és posztalveoláris résmássalhangzók szerkezetét elemezte Beke–Gyarmathy (2010) a spontán beszédben. A zöngésségi hasonulás érvényesülését is vizsgálták spontán beszédben (Gósy 2002), illetve különböző beszédfajtákban (Bóna et. al. 2008; Markó et. al. 2009). A magánhangzók nazalizációjának megvalósulását vizsgálta Beke–Horváth (2009) spontán beszédben.

A köznapi, spontán megnyilatkozásokat nagymértékű variancia jellemzi (Ács–Siptár 1994). A beszélő artikulációs sajátosságait spontán beszéd során számos tényező befolyásolja, még az azonos nyelvközösséghez tartozók között is nagyok a kiejtési különbözőségek (Gósy 2004). Ez az alapja Lindblom (1990) H&H (Hyper and hypo speech – 'túlartikulált és alulartikulált beszéd') elméletének, amely a beszédprodukció és –percepció összefüggéseit modellezi. Az elmélet alapja, hogy a beszédhangokat és a hangkapcsolatokat sohasem ejtjük egyformán, mert a kiejtést több tényező befolyásolja (szövegkörnyezet, hangerő, beszédtempó, beszédstílus stb.). A spontán beszéd során létrejövő megnyilatkozások a generatív fonológia kötelező szabályaival ke-

vésbé, funkcionális elvekkel sokkal inkább magyarázhatóak. A beszélő a legkisebb artikulációs energiával szeretné létrehozni közlését, a hallgató érdeke azonban a lehető legkisebb erőfeszítéssel történő megértése az elhangzottaknak. Az előbbi az artikulációs kényelmesség, az utóbbi a perceptuális szeparáció elvével magyarázható. Az artikuláció kényelmességével összefüggő folyamatokat a fonológiai szakirodalom lenizációs, simító vagy lazítási folyamatoknak nevezi (Ács–Siptár 1994).

Szende (1997) a lazítási folyamatot a következőképpen definiálja: a „'lazítás' (lenition) a realizáció egy poláris ellenpont felé tartó tendenciája, amely a nyelvi jelek elkülönítésének vesztesége árán szűkíti az artikulációs mozgások egy elemre vetített kitéréseit" (167). A lazítás érvényesülésének különböző előmozdító és gátló feltételeit sorolja fel: fonotaktikai pozíció, az alapalakmátrix összetevői, közlés műfaja, pragmatikai szerep, szemantikai érték, szintaktikai szerkezeten belüli helyzet, beszédtempó stb. A lenizációs folyamatokat többféleképpen osztályozzák, más-más típusait különítik el. Ács–Siptár (1994) négy, míg Szende (1997) hét fokozatát különbözteti meg a lenizációnak. Ács–Siptár (1994) a hasonulási, a rövidülési és a centralizációs folyamatokat, illetve a törlést, míg Szende a redukciót, a törlést, a kihagyást és csonkítást, a szekevenciaredukciót, a szekvenciazsugorítást, illetve az egybeolvadást, fúziót sorolja fel. A kétféle csoportosításban mást jelent a törlés fogalma. Szende a következőképpen definiálja: egy szegmentum fő és másodlagos képzési összetevőinek kiiktatása fonetikai nyomok hátrahagyásával a szekvencia más helyein. A beszédhang teljes eltűnését kihagyásnak, csonkításnak nevezi. Ács–Siptár (1994) a törlést a lenizáció legradikálisabb formájának tekinti, amely a teljes szegmentum eltűnését jelenti. Az artikuláció kényelmességével függ össze, és csökkenti a perceptuális szeparáció lehetőségét azáltal, hogy nehezíti a hallgató észlelési, megértési mechanizmusainak működését. A jelen kutatásunkban az utóbbi, funkcionális szemlélettel közelítettük meg a beszédhangok törlését.

Számos nyelvben a spontán beszéd fontos jellemzői a redukált, vagyis nem kodifikált ejtési változatok, ezt a tényt kvantitatív vizsgálatok is megerősítik. Johnson (2004) 40 amerikai angol anyanyelvű interjújából 88 000 szót elemzett. A szavaknak először a teljes alakját jegyezték le automatikusan, majd a lejegyzést képzett fonetikusok korrigálták manuálisan. Johnson azt találta, hogy a tartalmas szavak 6%-a, a funkciószavak 4,5%-a legalább egy szótaggal rövidebben realizálódott, mint a teljes alakjuk. Továbbá, a szavak 5–12%-ából legalább egy szegmentum hiányzott a teljes alakjukhoz képest. Ezek az eredmények összhangban vannak Dalby (1984) – amerikai angol televíziós interjúkat jegyzett le manuálisan – szótagtörlődésre vonatkozó vizsgálati eredményeivel, illetve Greenberg (1999), valamint Shattuck-Hufnagel és Veilleux (2007) korpuszokból származó adataival. Schuppler és munkatársai (2011) hasonló eredményeket kaptak dán nyelvű korpusz alapján. Ők automatikusan jegyeztek le egy 153 200 szóból álló korpuszt, amelyet 20 dán férfi beszélőtől

rögzítettek, és a szótagok törlése 19%-ban, míg a szegmentum váltakozása/törlése 40%-ban fordult elő. Bürki és munkatársai (2011) a franciában elemezték a svátörlést, amely jól ismert folyamat az angolban bizonyos fonológiai környezetben. Az extrém redukció jelensége, amelyet Niebuhr és Kohler (2011) vizsgált percepciós szempontból a németben, ismert más nyelvek spontán beszédében is, főképp a nagyon gyakori funkciószavak esetében. Más vizsgálatok azt mutatják, hogy még a tipológiailag rokonságban álló nyelvek is lényegesen különbözhetnek a redukció jellemzőiben. A redukció egyes típusait kategoriálisnak tekintik (például a szó belseji svá törlése a franciában), de a folyamatok többségére a fokozatosság jellemző. Mivel a beszélők létrehozzák, a hallgatók pedig feldolgozzák a redukált ejtési változatokat, ezért a beszédprodukció és a -percepció mechanizmusát bemutató pszicholingvisztikai modelleknek számolniuk kell ezekkel. A létező modelleket főképp gondozott, izolált ejtésű szavak vagy egyszerű mondatok laboratóriumi körülmények között rögzített vizsgálatának tanulságai alapján fejlesztették ki. Nem magától értetődő, hogy ezek a modellek a redukált kiejtési változatok produkcióját és percepcióját miként magyarázzák (Ernestus–Warner 2011).

Úgy tűnik, hogy a redukció elsősorban a beszéd legkevésbé informatív részét befolyásolja (Jurafsky et al. 2001), például a kontextusból nagy valószínűséggel megjósolható funkciószavakat, idiómákat, morfológiai elemeket (különösen a végződéseket), dátumokat, diskurzusjelölőket. A redukció következtében létrejöhetnek különböző (centralizált) fonémák és kevesebb szótagból felépülő realizációk is (Adda-Decker–Snoeren 2011).

Gyarmathy (2008) vizsgálta a magyar spontán beszéd fantomhangjait. A beszédfolyamat azon szegmentumait nevezte így, amelyek percepciósan igen, de vizuálisan nem észlelhetők, vagyis az oszcillogramon és a spektrogramon nem jelennek meg tipikus formájukban. Elemzései során arra a következtetésre jutott, hogy nem minden beszédhang realizálódik fantomhangként, és ez összefügg a mássalhangzók spontán beszédbeli előfordulásának gyakoriságával. A jelenség létrejöttét más tényezők is elősegítik: a beszélők artikulációs biztonsága, az intervokális hangkörnyezet – főképp a középső és az alsó nyelvállású magánhangzók –, a képzéshelyek közötti távolság. A szótövekben statisztikailag igazolhatóan magasabb arányú volt a fantomhangok realizációja, mint a toldalékokban.

A jelen kutatás célja, hogy spontán beszédanyagon elemezzük a beszédhangok törlésének sajátosságait. Számos tényező befolyásolhatja egy beszédhang „eltűnését", a jelen vizsgálatban ezek közül a beszédhang artikulációs sajátosságait (magánhangzó, ill. mássalhangzó, képzésmód) és a beszédszakaszok hosszát, illetve artikulációs tempóját elemeztük. Feltételezésünk szerint a mássalhangzók gyakrabban törlődnek, mint a magánhangzók (vö. Gyarmathy 2008), nem minden mássalhangzó hajlamos egyforma mértékben a törlődésre, és a törlődés mértékét a képzés módja is befolyásolja. Feltételez-

tük továbbá, hogy a beszédszakasz időtartama és artikulációs tempója összefügg a beszédhangok törlésének arányával.

Anyag, módszer, kísérleti személyek

A jelen kutatásunkban a magyar spontánbeszéd-adatbázis (BEA) 10 női adatközlőjének hangelvételeit elemeztük (Gósy 2008). Az adatközlők átlagéletkora 26 év volt (22–32 év). Az összesen 27 percnyi korpusz mintegy 3500 szóból és 16 000 beszédhangból állt. Az elemzéshez a Praat 5.1 szoftvert használtuk (Boersma–Weenink 2010). A hanganyagot manuálisan beszédszakaszszinten – két szünet közötti szöveges szakasz –, majd kényszerített automatikus címkézést követően kézi javítással (a rezgés- és hangszínkép alapján folyamatos auditív ellenőrzéssel) beszédhangszinten annotáltuk. Elemeztük a törlődött beszédhang artikulációs sajátosságait, egy erre írt Praat szkript segítségével automatikusan kiszámoltuk a beszédszakaszok hosszát és artikulációs tempóját. A törlődött beszédhangok arányát minden esetben az előforduló fonémák számához viszonyítottuk.

A stasztikai elemzésekhez az SPSS 19.0 verzióját használtuk (Pearson-féle korrelációelemzés).

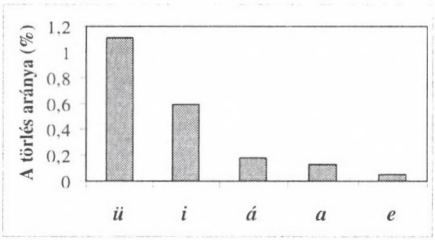
Eredmények

A korpusz 16 606 fonémájából 464 darab (2,79%) törlődött, a szavak 12%-át érintette a redukció legradikálisabb típusa. Angol anyanyelvűekkel végzett vizsgálat során hasonló eredményre jutottak, a beszédhang törlődése a szavak 5–12%-át érintette (Johnson 2004). A jelen korpuszban a törlés jelensége főképp a mássalhangzókat érintette (4,73%), összesen csak 10 magánhangzó (0,14%) törlődött.

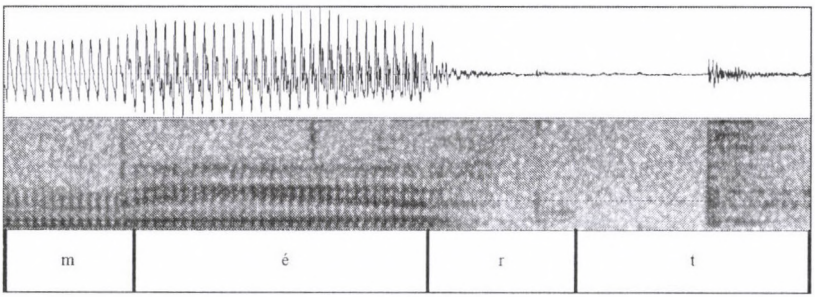
Öt magánhangzót [y, i, a:, ɔ, ɛ] érintett a törlés (1. ábra). A két felső nyelvállású palatális magánhangzó törlődött a legnagyobb arányban. A labiális felső nyelvállású palatális egy ([y] 1,11%-a), az illabiális öt esetben ([i] 0,59%-a) (vö. 2. ábra). Középső nyelvállású palatális és veláris magánhangzó [ø, e:] egy esetben sem törlődött, az alsó nyelvállásúk nagyon kis arányban törlődtek, a palatális egy ([ɔ] 0,05%-a), a veláris két esetben ([ɛ] 0,13%-a). A magánhangzótörlések mindig a szótagok szerkezetének megváltozásával, esetleg a szótagszám változásával járnak együtt, vagyis olyan folyamatok, amelyek magasabb nyelvi szinteken érvényesülnek (Szende 2011).

Adda-Decker és munkatársai (2005) francia nyelvű rádióinterjúkban elemezték a szótagszerkezeteket, és megfigyelték, hogy a korpuszban előforduló összes magánhangzó 15%-át érintette a törlés. Ha kizárták az elemzésükből a svát, akkor 6%-ra csökkent ez az arány. A következő magánhangzók törlését adatolták nagy arányban: [ɔ] 10%, [u] 8%, [ɛ̃] 7%, [ɛ] 7%, [i] 5%, [a] 5%. A két vizsgálat eredményei közötti eltérések feltehetően nyelvspecifikus sajátosságokkal és beszédstílusbeli különbségekkel (monologikus vs. dialogikus) magyarázhatók. A jelen vizsgálat anyagát magyar nyelvű monologikus spon-

tán beszéd, míg Adda-Decker és munkatársainak vizsgálatában francia nyelvű dialógusok képezték.



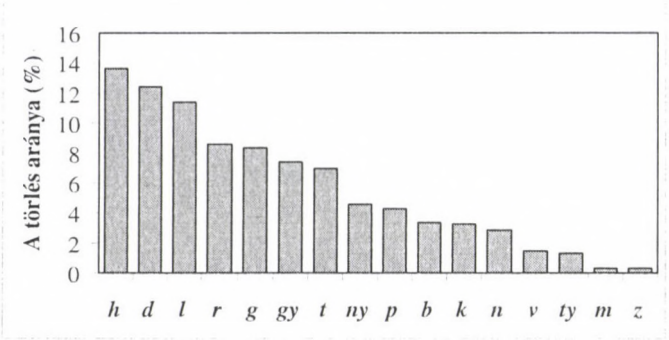
1. ábra
A magánhangzók törlésének aránya



2. ábra
A *miért* szó rezgés- és hangszínképe a törlődött [i] nélkül

Gyarmathy (2008) vizsgálatában egyetlen magánhangzó fantomizálódását sem adatolta. A két vizsgálat között megfigyelt különbségek érdekes kérdéseket vetnek fel a humán észlelési mechanizmus működéseinek sajátosságaival kapcsolatban (bár a két vizsgálat más-más módszerrel és korpuszon történt, ezért csak óvatos összehasonlításra van lehetőségünk). Feltételezésünk szerint a magánhangzó törlése nem kedvez a fantomizálódásnak, vagyis nem marad észrevétlen a hiánya a hallási percepció számára.

Szende (2011) megfigyelésen alapuló fonológiai szempontú tanulmányában megállapítja, hogy „a fonématörlés leggyakoribb áldozatai a mássalhangzók" (110). Hipotézisünknek megfelelően nem minden mássalhangzó hajlamos egyforma mértékben a teljes redukcióra. Korpuszunkban 16 mássalhangzóra volt jellemző a törlés, a 3. ábra szemlélteti, hogy milyen arányban törlődtek ezek a mássalhangzók a korpuszban adatolt összes előfordulásukhoz viszonyítva.



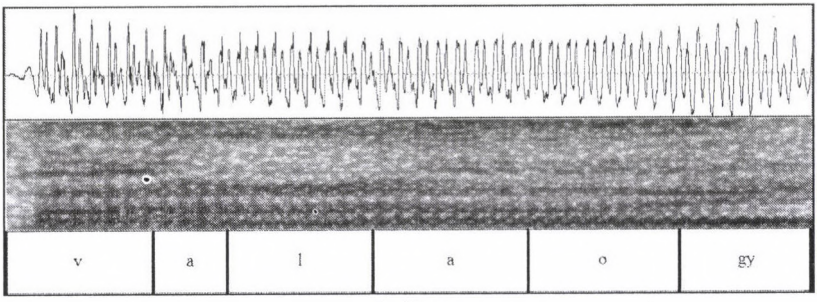
3. ábra
A mássalhangzók törlésének aránya

Az affrikáták, három zöngétlen ([f], [s], [ʃ]) és a [ʒ] zöngés spiráns, valamint a [j] approximáns egyetlen esetben sem törlődtek. Nagy arányban érintette a törlés a következő mássalhangzókat: [h] 13,67% (vö. 4. ábra), [d] 12,42%, [l] 11,39%. Gyarmathy (2008) korábban említett vizsgálata szerint a leggyakoribb fantomhangok a [h], a [l] és a [r]. A tremuláns a jelen korpuszban is nagy arányban törlődött (8,63%), de Gyarmathynál a zöngés alveoláris explozíva kisebb arányban fantomizálódott, mint a mi korpuszunkban törlődött. Gyarmathy azokat a törlődött beszédhangokat elemezte, amelyek akusztikusan észlelhetők voltak. Feltételezésünk szerint az eltérő sorrend hátterében az állhat, hogy egy zöngés explozíva hiánya feltűnőbb az észlelésben, mint más beszédhangoké. Ács és Siptár (1994) is megállapítja, hogy a [h] az egyik leggyakrabban törölt mássalhangzó, utána következik a [l], [r] és a [j]. Szende (2011) is megfigyelte, hogy a [l, r, j] mássalhangzókra jellemző a törlés, a nazálisokat azonban kevésbé érinti ez a változás. A különböző beszélők tendenciaszerűen azonos beszédhangokat töröltek, de nem várható el egy uniformizált stratégia.

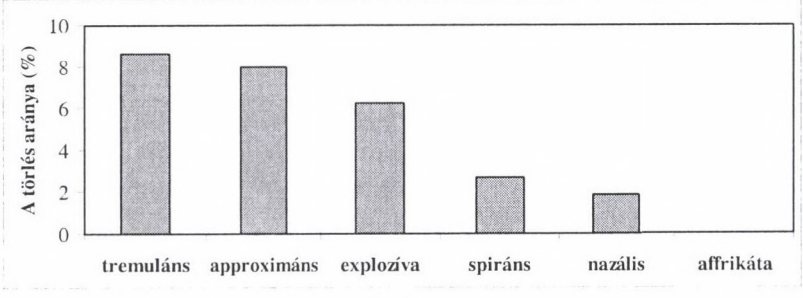
Elemeztük, hogy mely képzésmódú mássalhangzókra legjellemzőbb a törlés az adott képzésmódú mássalhangzók összes előfordulásához képest (5. ábra). A tremulánsok (8,63%), az approximánsok (7,97%) és az explozívák (6,25%) a legnagyobb, míg a spiránsok (2,70%) és a nazálisok (1,84%) kisebb arányban törlődtek.

Ács–Siptár (1994) fonológiai szempontú összefoglalása szerint leginkább a zöngétlen zörejhangok állnak ellen a törlésnek, még azonos magánhangzók között sem törlődnek. A spiránsok fesztelen beszédben még akkor sem törlődnek magánhangzó-környezetben, ha zöngések. A jelen, korpuszalapú vizsgálatunkban azonban legkevésbé az affrikátákra volt jellemző a törlődés, ilyet egyetlen esetben sem adatoltunk. Az affrikáták akusztikai szerkezetüket tekintve összetett hangok, vagyis nem állandóak a paraméterek az akusztikai

szerkezetükben, több képzési mozzanatból tevődnek össze (Gósy 2004). Feltételezésünk szerint az ilyen képzésmódú mássalhangzókra a redukció kevésbé radikális formái (az időtartam rövidülése; spektrális redukció, vagyis a spektrális minőség megváltozása) a jellemzőbbek a törlődés helyett.



4. ábra
A *valahogy* szó rezgés- és hangszínképe a törlődött [h] nélkül

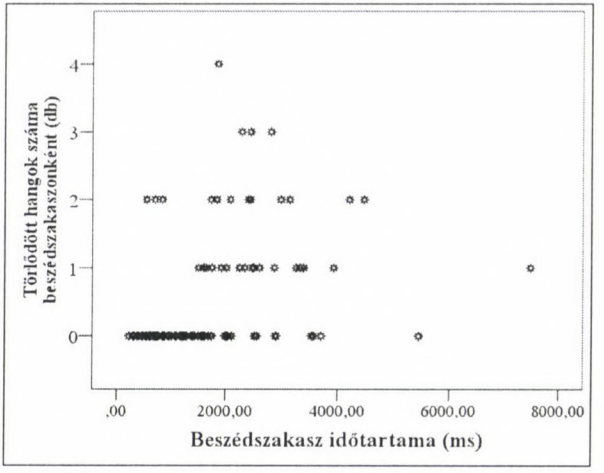


5. ábra
A mássalhangzók törlésének aránya a képzésmód szerint

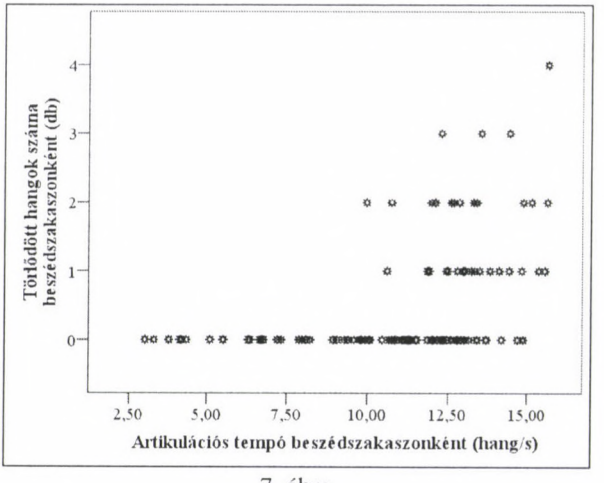
Feltételeztük, hogy a beszédszakaszok, vagyis a szünettől szünetig terjedő szöveges szakaszok hossza összefüggést mutat a beszédhangtörlések számával (6. ábra). Hipotézisünk részben igazolódott, a Pearson-féle korrelációelemzés szerint szignifikáns, de gyenge az összefüggés a két tényező között ($r = 0{,}366$; $p < 0{,}001$). Megfigyelhető az a tendencia, hogy a hosszabb időtartamban realizálódó beszédszakaszokból nagyobb arányban törlődnek beszédhangok.

A generatív fonológia gyorsbeszéd-folyamatoknak nevezi azokat a folyamatokat, amelyek nem illenek bele az invarianciaképbe, vagy amelyeket nem tud kezelni megfelelő szabályapparátus híján (Ács–Siptár 1994). Ezek közé a folyamatok közé sorolható a törlés jelensége is. Felmerül a kérdés, hogy való-

ban a sebesség-e a legmeghatározóbb tényező. A kérdés megválaszolása érdekében elemeztük a beszédszakaszonként mért artikulációs tempó és az adott beszédszakaszban adatolt törlések számának összefüggését (7. ábra).



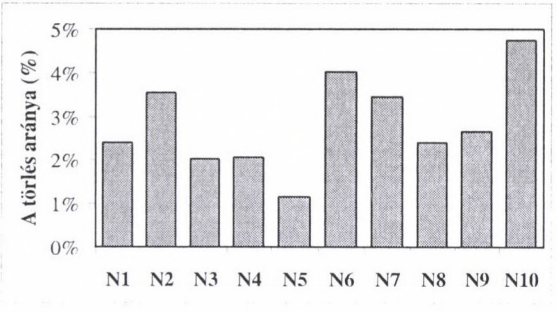
6. ábra
A törlés arányának és a beszédszakasz hosszának összefüggései



7. ábra
A törlés és a beszédszakaszok artikulációs tempójának összefüggései

A Pearson-féle korrelációelemzés szerint szignifikáns és közepesen erős az összefüggés ($r = 0{,}436$; $p < 0{,}001$). A törlések jellemzően a magasabb artikulációs tempójú beszédszakaszokban realizálódtak, de nagy az aránya azoknak a nagyobb artikulációs tempójú beszédszakaszoknak is, amelyekben nem ment végbe törlés. Eredményeink azt mutatják, hogy a törlés aránya nagyobb mértékben függ a beszédszakasz artikulációs tempójától, mint a hosszától.

A beszélők egyéni beszédjellemzői – ejtési tudatosság, artikulációs biztonság, beszédtapasztalat – szintén befolyásolhatják a beszédhangok törlésének arányát (8. ábra).



8. ábra
A beszédhangok törlésének aránya adatközlőnként

Az adatközlők között jelentős különbségeket adatoltunk a beszédhangtörlés mértékének tekintetében. Az N5 beszélőre jellemző a legkevésbé a beszédhangok törlése (1,16%), míg az N10 beszélőre sokkal nagyobb arányú törlés jellemző (4,75%).

Következtetések

A tanulmány célja a beszédhangok törlésének elemzése volt magyar spontánbeszéd-korpuszon. Számtalan tényező befolyásolhatja a beszédhangok törlődését, a jelen vizsgálatban ezek közül a képzési sajátosságokkal, a beszédszakasz hosszával és artikulációs tempójával foglalkoztunk.

A spontán beszéd variabilitásának, produkciós sajátosságainak köszönhetően a beszédhangtörlés a magyarban sem elhanyagolható jelenség, a szavak 12%-át érintette a korpuszban. Hipotéziseink igazolódtak: a mássalhangzók gyakrabban törlődtek, mint a magánhangzók; nem minden mássalhangzó törlődött egyforma arányban; és a mássalhangzók képzésmódja szerint eltérő mértékű volt a törlés. A beszédszakasz időtartama és artikulációs tempója is összefüggést mutatott a beszédhangok törlésének arányával, de a korreláció mértéke magasabb volt az artikulációs tempó esetén, mint az időtartamnál. Feltételezésünk szerint más tényezők, például a szó hossza is befolyásolhatja

a törlés arányát. A beszélők egyéni beszédjellemzői szintén befolyásolhatják a beszédhangok törlésének arányát, az adatközlők között jelentős eltéréseket adatoltunk a törlés mértékét illetően. Feltételezésünk szerint a beszédben való jártasság, a beszédtémához való érzelmi viszonyulás, az artikulációs tudatosság és biztonság, a beszédhelyzethez való alkalmazkodás számos egyéb tényező mellett befolyásolhatják, hogy hány és melyik beszédhang törlődik a szekvenciából. Ennek vizsgálatához elengedhetetlen a több beszélő bevonásával, többféle szituációban rögzített nagyméretű korpuszon történő elemzés.

A redukált kiejtési változatokkal, így a beszédhangtörlésekkel foglalkozó kutatásoknak köszönhetően hipotéziseket állíthatunk fel arra nézve, hogy az egyes szegmentumok artikulációs jellemzőinek mekkora szerepe van az adott szegmentum azonosításában. Ha a megfigyelt tendenciákat rávetítjük a nyelvtörténet idősíkjára, akkor betekintést nyerhetünk a hangtörténeti változások folyamatainak fonetikai hátterébe is (Szende 2011).

A spontán beszéd variabilitása, a nem kanonikusan realizálódott alakok és a beszédhangtörlés kezelése az automatikus beszédfelismerés egyik alapvető problémája. Korpuszokon végzett vizsgálatok eredményei azt mutatják, hogy az automatikus beszédfelismerő rendszerek jóval kisebb hibaszázalékkal működnek olvasott beszéd (10% körül), mint spontán telefonbeszélgetések (30–50% körül) esetén (vö. Nakamura et al. 2008 és Novotney et. al. 2010). A humán észlelési mechanizmus működése sem tökéletes a redukált ejtési változatok feldolgozása során. A hallgató nem feltétlenül tudja eldönteni egy fonémáról, hogy törlődött-e, ha a hiányzó fonéma akusztikai jellemzői jelen vannak a környező beszédhangokban.

Számos kutatás foglalkozott az agy hanghelyreállító hatásával (vö. Warren 1970, Dankovics–Pléh 2011, Kashino 2006). A kísérletekben köhögéssel vagy más zajjal helyettesítették egy szegmentumsor egy vagy több fonémáját, amely többszöri meghallgatás során is észlelhető maradt. Régóta ismeretes a beszédpercepció vizsgálatában, hogy az előre jelzés, illetve a szó fonológiai formájának aktiválása a percepcióban biztosítja a hallgató számára a sikeres dekódolást. Sokszor rejtve marad a törlés ténye, a hallgató tudatosan nem észleli a hiányt, mert a magasabb rendű megértési folyamatok működésének következtében sikeres a szófelismerés.

Irodalom

Ács Péter – Siptár Péter 1994. Túl a gondozott beszéden. In Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 2. Fonológia.* Akadémiai Kiadó, Budapest, 550–580.

Adda-Decker, Martine – Boula de Mareüil, Philippe – Adda, Gilles – Lamel, Lori 2005. Investigating syllabic structures and their variation in spontaneous French. *Speech Communication* 46. 119–139.

Adda-Decker, Martine – Snoeren, Natalie D. 2011. Quantifying temporal speech reduction in French using forced speech alignment. *Journal of Phonetics* 39. 261–270.

Beke András – Gyarmathy Dorottya 2010. Zöngétlen résmássalhangzók akusztikai szerkezete. *Beszédkutatás 2010.* 57–75.

Beke András – Horváth Viktória 2009. A nazális koartikuláció variabilitása a spontán beszédben. *Beszédkutatás 2009.* 28–45.

Boersma, Paul – Weenink, David 2010. Praat: doing phonetics by computer (Version 5.1.13). http://www.fon.hum.uva.nl/praat/dowload_win.html.

Bóna, Judit – Gráczi, Tekla Etelka – Markó, Alexandra 2008. Coarticulation rules and speaking style dependency. In Sock, Rudolph – Fuchs, Susanne – Laprie, Yves (eds.): International Seminar on Speech Production 2008. Strasbourg, 245–248.

Bürki, Audrey – Fougeron, Cécile – Gendrot, Cedric – Fraunfelder, Ulrich Hans 2011. Phonetic reduction versus phonological deletion of French schwa: Some methodological issues. *Journal of Phonetics 39.* 279–288.

Dalby, Jonathan Marler 1984. *Phonetic structure of fast speech in American English.* Dissertation. Indiana University Linguistic Club, Bloomington.

Dankovics Natália – Pléh Csaba 2001. Hangrestaurációs jelenségek és alaktani feldolgozás a magyarban: azt halljuk-e, amit várunk? In Pléh Csaba – Lukács Ágnes (szerk.): A magyar morfológia pszicholingvisztikája. Osiris Kiadó, Budapest, 55–83.

Ernestus, Mirjam – Warner, Natasha 2011. An introduction to reduced pronunciation variants. *Journal of Phonetics* 39. 253–260.

Gósy Mária 1997. Semleges magánhangzók a magyar beszédben. *Magyar Nyelvőr* 121. 9–19.

Gósy Mária 2000. A [p, t, k] mássalhangzók zöngekezdési ideje. *Magyar Nyelvőr* 124. 195–204.

Gósy Mária 2002. Temporal coding of voicing assimilation in speech production. *Acta Linguistica Hungarica* 49. 257–276.

Gósy Mária 2004. *Fonetika, a beszéd tudománya.* Osiris Kiadó, Budapest.

Gósy Mária 2008. Magyar spontánbeszéd-adatbázis – BEA. *Beszédkutatás 2008.* 194–207.

Gráczi Tekla Etelka 2008. Alveoláris spiránsok akusztikai fonetikai vizsgálata. *Beszédkutatás 2008.* 33–51.

Gráczi Tekla Etelka – Horváth Viktória 2010. A magánhangzók realizációja spontán beszédben. *Beszédkutatás 2010.* 5–16.

Greenberg, Steven 1999. Speaking in shorthand – A syllable-centric perspective for understanding pronunciation variation. *Speech Communication* 29. 159–176.

Gyarmathy Dorottya 2008. Fantomhangok a magyar spontán beszédben. *Beszédkutatás 2008.* 67–76.

Johnson, Keith 2004. Massive reduction in conversational American English. In Yoneyama, Kyoko – Maekawa, Kikuo (eds.): *Spontaneous speech: Data and analysis.* The National International Institute for Japanese Language, Tokyo, 29–54.

Jurafsky, Daniel – Bell, Alan – Gregory, Michelle – Raymond, William D. 2001. Probabilistic relations between words: Evidence from reduction in lexical production. In Bybee, Joan – Hopper, Paul (eds.): *Frequency and the emergence of linguistic structure.* John Benjamins, Amsterdam, 229–254.

Kashino, Makio 2006. Phonemic restoration: The brain creates missing speech sounds. *Acoustical Science and Technology* 27/6. 318–321.

Lindblom, Björn 1990. Explaining phonetic variation: A sketch of the H&H theory. In Hardcastle, William J. – Marchal, Alain (eds.): *Speech production and speech modelling*. Kluwer, Dordrecht, 403–439.

Markó Alexandra – Gráczi Tekla Etelka – Bóna Judit 2009. Zöngésségi hasonulás a spontán beszédben és a felolvasásban. (Esettanulmányok.) *Beszédkutatás 2009*. 5–27.

Nakamura, Masanobu – Iwano, Koji – Furui, Sadaoki 2008. Differences between acoustic characteristics of spontaneous speech and their effects on speech recognition performance. *Computer Speech and Language* 22. 171–184.

Niebuhr, Oliver – Kohler, Klaus J. 2011. Perception of phonetic details in the identification of highly reduced words. *Journal of Phonetics 39*. 319–329.

Novotney, Scott – Callison-Burch, Chris 2010. Cheap, fast and good enough: Automatic speech recognition with non-expert transcription. In: *Human Language Technologies: The 2010 annual conference of the North American chapter of the ACL*. Los Angeles, 207–215.

Schuppler, Barbara – Ernestus, Mirjam – Scharenborg, Odette – Boves, Louis 2011. Acoustic reduction in conversational Dutch: A quantitative analysis based on automatically generated segmental segmental transcriptions. *Journal of Phonetics* 39. 96–109.

Shattuck-Hufnagel, Stefanie – Veilleux, Nanette 2007. Robustness of acoustic landmarks in sponatneously-spoken American English. In *Proceedings of the 16th international congress of phonetic sciences ISPhS-07*. Saarbrücken. 925–928.

Szende Tamás 1997. *Alapalak és lazítási folyamatok*. MTA Nyelvtudományi Intézet, Budapest.

Szende Tamás 2011. Változási tendenciák a mai magyar beszédben. *Magyar Nyelvőr* 135. 107–112.

Wacha Imre 1974. Az elhangzó beszéd főbb akusztikus stíluskategóriáiról. *Általános Nyelvészeti Tanulmányok* X. 203–216.

Warren, Richard M. 1970. Perceptual restoration of missing speech sounds. *Science* 167. 392–393.

A FORMÁNSMENETEK RENDSZERE CVC KAPCSOLATOK MAGÁNHANGZÓIBAN A C KÉPZÉSI HELYÉNEK FÜGGVÉNYÉBEN

Abari Kálmán – Olaszy Gábor

Bevezetés

A magyar beszédhangok formánsainak vizsgálata a több mint 70 éves múltra tekint vissza (Tarnóczy 1941), és ennek megfelelően rendkívül gazdag szakirodalommal rendelkezik. Az akusztikai fonetika eme kiemelt témakörének alapfogalmai és a kezdeti kutatásoktól a napjainkig terjedő bemutatása Gráczi és Horváth (2010) kitűnő tanulmányának bevezetőjében olvasható. A magánhangzók formánsszerkezetét illetően ismert, hogy igényes ejtésben, izolált szóban az első és második formáns már egyértelműen meghatározza a magánhangzó minőségét (Gósy 2004). Mindeddig a mérések célzott hangfelvételeken alapultak, és sok esetben csak a magánhangzó közepéről vett formánsértéket vizsgálták átfogóan. Vértes O. András az *agyagya* logatom második [ɔ] magánhangzóját elemezte férfi beszélő ejtésében, és a tiszta fázisra jellemző 1000 Hz-es érték előtt és után 1200 Hz-es és a fölötti részt talált a második formánsban. Ez alapján írja: „Az említett jelenség nem kivétel. Éppen ezért kérdéses, hogy mennyire jogos a magánhangzó közepére korlátozni a méréseket" (Vértes 1982: 78). A magyar magánhangzók formánsszerkezetére vonatkozó első átfogó elemzést Magdics (1965) adta közre. Vizsgálatában 26 mássalhangzó és 8 magánhangzó összes CV és VC hangkapcsolódásának frekvenciaszerkezetét írja le. A CV hangpárok F_2 és F_3 formánsmozgásait ábrán követhetjük nyomon, de mindkét hangkapcsolat esetén (CV és VC) a formánsmozgás induló és cél frekvenciaértékeire számszerű adatokat is találunk. Attól függően, hogy a szomszédos beszédhangok artikulációja mennyire tér el egymástól, a formánsfrekvencia-értékek az időben kisebb-nagyobb mértékben változnak a magánhangzó képzési idején belül, azaz a formáns mozog a magánhangzón belül. Bolla (1978) szintén teljeskörű elemzést adott a magyar magánhangzók tiszta fázisát jellemző formánsokat illetően, vizsgálataihoz beszédszintézist is használt. Olaszy (1985) a CV és VC kapcsolatokat vizsgálja analízis-formánsszintézis módszerrel, melyhez szóadatbázist készít (407 elem egyetlen férfi bemondóval). Törtvonalas közelítéssel megadja az F_1, F_2, F_3 mozgását a CV és a VC diád magánhangzójában 23 C-re és 9 V-re. Minden hangkapcsolódásra találunk adatot, tehát a teljes magyar hangkapcsolódási teret lefedi. A formánsmozgásokat számadatokkal írja le, amelyek korrektségét szintetizált szavakon igazolja. A hangátmeneti részben bekövetkező for-

mánsmozgás percepcióban betöltött szerepét Gósy (1995) vizsgálta CV hangkapcsolatokban. A hangátmenettől megfosztott hangkapcsolat mássalhangzójának felismerhetőségével kapcsolatos kísérlete alapján megállapítja, hogy a hangátmenetek elsődleges vagy másodlagos akusztikus kulcsok az észlelésben, de sok esetben redundáns tényezőként funkcionálnak. Szerepük függ a mássalhangzók minőségétől és a kontextustól.

A formánsok mozgását leíró matematikai modellekre csak a nemzetközi szakirodalomban találtunk példákat. A magánhangzó két átmeneti részét egy-egy másodfokú görbével írja le az egyik legkorábbi tanulmány e témában (Stevens et al. 1966), de van kísérlet a mozgási formák megragadására egyetlen parabolával (van Bergem 1993), exponenciális függvénnyel (Lindblom 1963) vagy diszkrét koszinusz-transzformációval (Watson–Harrington 1999). A mássalhangzó-magánhangzó találkozások formánsváltozásainak könnyen implementálható megközelítése a locus egyenlet (Krull 1987; vö. Csapó–Németh 2009), melyet főképp felpattanó zárhangok és magánhangzók átmeneteinek vizsgálatához alkalmaznak. A locus egyenlet a koartikulációs hatásokra két mérési pont alapján következtet. A módszer egyfajta általánosítását Broad és Clermont (2010) adja közre. Az általuk definiált lineáris modellben több mérési pont adata szerepelhet, például idézett tanulmányukban 11 pontot használnak a magánhangzó teljes hangidőtartamából a második formáns reprezentálására.

A jelen kutatás magyar nyelven elsőként vizsgálja a magánhangzók formánsmozgásait 5 mért pont alapján, formánsadatbázisból vett frekvenciaértékek vizsgálatával. A formánsmozgásokat saját modellel közelítjük, amely a magánhangzó teljes időtartamában jellemzi a formáns időbeli változását. A vizsgálati tér leszűkítésnek érdekében csak a CVC hangkapcsolatokat elemezzük, és ezeken belül is további hangösszevonásokat alkalmazunk a mássalhangzók képzési helyei (C_{kh}) és a magánhangzó minőségek (V_m) szerint. Ezzel olyan méretűre szűkítettük a mérési teret, ami kezelhetővé teszi a kérdéskört. Összefoglalva, a $C_{kh.1}V_mC_{kh.2}$ hármasok rendszerében a formánsmozgásokon keresztül vizsgáljuk a koartikulációs hatásokat, és kísérletet teszünk a formánsmozgások formáinak meghatározására és rendszerezésére.

Ismert, hogy a mássalhangzók képzési helyei egy-egy jellemző artikulációs pozíciót jelentenek a toldalékcső akusztikai átviteli tényezőinek formálása során. Vizsgálatunkban úgy tekintjük, hogy egy adott képzési helyet jellemző artikulációs konfigurációhoz hozzárendelhető a rá jellemző spektrális vetület, amelyet itt most csupán az F_1 és F_2 formánssal fejezünk ki (ez nem beszédhanghoz köthető formánsokat jelent). Ez a vetület a koartikuláció eredményeként hat a magánhangzón belül végbemenő formánsmozgásra a hangátmeneti szakaszokban és kisebb-nagyobb mértékben a magánhangzó tiszta fázisában is (Vértes 1989: 80). Úgy tekintjük, hogy a mássalhangzó képzési helyének (C_{kh}) ilyen fajta lenyomata a frekvencia tartományban nem függ lényegesen az ott képzett mássalhangzó képzési módjától és gerjesztésétől,

ezért általánosan alkalmasnak tartjuk mássalhangzó csoportok képviseletére (Olaszy 2003). Első hipotézisünk szerint az azonos $C_{kh.1}V_mC_{kh.2}$ cellákhoz azonos jellegű formánsmozgások tartoznak a magánhangzóban.

Célunk, hogy a formánmozgások alakjáról is képet kapjunk a magánhangzó teljes terjedelmében. Az általunk javasolt, matematikailag megfogalmazott alakok rendszerén keresztül tipikus formánsmozgások kategóriáit hozzuk létre. Második hipotézisünk szerint ez a kategorizálás megfelelően (kevés hibával) sorolja be az F_1 és az F_2 formánsmozgásokat az egyes kategóriákba, illetve alkalmas a $C_{kh.1}V_mC_{kh.2}$ cellák formánsmozgásainak jellemzésére.

A harmadik hipotézisünk az, hogy a $C_{kh.1}V_mC_{kh.2}$ magánhangzóiban lezajló formánsmenet típusok száma sokkal kisebb, mint a kapcsolatokból adódó hanghármas típusok száma. Ez azt jelenti, hogy lesznek olyan formánsmozgástípusok, amelyekkel többféle $C_{kh.1}V_mC_{kh.2}$-t lehet jellemezni. Így adattömörítésre nyílik lehetőség a formánsmozgások rendszerének leírásánál, esetleges algoritmusok kialakításánál.

Módszer

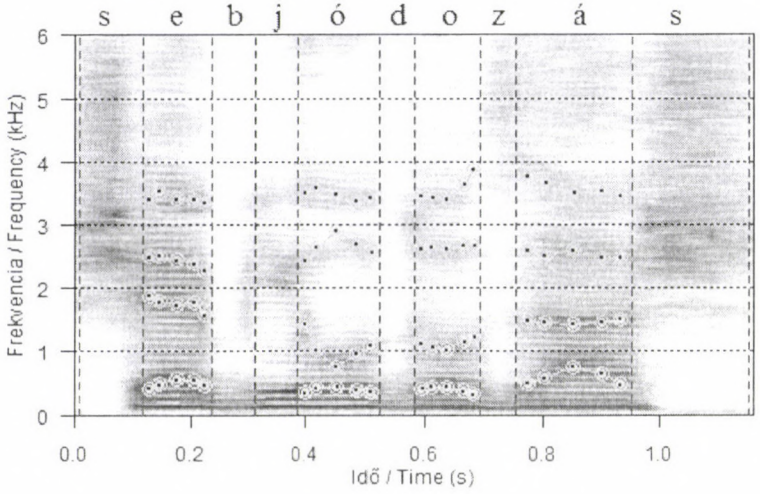
A vizsgálatokhoz az első, interneten elérhető magyar formánsadatbázis frekvenciaadataiból (Abari–Olaszy 2011) indultunk ki. Az adatbázis beszédkorpusza szólistát tartalmaz, egy férfi (60 éves) és egy nő (32) felolvasásában (1500 szó). Magánhangzónként 3 mérési pont (az időtartam 25%-ában, 50%-ában 75%-ában) jellemzi az egyes formánsok mozgását. Azért, hogy a hangátmeneti részeket pontosabban tudjuk leírni, minden magánhangzóban 5 mérési pontot (10%, 25%, 50%, 75%, 90%) határoztunk meg a formánsmozgás jellemzésére. Így a hangátmeneti részek formánsmozgásait pontosabban tudjuk modellezni. Ez képezte a mérési teret. Az adatbázis használata biztosítja, hogy a mérési eredmények bárki számára ellenőrizhetők (vizuálisan is), a mérések bármikor megismételhetők, más mérési eredményekkel összehasonlíthatók. Az 5 mérési pontos formánsadatbázis kiegészítő méréseinek adathelyességét vizuálisan ellenőriztük, tehát ez a formánsadatbázis is referenciának tekinthető, hasonlóan a 3 pontoshoz.

A vizsgálatokban a 9 magánhangzó-minőség szerepel (V_m): [ɔ], [aː], [ɛ], [eː], [i, iː], [o, oː], [ø, øː], [u, uː] és [y, yː]. A mássalhangzókat 6 képzési hely szerint csoportosítottuk (C_{kh}): BL = bilabiális, LD = labiodentális, DA = dentialveoláris, AL = alveoláris, PA = palatális, VE = veláris. Ennek megfelelően 6 × 9 × 6 = 324 féle $C_{kh.1}V_mC_{kh.2}$ hanghármas a vizsgálat tárgya. Ezeket a hanghármas típusokat röviden *cellák*nak nevezzük. A BL[ɔ]VE cella például a [bɔg], [pɔg], [mɔg], [bɔk], [pɔk] és [mɔk] hanghármasokat fedi le. A cellákban előforduló minták száma változó, ezt az adatbázis jelenlegi tartalma határozza meg.

A nyelvi anyagban 9778 db magánhangzó található, és kettő kivételével minden $C_{kh.1}V_mC_{kh.2}$ kapcsolatra van minta (a kivétel az LD[ø, øː]LD és a BL[u, uː]VE cella). Az adatbázisban normál köznapi ejtéssel kiejtett magyar

szavak szerepelnek; a bemondókat semleges hanglejtésre kérték. A formánsadatbázis készítésének további részleteiről lásd Olaszy et al. 2009.

Az 1. ábrán az adatbázis *sebjódozás* szóalakjában lévő 4 magánhangzó formánsmozgása (F_1, F_2, F_3 és F_4) látható. A fekete pontok az adatbázisban rögzített frekvenciaértékeket jelzik. A 4 magánhangzó rendre az AL[ɛ]BL, PA[o, o:]DA, DA[o, o:]DA és DA[a:]AL hanghármas kategóriák egy-egy adatbázisbeli előfordulása.



1. ábra
Példa három $C_{kh.1}V_mC_{kh.2}$ hangkapcsolat magánhangzójának az adatbázisban rögzített mérési pontjaira (a *sebjódozás* szóból)

A $C_{kh.1}V_mC_{kh.2}$ kapcsolatok magánhangzójának csak az F_1, F_2 értékeit vizsgáltuk az 5 mérési ponton. Ez közel 100 000 formánsfrekvencia-értéket jelent. A formánsmozgások modellezésére matematikai módszereket használunk. Ezek részleteiről az egyes vizsgálatoknál szólunk. Az adatok rendszerezésére, lekérdezésére, statisztikai és fonetikai elemzésére az R statisztikai programot (2011) használtuk. (E tanulmányban a formánsmozgás szóösszetételen a konkrét formánsmozgásokat értjük egy-egy formánsra vonatkoztatva. A modellünkben szereplő 27 db (5. ábra) mozgásformára a formánsmenet összetett szót használjuk.)

A vizsgálat véghezviteléhez lépcsőzetesen közelítettünk. Az első lépésben normalizáltuk a mérési tér adatait, hogy nemtől függetlenül vizsgálhassuk az adatokat. A második lépésben a formánsmozgásokat összesítettük, hasonló-

ságukat statisztikai mutatóval jellemeztük. A következő fázisban modelleztük a formánsmozgásokat, és a modellek alapján meghatároztuk minden $C_{kh.1}V_mC_{kh.2}$ hanghármasra a jellemző formánsmenettípust. Statisztikailag értékeltük, hogy a formánsmenetek milyen mértékben kapcsolhatók a mássalhangzó képzési helyéhez a magánhangzó függvényében. Az utolsó lépésben a formánsmeneteket egyesítettük, és létrehoztuk a jellemző formánstérképeket, amelyekben mind az F_1-re, mind az F_2-re jellemző formánsmeneteket – hasonlóan a hangspektrogramhoz – megadtuk az idő függvényében. A formánstérkép jellemzi a magánhangzót, valamint a megelőző és a követő mássalhangzót a képzési helyük függvényében.

Eredmények

Normalizálás

A formánsadatbázis adatainak együttes elemzéséhez a formánsértékeket normalizálni kell. Minden egyes formánsérték két koordinátaponttal jellemezhető: egy időkoordinátával (a hangon belül) az *x*-tengelyen (mp-ben) és egy frekvenciakoordinátával az *y*-tengelyen (Hz-ben). Az időben változó formánsmozgások vizsgálatához mindkét koordinátában szükség van az összemérhetőséget biztosító normalizált adatokra.

A normalizált állapotot az időtengelyen a mérési pontok egységes meghatározása eleve biztosítja. Itt nem volt szükség további teendőre, mert a mérési pontok a magánhangzó hangidőtartamának százalékos arányában lettek megállapítva (10%, 25%, 50%, 75%, 90%) az adatbázis összes magánhangzója esetén. A frekvenciatengelyen a sok szóba jöhető ún. magánhangzó-normalizálási eljárás közül a Lobanov-féle módszert (Lobanov 1971) választottuk, amely könnyen értelmezhető és jól használható a beszélők közötti különbségek eliminálására (pl. fiziológiai, anatómiai különbségek), míg a hangsor beszélőfüggetlen része, a fonemikus információ, változatlan marad (Flynn–Foulkes 2011).

A Lobanov-féle módszer a matematikai statisztikában megszokott *Z*-transzformáción alapul:

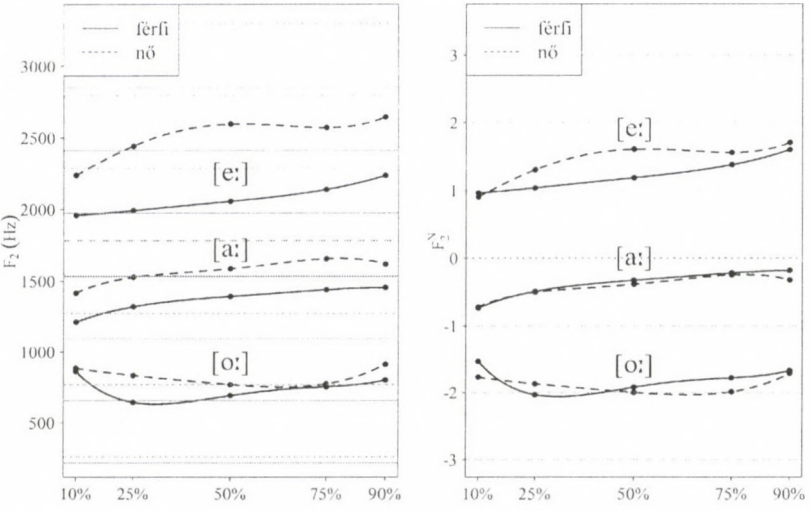
$$F_i^N = \frac{F_i - \overline{F_i}}{s_i}, i = 1,2$$

A módszer alapján beszélőnként és formánsonként kiszámoljuk a formánsértékek átlagát ($\overline{F_i}$) és szórását (s_i), melyekkel korrigáljuk az összes magánhangzó megfelelő formánsértékeit. Az 1. táblázatban összefoglaltuk az adatbázisunk magánhangzóira kiszámolt statisztikai mutatókat, két beszélőre és két formánsra csoportosítva.

A normalizálás fontos szerepét a 2. ábra mutatja be egy példán keresztül. Három magánhangzó F_2 formánsmozgását láthatjuk három szó (*megóv*, *lékvágó*, *különbség*) esetében. A bal oldali ábrán a normalizálás előtti állapotot, a jobb oldalin pedig a Lobanov-féle módszerrel korrigált mozgásokat láthatjuk.

1. táblázat: A Lobanov-féle normalizálás során felhasznált átlag- és szórásadatok az összes magánhangzó alapján

	Átlag ($\overline{F_i}$)		Szórás (s_i)	
	F_1 (Hz)	F_2 (Hz)	F_1 (Hz)	F_2 (Hz)
1. beszélő (férfi)	470	1535	122	439
2. beszélő (nő)	534	1781	153	506



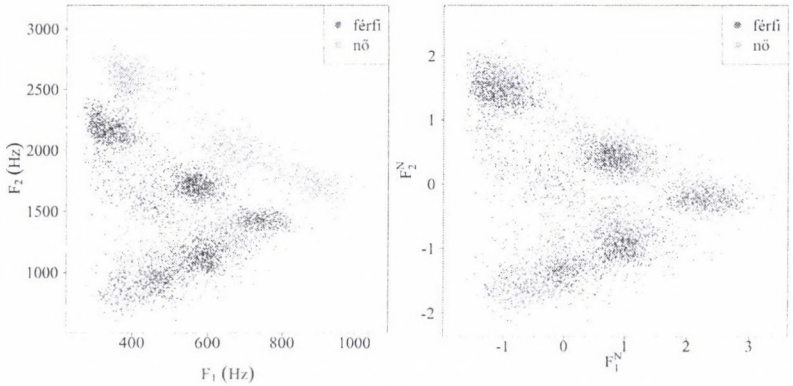
2. ábra
A normalizálás hatása a formánsértékekre. Az átlagtól azonos szórásnyi távolságra lévő frekvenciaértékek (bal oldali ábra) a normalizált formánstérben azonos pozícióba kerülnek (jobb oldali ábra)

A normalizálás eredményének jobb megértését rácsvonalakkal segítjük. A bal oldali ábrán folytonos vízszintes vonallal jelöltük a férfi beszélő F_2 formánsmozgás-átlagát (1535 Hz) a teljes adatbázisra vonatkozóan (lásd 1. táblázat), továbbá az egy szórásra lévő frekvenciákat (1096 Hz = 1535 Hz – 439 Hz és 1974 Hz = 1535 Hz + 439 Hz), valamint a két szórásra (657 Hz és 2413 Hz) és a három szórásra lévő értékeket (218 Hz, 2852 Hz). Szaggatott vízszintes vonallal jelöltük a női beszélő F_2 formánsmozgásának hasonló adatait: átlag (1781 Hz), egy szórásnyi távolságok (1275 Hz és 2287 Hz), két szórásnyi távolságok (769 Hz, 2793 Hz) és a három szórásnyi távolságok (263 Hz és 3299 Hz). A jobb oldali ábrán a normalizált formánstér –3, –2, –1, 0, 1, 2 és

3 pontjait emeltük ki. Az egyes értékek az átlagtól való szórásnyi eltérések együtthatóit jelentik.

A normalizálás úgy értelmezhető, hogy az átlagtól azonos szórásnyira lévő formánsértékek átfedik egymást a normalizált formánstérben különböző beszélők esetén. A 2. ábra bal oldali részén az [a:] hangra megadott középső két formánsmozgás mutatja, hogy a görbék a beszélőre jellemző saját átlaguknál csaknem azonos (saját szórásukban számolt) mértékben alacsonyabban helyezkednek el, így a jobb oldali ábrán ezek a görbék szinte átfedik egymást.

Hasonló átfedést láthatunk a 3. ábrán is, amely az F_1-F_2 síkon a vizsgált formánsadatbázis összes formánsértékét mutatja be normalizálás előtt és normalizálás után. A tanulmány további részében ezeket a normalizált frekvenciaadatokat használjuk a Hz-ben mért értékek helyett.



3. ábra

Az adatbázis formánsértékei az 50%-os mérési pontokon mérve az F_1-F_2 síkon, normalizálás előtt (balra) és normalizálás után (jobbra)

A formánsmozgások stabilitása

Adataink normalizálása után lehetőségünk van az összes formánsmozgás megrajzolására, az összes magánhangzóra az összes mássalhangzó-képzési hely függvényében beszélőtől függetlenül. Egy adott magánhangzó esetén a 6 képzési helyből képzett mátrixban, összesen 36 cellában rajzolhatjuk meg a különböző hangkörnyezetekben megvalósuló formánsmozgásokat. Az összes magánhangzóra vonatkoztatva 9 db ilyen mátrixot készítettünk formánsonként. A formánsmozgásokat az 5 pontra illesztett negyedfokú polinommal közelítjük. Az F_1 és F_2 formánsmozgások egy-egy mátrixára láthatunk példát a 4. ábrán az [o, o:] hangra. A 324 különböző $C_{kh.1}V_mC_{kh.2}$ kategóriában megvalósuló formánsmozgást áttekintve azt mondhatjuk, hogy F_1 és F_2 tekintetében, minden kategória esetén, szemre, jellemzően azonos formánsmenet álla-

pítható meg. Az objektívebb értékelés érdekében egy ún. stabilitási mutatót (SM) vezettünk be, amely minden kategória esetén egy számértékkel jelzi a benne megvalósuló formánsmozgások hasonlóságát. Minél közelebb van nullához a mutató értéke, annál hasonlóbbak a cella formánsmozgásai.

Egy adott $C_{kh.1}V_mC_{kh.2}$ hanghármasra jellemző stabilitási mutató kiszámításához először meghatározzuk a cella átlagos görbéjét, az átlaggörbét (vagy a klaszteranalízisből kölcsönzött szóval a centroidját), amelyet az azonos időkoordinátával rendelkező formánsértékek átlagolásával kapunk meg. Az átlaggörbe 5 pontjának formánsértékeit jelöljük $\bar{f}_i$ -vel (i = 1, 2, 3, 4, 5). A stabilitási mutató az átlaggörbétől mért átlagos euklideszi távolság, azaz

$$SM = \frac{\sum_{j=1}^{n} \sqrt{\sum_{i=1}^{5} (f_{ji} - \bar{f}_i)^2}}{n},$$

ahol n a cellában előforduló görbék száma, f_{ji} a j. görbe i. pontjának normalizált frekvenciaértéke. A stabilitási mutató alkalmas az egyes cellákban megvalósuló formánsmozgások hasonlóságának jellemzésére. A mutató biztosítja, hogy az adatbázisban gyakrabban előforduló hanghármasok (pl. DAV_mDA) és a ritkább kapcsolatok (pl. VEV_mPA) is összehasonlíthatók legyenek, hiszen átlagértékről van szó. A kiszámolt stabilitási mutatók segítségével a 324 $C_{kh.1}V_mC_{kh.2}$ hanghármas akár sorba is rendezhető, továbbá csak a magánhangzók és csak a hangkörnyezetek esetében is felállítható egy-egy stabilitási sorrend.

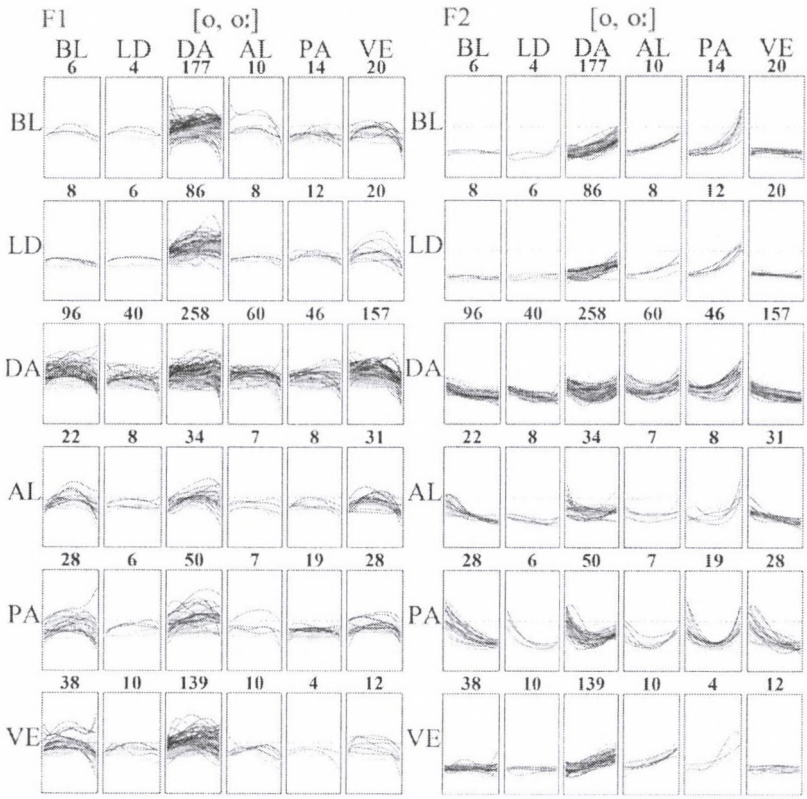
Az [o, o:] magánhangzó kategória stabilitási mutatóit különböző hangkörnyezetekben a 2. táblázatban közöljük. A kisebb érték stabilabb formánsmozgás halmazt jelent a cellán belül.

2. táblázat: Az [o, o:] magánhangzó minőség stabilitási mutatói, különböző hangkörnyezetekben F_1-re és F_2-re

	F_1						F_2					
	BL	**LD**	**DA**	**AL**	**PA**	**VE**	**BL**	**LD**	**DA**	**AL**	**PA**	**VE**
BL	0,47	0,55	0,91	1,13	0,65	0,78	0,19	0,39	0,42	0,36	0,72	0,31
LD	0,45	0,44	0,84	0,45	0,41	0,77	0,31	0,28	0,38	0,56	0,48	0,24
DA	0,98	0,81	0,84	0,98	0,87	0,81	0,42	0,37	0,55	0,59	0,78	0,36
AL	0,74	0,44	0,70	0,64	0,57	0,54	0,47	0,35	0,67	0,78	0,89	0,33
PA	1,06	0,53	1,09	0,85	0,47	0,66	0,87	0,59	0,81	0,53	0,90	0,83
VE	1,02	0,79	0,87	0,66	0,62	0,79	0,32	0,27	0,40	0,50	0,84	0,26

Az átlagok az adatbázis összes első és második formánsában megvalósuló formánsmozgás alapján: $SM(F_1)$ = 0,613, $SM(F_2)$ = 0,474. A két formánsban megvalósuló formánsmozgások stabilitása szignifikánsan eltér (Wilcoxon-próba: $p < 0{,}001$). Ezek alapján azt mondhatjuk, hogy a különböző

$C_{kh.1}V_mC_{kh.2}$ hanghármasokon belül az F_2 formánsmozgásai egységesebbek az F_1 formánsmozgásaihoz viszonyítva, vagyis a hangkörnyezet egymáshoz hasonlóbb mozgásformákat alakít ki a második formánsban, mint az első formánsban. A stabilitási mutató alkalmas a $C_{kh.1}V_mC_{kh.2}$ hanghármasokban előforduló formánsmozgások hasonlóságának leírására, de a formánsmenet alakjáról nem nyújt tájékoztatást. A továbbiakban a cellákat leíró jellemző formánsmenetek meghatározását mutatjuk be.



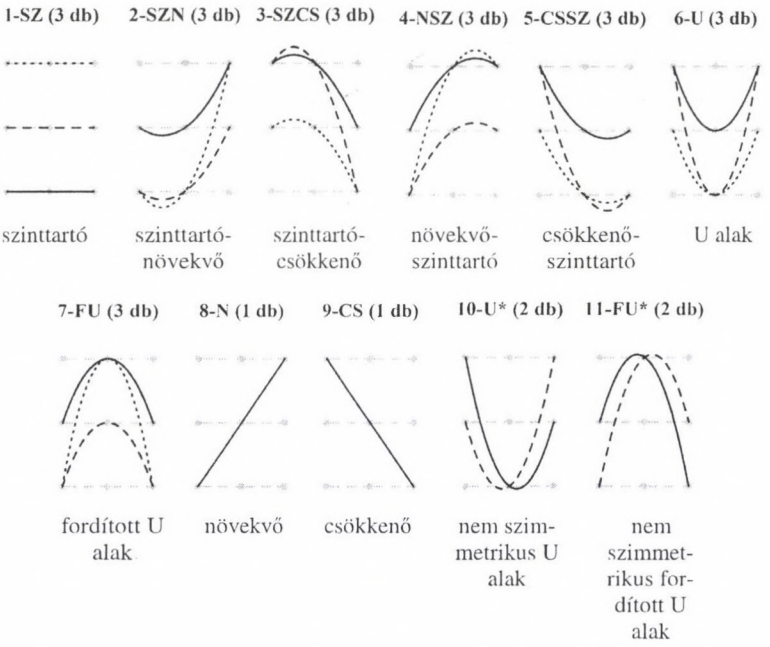
4. ábra

Az [o, o:] magánhangzók F_1 és F_2 formánsmozgásainak grafikus összesítése $C_{kh.1}V_mC_{kh.2}$ helyzetekben a mérési pontokból származtatva. A megelőző mássalhangzó a mátrix sorainak, a követő a mátrix oszlopainak nevében szerepel. A cellák felett a hanghármas előfordulások (görbék) száma olvasható

A formánsmozgások modellezése: formánsmenettípusok

Eljárást dolgoztunk ki a $C_{kh.1}V_mC_{kh.2}$ kapcsolatokban előforduló formánsmozgások kategorizálására. Először a formánsmozgások tipizálását végeztük el. Az 5. ábrán a matematikailag lehetséges ideális formánsmenetek rendszerét mutatjuk be. 27 db lehetséges ideális formánsmozgást definiáltunk, amelyeket 11 formánsmenettípusba soroltunk. Minden formánsmenettípus 9 pontot tartalmaz, 3 sorban és 3 oszlopban elrendezve. A 3 oszlop jelentése: a $C_{kh.1}V_mC_{kh.2}$ hármas egyes elemeinek megfelelő artikulációs konfiguráció akusztikus vetülete. A három sor pedig a frekvenciaértékek 3 lehetséges szintje, melynek alapja, hogy a mássalhangzó-környezetbe helyezett magánhangzók frekvenciaértékeinek 2, ill. 3 fokozata figyelhető meg (Magdics 1965: 24). Az 5. ábra első formánsmenettípusának felirata: 1-SZ (3db), a neve szinttartó, és összesen 3 db lehetséges formánsmenetet tartalmaz (a pontozott, szaggatott és folytonos vonaltípusok különböztetik meg őket). Az egy típusba tartozó formánsmenetek a grafikon 9 pontjából úgy kötnek össze 3 pontot, hogy a 3 pont által meghatározott két átmeneti szakaszban, formánsmenetenként azonos a legyen mozgás iránya. A mozgás iránya lehet szinttartó, növekvő vagy csökkenő. Az 1. formánsmenettípus esetén mindkét átmenet szinttartó mindhárom formánsmenetben. A 2–5. grafikonokon további 3-3 ideális formánsmenetet látunk, az átmenetek itt: szinttartó-növekvő, szinttartó-csökkenő, növekvő-szinttartó és csökkenő-szinttartó. A 6. és 7. grafikon 3-3 szimmetrikus menete a csökkenő-növekvő és növekvő-csökkenő eseteknek felel meg. A 8. és 9. grafikon egy-egy ideális formánsmenete növekvő-növekvő illetve csökkenő-csökkenő mozgást mutat. A 10. és 11. csillaggal jelölt formánsmenettípusokról később ejtünk szót. Az 5. ábra egy formánsmenete tehát kiindul valamely hangátmeneti frekvenciából, átlagosból (középen), átlag felettiből vagy átlag alattiból (3 eset), elér egy tiszta fázisbeli értékhez [itt is 3 eset van: átlagos (középen), átlag feletti és átlag alatti], és zárul egy hangátmeneti frekvenciaértékkel (szintén 3 eset). Az ideális formánsmenetek száma így $3 \times 3 \times 3 = 27$ darab.

Az 5. ábra 11 darab formánsmenettípusa a görbéket alkotó 3 pont mozgatásával átvihető egymásba. Távolságot definiálhatunk két formánsmenettípus között: minimálisan hány pontnyi mozgással juthatunk el az egyik típusból a másikba (3. táblázat). A távolság szimmetrikus tulajdonsága miatt a 3. táblázatban elegendő a felső háromszöget kitölteni. Az utolsó, Σ sorból azt olvashatjuk ki, hogy az adott típushoz képest az összes többi típus együttvéve milyen távol van. Minél kisebb ez a szám, annál hasonlóbb a többihez az adott típus (a szinttartó van a legközelebb a többihez). A Σ érték kiszámításához a felső háromszögből tükröztük az adatokat az alsó háromszögbe, majd összegeztük az oszlopokat.



5. ábra
A formánsmenettípusok 11 elméleti kategóriája. Zárójelben a típus által lefedett formánsmenetek száma látható

Amennyiben rendszerezni szeretnénk az adatbázisbeli formánsmozgásokat a fenti modell szerint, akkor minden konkrét formánsmozgást az 5. ábrán megjelenő 11 formánsmenettípus valamelyikébe kell besorolnunk. Nem gondoljuk, hogy a fenti séma minden formánsmozgást tökéletesen leír, hiszen alapvetően egyenes (1, 8–9), illetve parabola jellegű mozgással (2–7, 10–11) igyekszünk leírni a mozgásformákat. Ismert, hogy a formánsmozgások mindegyike pontosan nem írható le parabolikus görbével (Harrington 2010: 143), és már a mi 5 ponton alapuló mozgásaink között is elképzelhető az 5. ábrán látható formánsmenetektől lényegesen eltérő forma. Szükség van tehát egy 12. egyéb kategóriára, amelybe azok a formánsmozgások kerülnek, amelyek egy adott hibaküszöböt meghaladó mértékben illeszkednek csak az 1–11 típusokra. A módszerünk annál hatékonyabb, minél kisebb elemszámú a 12. kategória, vagyis minél kevesebb elemet kell figyelmen kívül hagynunk a további elemzésből. A módszer másik sarkalatos pontja a kategóriák közötti döntés biztossága. Látjuk, hogy sok 1 távolsággal rendelkező típuspár létezik, az ezek közötti döntés nagy körültekintést igényel.

3. táblázat: A 11 formánsmenettípus egymáshoz mért távolsága. Az utolsó sorban a teljesen kitöltött mátrix oszlopainak összege látható. A sor és oszlopnevek magyarázata az 5. ábrán szerepel

	1	2	3	4	5	6	7	8	9	10	11
1-SZ	0	1	1	1	1	1	1	2	2	2	2
2-SZN		0	2	1	2	1	2	1	3	1	3
3-SZCS			0	2	1	2	1	3	1	3	1
4-NSZ				0	2	2	1	1	3	3	1
5-CSSZ					0	1	2	3	1	1	3
6-U						0	3	2	2	1	3
7-FU							0	2	2	3	1
8-N								0	4	2	2
9-CS									0	2	2
10-U*										0	4
11-FU*											0
Σ	**14**	**17**	**17**	**17**	**17**	**18**	**18**	**22**	**22**	**22**	**22**

Az adatok illesztése a formánsmenetekhez

A magánhangzó hangidőtartamában az 5 mérési ponton mért értékeket jelöljük m_i-vel, ahol i = 1, 2, 3, 4, 5. Az m_1 a 10%-os mérési pontban kapott normalizált formánsérték, az m_2 a 25%-os mérési ponthoz tartozó érték és így tovább. Ha képezzük az egymás mellett lévő pontok különbségét, akkor a mozgás irányáról kapunk képet. A $d_i = m_{i+1} - m_i$ (i = 1, 2, 3, 4) összefüggés segítségével kapott különbség pozitív, ha növekedés van a formánsmogásban, negatív, ha csökkenés, illetve nulla (körüli), ha nincs változás a mozgásban.

Az 5. ábrán közölt 11 formánsmenettípust használjuk fel egy adott pontötössel meghatározott formánsmozgás jellemzésére. Az eljárás során a fent definiált d_i különbségek játszanak központi szerepet. Azzal, hogy áttérünk az 5 db m_i formásértékről a 4 db d_i különbségre a formánsmozgás y-koordináta szerinti helyét, az y-eltolás mértékét, elveszítjük. Mindössze a mozgásról marad információ, ami, mint látni fogjuk nagyon hasznosnak bizonyul az eljárás további részében.

Az 5. ábra 1–9-ig sorszámozott formánsmenet típusait jellemezhetjük pozitív c, c_1 és c_2 konstansok (valós számok), valamint a nulla segítségével. A 4. táblázat összefoglalja a formánsmenettípusokhoz tartozó konstansnégyeseket. A konstansnégyes elemei a mért pontok különbségeinek (d_i) felelnek meg. Az egyes formánsmenettípusok esetén a konstansok előírt értékei a lehetséges mozgási formákat (vagyis a d_i különbségeket) úgy korlátozzák, hogy segítségükkel az 5. ábrán szereplő formánsmeneteknek megfelelő kategorizálást el tudjuk végezni. Ezek a korlátozó tényezők a 4. táblázat „Konstansok" oszlopából olvashatók ki. A korlátozó előírások 3 típusba sorolhatók. Az említett oszlopban megjelenő nulla jelentése: a konstans értéke legyen nulla (kö-

rüli). Az előjel nélküli és negatív előjellel szereplő betűk (pl. c_1, $-c_2$) jelentése: a konstans nem egyenlő nullával, ha nincs előjel, pozitív, ha van, negatív értékű. Az utolsó előírás, hogy egy formánsmenettípuson belül az azonos konstansok azonos értéket jelölnek. Ha tekintjük például az 1-SZ szinttartó formánsmenet típust, akkor 5 olyan mért pont, amely illeszkedik a vízszintes egyenesre, különbségként 4 db nulla értéket szolgáltat. A 8-N növekvő esetben az 5 pont különbsége 4 db azonos nagyságú pozitív érték lesz (c, c, c, c). Utolsó példaként a 6-U típust leíró 5 képzeletbeli pont különbsége két negatív érték ($-c_2$, $-c_1$), majd fordított sorrendben pozitív előjellel ugyanaz a két érték (c_1, c_2).

4. táblázat: A formánsmenettípusok illesztése során használt elnevezések, konstansok és kiszámításuk

Formánsmenettípusok		Konstansok	Kiszámítás
1-SZ	szinttartó	0, 0, 0, 0	–
2-SZN	szinttartó-növekvő	0, 0, c_1, c_2	$c_1 = d_3$ $c_2 = d_4$
3-SZCS	szinttartó-csökkenő	0, 0, $-c_1$, $-c_2$	$c_1 = -d_3$ $c_2 = -d_4$
4-NSZ	növekvő-szinttartó	c_2, c_1, 0, 0	$c_1 = d_2$ $c_2 = d_1$
5-CSSZ	csökkenő-szinttartó	$-c_2$, $-c_1$, 0, 0	$c_1 = -d_2$ $c_2 = -d_1$
6-U	U alak	$-c_2$, $-c_1$, c_1, c_2	$c_1 = (d_3 - d_2)/2$ $c_2 = (d_4 - d_1)/2$
7-FU	fordított U alak	c_2, c_1, $-c_1$, $-c_2$	$c_1 = (d_2 - d_3)/2$ $c_2 = (d_1 - d_4)/2$
8-N	növekvő	c, c, c, c	$c = (d_1 + d_2 + d_3 + d_4)/4$
9-CS	csökkenő	$-c$, $-c$, $-c$, $-c$	$c = (-d_1 - d_2 - d_3 - d_4)/4$
10-U*	nem szimmetrikus U alak	$-c_2$, $-c_1$, c'_1, c'_2	–
11-FU*	nem szimmetrikus fordított U alak	c_2, c_1, $-c'_1$, $-c'_2$	–

Minden adatbázisbeli formánmozgást a fenti 1–9 típushoz illesztettünk, a formánsmozgást leíró (d_i) különbségek segítségével. Az illesztés során a hiba minimalizálása mellett kerestük a c (8. és 9. esetben), illetve a c_1, c_2 konstansokat (2–7. esetekben). Ha a 4. táblázat konstansait és nulláit egy formánsmenettípus esetén rendre c_1^*, c_2^*, c_3^*, c_4^*-gal jelöljük, akkor az illesztés során a következő kifejezést kell minimalizálnunk:

$$\sum_{i=1}^{4} (d_i - c_i^*)^2 \rightarrow \min$$

A minimum keresése során kapott c, c_1, c_2 konstansok kiszámítása a 4. táblázat harmadik oszlopából olvasható ki. Az adatbázis minden formánsmozgásához hozzárendeltük a konstansok meghatározása során kapott legkisebb hibával rendelkező formánsmenettípus sorszámát. Így minden formánsmozgást azzal a számmal jelölt mozgáskategóriával jellemeztünk a 4. táblázat szerinti sorszámokból 1–9-ig, amelyre a legjobban illeszkedik. Itt két fontos kiegészítésre van szükség. *a)* Az 1. formánsmenettípus minden további mozgás (2–9) speciális esete. Egy 8. kategóriába tartozó formánsmozgás pozitív c értékei annyira közel lehetnek a 0-hoz, hogy az eredeti növekvő kategóriát szinttartóra kell cserélnünk. Bizonyos esetekben tehát a 2–9. sorszámozott formánsmozgásokat az 1. kategóriába kell átsorolnunk. Az átsorolás elvégzéséhez a c, c_1, c_2 konstansokat használjuk fel. Azokat az eseteket soroltuk át az 1. kategóriába, amelyeknél minden $|c_i^*| < 0{,}12$ ($i = 1, 2, 3, 4$). A küszöb értékének meghatározásánál azt a szempontot vettük figyelembe, hogy az 1. kategória elemszáma a többi átlagával egyezzen meg. *b)* A másik kiegészítés az 5. ábra 10. és 11. kategóriájára vonatkozik. Ezek a kategóriák 4 különböző konstanssal írhatók le, azaz az illesztés során mindig 0 minimumot szolgáltattak volna előjelhelyes esetekben. Ezen két kategóriába azokat az eseteket soroljuk át az 1–9. kategóriákból, amelyeknél az előjel megfelelő, és az illesztés hibája a saját formánsmenet-kategóriához egy adott küszöbértéket meghalad. Külön küszöbértéket használtunk fel F_1-re és F_2-re: a felső 15%-os legnagyobb illeszkedési hibával rendelkező formánsmozgást soroltuk át mindkét esetben. Az átsorolás során a küszöbérték feletti hibaillesztéssel rendelkező, de a konstansok előjelei alapján a 10–11. kategóriákba be nem sorolható görbéket egy új 12. (egyéb) kategóriába helyeztük. Miután az adatbázis összes formánsmozgását besoroltuk valamely kategóriába (1–12), lehetőségünk van a formánsmenettípusok gyakorisági táblázatát megadni külön az első és külön a második formánsra (5. táblázat).

5. táblázat: Az 1–11 formánsmenettípusok eloszlása és a 12. kategória százalékos gyakorisági táblázata F_1 és F_2 esetén. (Az adatok egész %-ra kerekítettek.) A cellákban szereplő számok az adott formánsmenettel leírható formánsmozgások százalékát jelentik a teljes adatbázisban. Az oszlopnevek magyarázata az 5. ábrán szerepel

	1	2	3	4	5	6	7	8	9	10	11	12	
F1	11	6	22	13	1	1	30	2	1	0	8	7	100%
F2	13	14	11	9	9	7	7	8	5	1	1	12	100%

Az 5. táblázat szerint az F_1 mozgási formái kevesebb számú formánsmenettel jellemezhetők, mint az F_2-éi: az F_1 formánsmozgásainak többsége leír-

ható 5-6 db formánsmenettípussal, míg F_2 leírásához 8-9 ilyen kategória szükséges. A mássalhangzó-környezet tehát elsősorban a magánhangzó második formánsának mozgására hat.

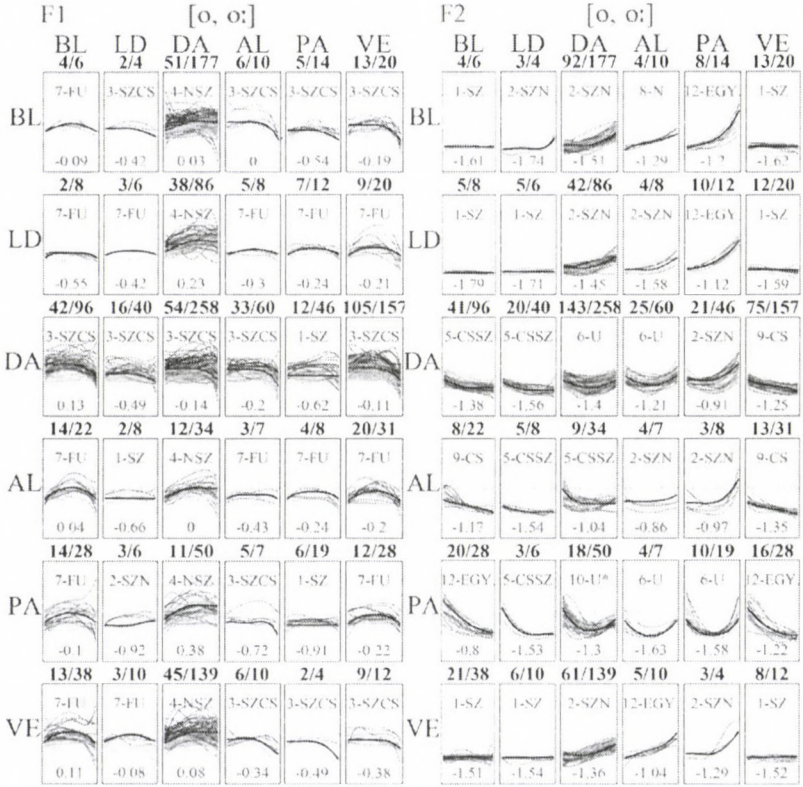
A képzési helyek és a formánsmenettípusok kapcsolata

Ha hozzárendeltünk minden formánsmozgáshoz egy formánsmenettípust, akkor megvizsgálhatjuk, hogy adott magánhangzó esetén vajon az azonos hangkörnyezethez azonos formánsmenettípussal rendelkező formánsmenetek tartoznak-e, a képzési helyre vonatkoztatva. Ez a korábban már látott stabilitásnak egy szigorúbb megközelítése, mert úgy kell hasonlónak lennie a magánhangzón belüli formánsmozgásoknak, hogy azok az 1–11 formánsmenettípus egyikével is jól egyezzenek. Most a legstabilabb esetet a csupa azonos típusú mozgások jelentik, a leginstabilabbat pedig az a helyzet, amikor mindegyik típus (1–11) azonos arányban fordul elő az adott hangkörnyezetben. Végső soron arra a kérdésre keressük a választ, hogy az azonos képzési hely szerinti kategóriák a C-ben azonos formánsmenetet eredményeznek-e a V_m-ben. Ha a kérdésre igenlő a válasz, akkor az egyes cellák formánsmenetei jól leírhatók valamely ideális formánsmenettel, természetesen cellánként kissé eltérő mértékben. A cellához tartozó jellemző formánsmenettípus (1–12) meghatározásához a cellában lévő *N* darab görbe típusainak megfigyelt gyakoriságait (O_i) használtuk fel: a cellához azt az *i* típust rendeljük hozzá, amelyből a legtöbb volt a cellában, azaz amelyre az $SM_2 = O_i/N$ hányados maximális értéket ad. A szigorú stabilitási mutató (SM_2) értéke 0 és 1 között változhat, minél közelebb van az egyhez, annál stabilabban írja le a formánsmenettípus az adott cellát. A 6. ábrán az [o, o:] magánhangzó celláiban – a formánsmozgások felett és alatt – a hanghármashoz rendelt formánsmenettípust és annak az 50%-os mérési pont alapján számolt *y*-eltolás értékét tüntettük fel. Minden cella esetében az SM_2 értéket is feltüntettük a cellák felett O_i/N formában, ahol az adott típusú görbék számát O_i és az összes görbe számát *N* adja meg.

A 6. táblázatban összefoglaltuk, hogy a $C_{kh.1}V_mC_{kh.2}$ hanghármasokat milyen arányban és milyen átlagos stabilitással jellemzik az egyes formánsmenettípusok. Az első sorban lévő 20 és 0,47 érték azt fejezi ki, hogy az első formánsban az 1-SZ formánsmenettípus a lehetséges 324 $C_{kh.1}V_mC_{kh.2}$ hanghármas 20%-át fedi le, és a cellák átlagos SM_2 értéke 0,47. A 6. táblázat alapján levonható konklúzió hasonló az 5. táblázatnál leírtakhoz, de itt nem az összes magánhangzóra vonatkoznak a következtetések, hanem a 324 darab $C_{kh.1}V_mC_{kh.2}$ hanghármasra. A $C_{kh.1}V_mC_{kh.2}$ hanghármasokhoz rendelt formánsmenettípusok F_2-ben változatosabbak, mint F_1-ben, és a hozzárendelés stabilitása is nagyobb F_2-ben.

A 6. ábra $C_{kh.1}V_mC_{kh.2}$ hanghármasait jellemző formánsmenettípusok *y*-eltolását a cellához rendelt formánsmenettípussal megegyező kategóriájú formánsmozgások 50%-os mérési pontjainak átlagolásával kapjuk meg. A 7. áb-

ra az összes *y*-eltolás értékének eloszlását szemlélteti. Az ábrán mindkét formáns esetén 3 módusz határozható meg, amely alapján az *y*-eltolásokat 3 típusba soroljuk (alacsony, közepes és magas). Például az első formáns esetében alacsony kategóriába soroltuk a 0,2-nél kisebb normalizált frekvenciaértékkel rendelkező formánsmozgásokat. Az *y*-eltolás típusainak eloszlása F_1 esetén alacsony: 67%, közepes: 22% és magas: 11%, míg F_2 esetén alacsony: 29%, közepes: 47% és magas: 24%. Az *y*-eltolások tekintetében azt állapíthatjuk meg, hogy F_1-ben gyakoribbak az átlag alatti tiszta fázisban mért normalizált frekvencia értékek, míg F_2-ben az átlag körüliek a jellemzőbbek.

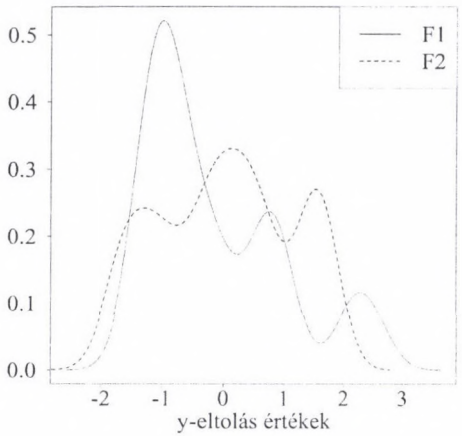


6. ábra

Az [o, o:] magánhangzó különböző hangkörnyezetbeli F_1 és F_2 formánsmozgásaihoz rendelt formánsmenettípusok és az *y*-eltolás mértéke. A vastagított görbe az adott cella standard formánsmenetét közelíti, a cella típusával egyező formánsmozgások c_1^*, c_2^*, c_3^*, c_4^* értékei alapján

6. táblázat: A formánsmenettípusok eloszlása és az átlagos SM_2 értékek a $C_{kh.1}V_mC_{kh.2}$ hanghármasok között. A sornevek magyarázata az 5. ábrán szerepel

Formánsmenet	F1		F2	
	Relatív gyakoriság (%)	SM_2-átlag	Relatív gyakoriság (%)	SM_2-átlag
1-SZ	20	0,47	14	0,55
2-SZN	5	0,43	16	0,49
3-SZCS	27	0,53	15	0,49
4-NSZ	5	0,46	6	0,52
5-CSSZ	0	–	10	0,51
6-U	0	0,50	8	0,54
7-FU	36	0,56	8	0,53
8-N	1	0,38	7	0,53
9-CS	0	–	6	0,60
10-U*	0	–	1	0,35
11-FU*	6	0,53	0	–
12-EGYÉB	0	–	9	–



7. ábra
Az *y*-eltolások eloszlása F_1 és F_2 esetén

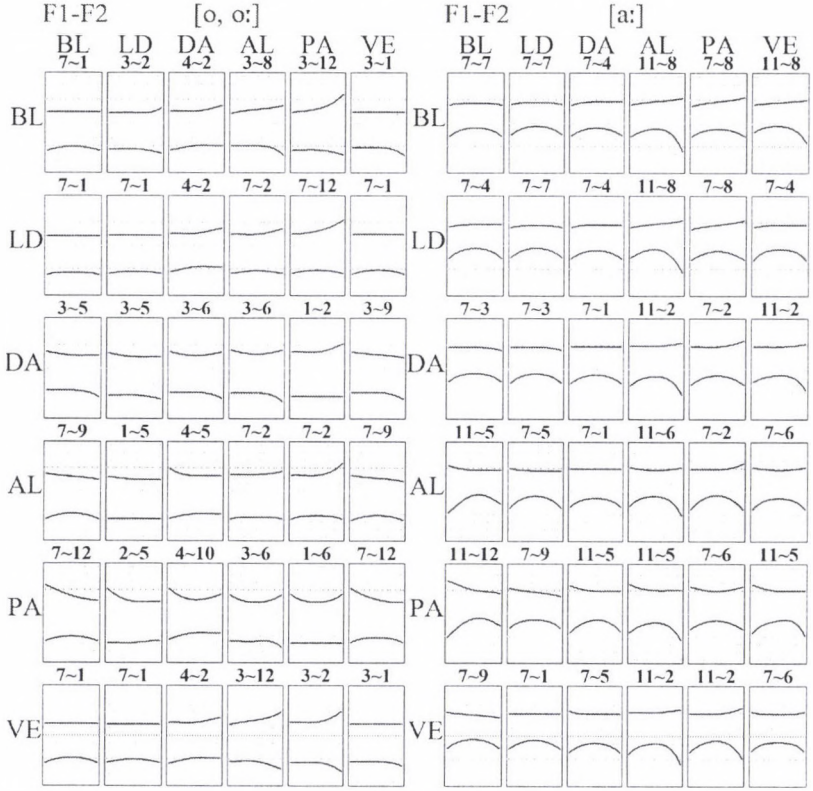
Formánstérképek

A jelen kutatás egyik végcélja a formánstérképek elkészítése volt. A formánstérkép az első és második formánsban eddig végrehajtott kategorizálások eredményeit egyetlen ábrán jeleníti meg, hasonlóan ahhoz, ahogyan a

hangspektrogramokon is látni a formánsokat. Létrehoztunk 9 × 36 db formánstérképet – a már megszokott mátrixos elrendezésben – a 9 magánhangzótípusra a mássalhangzó képzési helyeinek függvényében. A 8. ábrán az [o, o:] és az [a:] magánhangzó formánstérképeit láthatjuk. Itt is, ahogyan a 6. ábrán is, a berajzolt formánsmenetek a cella típusával egyező formánsmozgások átlagolt c_1^*, c_2^*, c_3^*, c_4^* értékein alapulnak. Egy formánstérkép az adott cellára jellemző első és második formánsmenettípust és azok *y*-eltolásait tartalmazza. Az F_1 és F_2 formánsmenetét az *y*-eltolások figyelembe vételével tudjuk egymás fölé rajzolni, úgy, hogy az összeillesztés pontja F_1 esetében a +3-as, F_2 esetében a −3-as *y*-koordináta volt. A összeillesztési pontot a formánstérkép közepén egy folytonos vastag vonal jelzi. A vonal alatt az első formáns mozgását, a vonal felett a második formáns mozgását követhetjük nyomon. Mindkét részben a vastagabb szaggatott vonal az átlagos értéknek (nulla) felel meg, a pontozott vonalak pedig 1 szórással távolabbi pozíciót jelölnek. A formánstérképről cellánként a következő információ olvasható le: (1) F_1 menete, (2) F_1 *y*-eltolása a normalizált formánstérben, (3) F_2 menete, (4) F_2 *y*-eltolása és végül (5) F_1 és F_2 egymáshoz mért távolsága. A 8. ábrán a formánstérképek felett az F_1 és F_2 formánsmenettípusok sorszámát adtuk meg, az *y*-eltolások és az *y*-eltolások különbségeként kiszámítható F_1–F_2 távolságok leolvasását a rácsvonalak teszik lehetővé.

A formánstérképek megadásával létrehoztuk a magyar formánsmozgásokat általánosságban képviselő rendszerezett adathalmazt a vizsgált hármas hangkapcsolatokra. A formánstérképek lehetőséget adnak arra, hogy az összes F_1-F_2 formánsmenettípust együtt vizsgáljuk, összehasonlítsuk. A formánstérképeken leggyakrabban előforduló formánsmenettípus-párt és annak gyakoriságát a 7. táblázat tartalmazza. Adatbázisunk 324 db különböző hanghármas kategóriája 57 db formánstérképpel írható le, ha az *y*-eltolásokat figyelmen kívül hagyjuk. A leggyakrabban előforduló típuspár a 3-SZCS~2-SZN (vagyis F_1-ben szinttartó-csökkenő, F_2-ben szinttartó-növekvő), amely 6,52%-ban fordul elő a 324 cellából (21 cella, azaz ennyiféle $C_{kh.1}V_mC_{kh.2}$ kapcsolat). A 7. táblázat „SM_2 (F_1)" és „SM_2 (F_2)" oszlopai a cellákban lévő formánsmozgások adott típusú formánsmenetként való beazonosításának átlagos arányát mutatják. Az első sort figyelembe véve, a 21 cellában átlagosan 53%-ban található 3-SZCS típusként azonosított formánsmozgás F_1-ben, és 51%-ban 2-SZN F_2-ben. Az ideális, 100%-os eset nagyon ritka, mindössze 8, illetve 7, kevés formánsmozgást tartalmazó cellában fordul elő F_1 és F_2 esetében (F_1-ben pl.: BL[ɔ]BL, BL[a:]LD, AL[a:]DA és VE[a:]VE; F_2-ben pl.: BL[ɔ]BL, VE[i, i:]BL, AL[ø, ø:]LD, PA[ø, ø:]DA). A táblázat utolsó sorában látható átlagok (49% F_1-ben és 54% F_2-ben) magyarázatához tisztában kell lennünk azzal, hogy a cellában lévő formánsmozgások elvileg 12 típus bármelyikébe is tartozhatnának, a legjellemzőbb azonban a táblázatban feltüntetett arányban van jelen. Ezen túlmenően a cella második leggyakoribb formánsmenetét is megvizsgáltuk, és amennyiben 1 távolságra volt a cella meg-

határozó formánsmenettípusához (a távolságokhoz lásd a 3. táblázatot), akkor összevontuk a két típust. Összevonást tudtunk végrehajtani F_1-ben az esetek 94%-ában, F_2-ben pedig a cellák 96%-ában. Az egyesítés megnövelte az egyezési arányt, ahogyan azt a 7. táblázat utolsó két oszlopa is tartalmazza, miközben a mozgás alakján lényegesen nem változtattunk. Az átlagos értékek 75%-ra nőttek F_1-ben és 79%-ra F_2-ben. Ezek alapján azt mondhatjuk, hogy az 5. ábrán leírt formánsmenetmodellel a $C_{kh.1}V_mC_{kh.2}$ hanghármasok formánsmozgásai jól leírhatók, de a formánsmozgások folytonosan töltenek ki egy határozottan körülírható részt a normalizált formánstérben.



8. ábra

Az [o, o:] és az [a:] magánhangzó formánstérképei a mássalhangzó képzési helyének függvényében

7. táblázat: A formánsmenet térképeken előforduló F_1-F_2 formánsmenettípuspárok gyakoriság szerint csökkenő sorrendben

	F_1-F_2 formánsmenettípusok	Relatív gyakoriság (%)	SM_2 (F_1)	SM_2 (F_2)	Összevont SM_2 (F_1)	Összevont SM_2 (F_2)
1.	3-SZCS~2-SZN	6,52	0,53	0,51	0,79	0,78
2.	7-FU~1-SZ	6,21	0,60	0,51	0,84	0,72
3.	1-SZ~3-SZCS	5,90	0,51	0,58	0,75	0,83
4.	7-FU~3-SZCS	4,97	0,51	0,60	0,75	0,79
5.	7-FU~2-SZN	4,66	0,54	0,50	0,82	0,74
6.	3-SZCS~1-SZ	3,73	0,55	0,47	0,79	0,73
...	...	...	...	...	...	...
57.	8-N~3-SZCS	0,31	0,50	0,83	0,66	1,00
Átlag			0,49	0,54	0,75	0,79

Választ kerestünk arra a kérdésre, hogy hányfajta formánstérkép különböztethető meg a magyar formánstérben a V_m-ekre vonatkoztatva (a mért hangkapcsolati formákban). Az előző pontban láttuk, hogy 57 formánsmenettípuspár leírja a lehetséges 324 kategóriát, de ebben nem szerepelnek az F_1 és F_2 *y*-eltolásai. Ez a két adat tovább finomítja a lehetséges kategóriákat. Kérdés lehet, hogy minden cella azonosítható-e ezen négy információ alapján (F_1-F_2 menete és *y*-eltolásaik típusa). A négy faktor együttes figyelembe vétele az 57 típust 132 kategóriára bontotta tovább, azaz 132 különböző formánstérkép létezik az adatbázisunkban. A legnagyobb gyakoriságú 5-féle formánstérképet a 8. táblázat tartalmazza. A 324 cella 40%-ában olyan jellemző négyest kapunk, amely csak az illető $C_{kh.1}V_mC_{kh.2}$ hanghármasra jellemző, vagyis ezekben az esetekben egyértelműen meghatározza a formánstérkép a $C_{kh.1}V_mC_{kh.2}$ hármast. A cellák 24%-át olyan formánstérkép írja le, amely 2 cellához tartozik, továbbá a 14%-át 3 értékű, 10%-át 4 értékű és 5%-át 5 értékű formánstérképek jellemzik. A fentiek alapján azt mondhatjuk, hogy a $C_{kh.1}V_mC_{kh.2}$ magánhangzóiban lezajló formánsmozgások az F_1 és F_2 formánsmeneteinek figyelembevételével 57 különböző formánstérképet adnak, de ha az *y*-eltolások típusait is figyelembe vesszük, a cellák 60%-ában lehetőségünk van tömörítésre.

Az eredmények összehasonlítása, a formánstérkép mint mérőeszköz

A formánstérképek segítségével új típusú vizsgálatokat végezhetünk a magyar beszédhangok körében. Az eddigi vizsgálatok nagy része a mostani fogalmainkkal az *y*-eltolás értékére vonatkozott. A formánstérképek azonban a magánhangzó tiszta fázisát megelőző és követő részekről is nyújtanak információt. Egyelőre a szakirodalomban fellelhető néhány állítás ellenőrzésére mutatunk példát, ugyanis tézisszerű megfogalmazásokhoz az adatbázis méretét jelentősen növelnünk kell. Magdics szerint az első formáns mozgása kevésbé érzékeny a hangkörnyezetekre, de veláris magánhangzók esetén – ahol

az F_1 és F_2 távolsága kisebb – a frekvenciaértékek nagyobb változékonyságot mutatnak a palatálisokkal szemben (Magdics 1965: 47). Ennek ellenőrzésére az 1–12 formánsmenetek két osztályát hoztuk létre. A neutrális osztályba az 1. szinttartó formánsmozgástípus tartozott, a változó osztályba soroltuk be az összes többi mozgásformát. A hátul képzett magánhangzókhoz tartozó cellák 91%-ában változó mozgásformát találtunk F_1-ben, míg elöl képzett esetekben ez a $C_{kh.1}V_mC_{kh.2}$ hanghármasok 72%-át jellemzi. Meghatároztuk az első két formáns *y*-eltolásának különbségét minden cella esetén. Ezt úgy kaphatjuk meg, hogy a második formáns *y*-eltolás értékéhez hozzáadunk 6-ot, és ebből vonjuk ki az első formáns *y*-eltolás értékét. A művelet a 8. ábra összeállításának magyarázatából következik. Azt találtuk, hogy a neutrális F_1-mozgással rendelkező hanghármasok esetében az *y*-eltolás távolságok szignifikánsan nagyobbak, mint a változó osztályba soroltaké (kétmintás Wilcoxon-próba: $p < 0{,}001$).

8. táblázat: Az 5 legnagyobb gyakoriságú formánstérkép

Formánstérkép	Gyakoriság	$C_{kh.1}V_mC_{kh.2}$ hanghármas
1-SZ~3-SZCS alacsony~magas	4,35%	AL[e:]BL, AL[e:]DA, BL[i, i:]LD, LD[i, i:]DA, DA[i, i:]BL, DA[i, i:]LD, AL[i, i:]BL, AL[i, i:]LD, AL[i, i:]DA, PA[i, i:]BL PA[i, i:]LD, PA[i, i:]DA, PA[i, i:]AL, VE[i, i:]LD
3-SZCS~2-SZN alacsony~alacsony	2,48%	BL[o, o:]LD, VE[o, o:]PA, BL[u, u:]BL, BL[u, u:]AL, BL[u, u:]PA, LD[u, u:]PA, DA[u, u:]PA, VE[u, u:]AL
7-FU~1-SZ közepes~alacsony	2,17%	BL[ɔ]BL, BL[ɔ]VE, LD[ɔ]BL, LD[ɔ]VE, AL[ɔ]LD, VE[ɔ]LD, VE[ɔ]VE
3-SZCS~2-SZN alacsony~közepes	2,17%	BL[ø, ø:]AL, BL[ø, ø:]PA, LD[ø, ø:]AL, AL[ø, ø:]AL, AL[ø, ø:]PA, VE[ø, ø:]AL, DA[y, y:]PA
3-SZCS~4-NSZ alacsony~magas	2,17%	BL[e:]AL, BL[e:]PA, LD[e:]PA, DA[e:]PA, BL[i, i:]VE, DA[i, i:]AL, BL[y, y:]PA

A második formánsban neutrális formánsmenettel rendelkező $C_{kh.1}V_mC_{kh.2}$ hanghármasok esetében azt várjuk, hogy a környező mássalhangzók locusa (körülbelül) a magánhangzó F_2-jével egy magasságban van (Vértes 1982: 92). Neutrális mozgással mindössze a cellák 14%-a rendelkezik, a legtöbb az [o, o:], majd az [ɔ], [u, u:], [i, i:], [ø, ø:], [e:], [y, y:], [a:] és [ɛ] magánhangzóminőségek esetén fordul elő.

Magdics (1965) szerint a [p], [b], [m] locusainak az egyes magánhangzókra gyakorolt hatása nagyon hasonló, így azt várjuk, hogy a BLV_mBL cella stabilitása nagy lesz. Adatbázisunkban az SM stabilitási mutató átlaga ebben

a cellatípusban 0,333, amely a negyedik legjobb a lehetséges 36 kontextusból. A három legnagyobb stabilitású mássalhangzó-környezet F_2 formánsmozgásai alapján: LDV_mLD (0,308), BLV_mVE (0,309) és LDV_mBL (0,313). Megvizsgálhatjuk, hogy a szimmetrikus $C_{kh}V_mC_{kh}$ hanghármasok milyen sorrendben következnek egymás után az adatbázisban mért stabilitási mutató alapján. Méréseink szerint a szimmetrikus hangkörnyezetek sorrendje a hasonlóbb formánsmozgásokat tartalmazótól a kevésbé hasonlóig a következő: LD, BL, VE, AL, DA, PA. Ez a sorrend magyarázható Magdics (1965) magyar mássalhangzólocusokra vonatkozó megállapításaival. Mérései alapján a palatálisok [ɲ], [c, ɟ] és [j] csoportokba sorolva, három határozottan különböző locusértékkel rendelkeznek, a dentialveolárisok pedig számosságuknál fogva egy nagyobb locustartományt fognak át.

A szimmetrikus $C_{kh}V_mC_{kh}$ hanghármasok segíthetnek annak a kérdésnek a megválaszolásában is, hogy a CV vagy VC helyzetben érvényesül-e jobban a hangkörnyezet hatása. A nagyobb hatás nagyobb formánsmozgás-elhajlással jár, tehát ha például CV-ben várunk nagyobb kontextushatást, akkor gyakoribbak a 4-es és 5-ös típusú, CV-ben jobban elhajló formánsmenetek, mint a 2-es és 3-as kategóriájúak, amelyeknél kisebb az elhajlás a követő mássalhangzó locusa felé. A lehetséges $6 \times 9 = 54$ szimmetrikus kontextusból 13 db 1-szinttartó, 12 db 6-U alak, 11 db 7-fordított U alak formánsmenet található az adatbázisban, melyek mindegyike szimmetrikus mozgással írja le a magánhangzó második formánsának változását. A nem szimmetrikus formánsmenetek közül 6 db 2-szinttartó-növekvő és 5 db 3-szinttartó-csökkenő, illetve 1 db 4-növekvő-szinntartó és 4 db 5-csökkenő-szinttartó, amely inkább a VC kapcsolatok erőteljesebb hatását mutatja. A Vértes (1982) által vizsgált *agyagya* logatom második magánhangzója a mi rendszerezésünkben a PA[ɔ]PA cellához tartozik. A 7. ábráról leolvasható, hogy ehhez a cellához F_2-ben a 6-os, U-alakú formánsmenettípus tartozik.

Következtések

A 9 magyar magánhangzó-minőség formánsmozgásainak tipizálását végeztük el $C_{kh.1}V_mC_{kh.2}$ hanghármasokban, ahol C_{kh} a mássalhangzó képzési helyét jelenti (6 képzési hely). A tipizálást formánsadatbázis adataiból végeztük, magánhangzónként 5 mérési pont alapján. A kidolgozott eljárással a normalizált formánstérbe konvertált formánsmozgásokat 12 formánsmenettípusba soroltuk. A 324-féle $C_{kh.1}V_mC_{kh.2}$ hanghármas egyes elemeihez hozzárendeltük a jellemző formánstérképeket, amelyek cellánként 4 adatértéket jelentenek: az F_1 és F_2 formánsmenettípusok és azok *y*-eltolás értékei. A kategorizálás részben sikeresnek mondható, hiszen a 12. egyéb (hibakategória) 7 és 12%-ban jelent meg az első és második formáns esetén, de a cellák formánsmenettípusainak meghatározásánál már csak F_2-ben fordult elő 9%-kal. Az eljárás során definiált stabilitási mutatók (SM és SM_2) segítségével igazoltuk, hogy az azonos $C_{kh.1}V_mC_{kh.2}$ cellákhoz azonos jellegű formáns-

mozgások tartoznak. A formánsmozgások a második formánsban nagyobb stabilitást mutatnak az első formáns mozgásaihoz képest. A cellákhoz rendelt formánsmenettípusok szigorú stabilitását az SM_2 mutatóval mértük, amelynek 0,49 és 0,54 értékei (első és második formáns) nem tekinthetők túl magasnak. Azonban ha a cella formánsmenetéhez a második leggyakoribb típust is hozzávesszük (feltéve, hogy köztük 1 a távolság), akkor a mutató értéke 0,75-re és 0,79-re nő.

A cellához rendelt formánsmenettípusok áttekintése után azt mondhatjuk, hogy a második formánsban 8-9 formánsmenettípussal tudjuk jellemezni a lehetséges formánsmozgásokat, míg F_1-ben a szinttartó, a szinttartó-csökkenő és a fordított U alak a cellák 83%-át lefedi. A formánstérképek együtt írják le a két formáns menetét és az *y*-eltolást. A különböző formánstérképek száma jóval kisebb, mint a $C_{kh.1}V_mC_{kh.2}$ kapcsolatokból adódó hanghármas típusok száma, így adattömörítésre nyílik lehetőség.

Az adatbázis jelentős bővítése folyamatban van, amelynek eredményeképpen a kevésbé gyakori $C_{kh.1}V_mC_{kh.2}$ hanghármasok számát úgy tervezzük megnövelni, hogy minden esetben lehetővé váljon a tendenciák kiolvasása az adatokból. A formánsmozgások rendszerezése további formánsvizsgálatok kiindulópontjául szolgálhat, például a koartikulációs hatások pontosítása terén. A formánstérképek jó alapot adnak formánsérték-javító algoritmusok kidolgozásához, így megnyílhat az út nagyméretű és sokbeszélős beszédadatbázisok formánsterének adatbázisba rendezéséhez is.

Irodalom

Abari Kálmán – Olaszy Gábor 2011. Magyar formánsadatbázis az interneten. *Beszédkutatás 2011.* 73–82. (vö. http://magyarbeszed.tmit.bme.hu/formans).

van Bergem, Dick R 1993. Acoustic vowel reduction as a function of sentence accent, word stress, and word class. *Speech Communication* 12. 1–23.

Bolla Kálmán 1978. A magyar magánhangzók akusztikai analízise és szintézise. *Magyar Fonetikai Füzetek* 1. 53–67.

Broad, David – Clermont, Frantz 2010. Target-locus scaling methods for modeling families of formant transitions. *Journal of Phonetics* 38/3. 337–359.

Csapó Tamás Gábor – Németh Géza 2009. Mássalhangzó-magánhangzó kapcsolatok automatikus osztályozása szubglottális rezonanciák alapján. In Tanács Attila – Szauter Dóra – Vincze Veronika (szerk.): *Magyar Számítógépes Nyelvészeti Konferencia.* Szeged, 226–237.

Flynn, Nicholas – Foulkes, Paul 2011. Comparing vowel formant normalization methods. In: *Proceedings of the 17th International Congress of Phonetic Sciences ICPhS 2011.* Hong Kong, 683–686.

Gósy Mária 1995. Szükséges és szükségtelen hangátmenetek. *Beszédkutatás 1995.* 20–32.

Gósy Mária 2004. *Fonetika, a beszéd tudománya.* Osiris Kiadó, Budapest.

Gráczi Tekla Etelka – Horváth Viktória 2010. A magánhangzók realizációja spontán beszédben. *Beszédkutatás 2010.* 5–16.

Harrington, Jonathan 2010. *The phonetic analysis of speech corpora.* Wiley-Blackwell. http://www.phonetik.uni-muenchen.de/~jmh/research/pasc010808/pasc.htm (A letöltés ideje: 2011. szeptember)

Krull, Diana 1987. Second formant locus patterns as a measure of consonant-vowel coarticulation. *Phonetic Experimental Research at the Institute of Linguistics University of Stockholm Sweden* V. 43–61.

Lindblom, Björn 1963. Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* 35/11. 1773.

Lobanov, Boris 1971. Classication of Russian vowels spoken by different speakers. *Journal of the Acoustical Society of America* 49. 606–608.

Olaszy Gábor 1985. *A magyar beszéd leggyakoribb hangsorépítő elemeinek szerkezete és szintézise.* Nyelvtudományi Értekezések 121. Akadémiai Kiadó, Budapest.

Olaszy Gábor 2003. Az artikuláció akusztikai vetülete – a hangsebészet elmélete és gyakorlata. In Hunyadi László (szerk.): Kísérleti fonetika, laboratóriumi fonológia a gyakorlatban. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen, 241–254.

Olaszy Gábor – Rácz Zsuzsanna – Bartalis Mátyás 2009. Formánsmérések automatizálása, formánsadatbázisok létrehozása. *Beszédkutatás 2009.* 134–147.

R Development Core Team 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/ (A letöltés ideje: 2011. szeptember)

Stevens, Kenneth – House, Arthur – Paul, Allan 1966. Acoustical description of syllabic nuclei: an interpretation in terms of a dynamic model of articulation, *Journal of the Acoustical Society of America* 40. 123–132.

Tarnóczy Tamás 1941. *A magyar magánhangzók akusztikai szerkezete.* Királyi Magyar Pázmány Péter Tudományegyetem Általános Nyelvészeti és Fonetikai Intézete, Budapest.

Vértes O. András 1982. A magyar beszédhangok akusztikai elemzésének kérdései. In Bolla Kálmán (szerk.): *Fejezetek a magyar leíró hangtanból.* Akadémiai Kiadó, Budapest, 71–115.

Watson, Catherine – Harrington, Jonathan 1999. Acoustic evidence for dynamic formant trajectories in Australian English vowels. *Journal of the Acoustical Society of America* 106. 458–468.

A kutatást támogatta a CESAR project (Grant No 271022).

SUTTOGOTT BESZÉDHANGOK AUTOMATIKUS OSZTÁLYOZÁSA FELOLVASOTT BESZÉDBEN

Tóth Andrea – Beke András

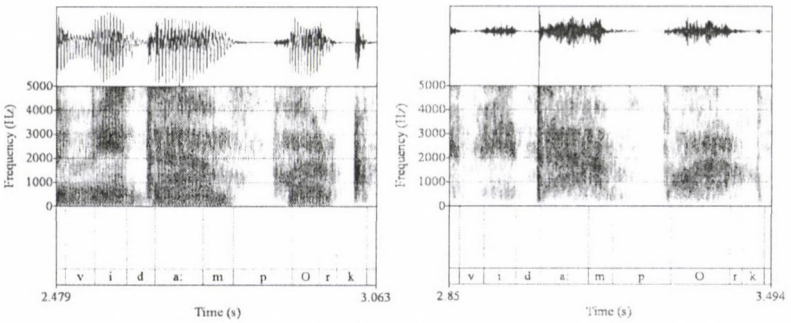
Bevezetés

A mindennapi kommunikációban számtalanszor adódhat olyan beszédhelyzet, amelynek során suttognunk kell (pl. könyvtári beszélgetés, titkolózás, segélykérés telefonon). A suttogás a beszédnek egy természetes formája, amelyet a beszédképzés módjában megjelenő változások tesznek különlegessé. Köznapi tapasztalataink azt mutatják, hogy a suttogott beszédet képesek vagyunk – szituációtól függően – megérteni. Mindez azért lehetséges, mert a suttogáskor a beszédhangokban megőrződnek olyan akusztikai jegyek, amelyek segítségével azok felismerhetők. Természetesen a felismerésben nagy szerepe van a kontextusnak is.

A suttogott beszéd sajátosságai a beszédképzés módjából erednek. Fiziológiailag az emberi hang keletkezésének alapja a kilégzés okozta levegőáramlás, a hangszalagok fonációs beállítása és rezgése és a toldalékcső rezonátortevékenysége (Stevens 2000; Reetz–Jongman 2009). A hangszalagok és a toldalékcső működésének folyamatait a központi idegrendszer, illetve a perifériás idegrendszeri és reflexes tevékenység koordinálja (Gósy 2002). A hangképzés alapja a fonáció, azaz a zöngeképzés, aminek a legelfogadottabb mai modellje a mioelasztikus-aerodinámiás modell. Lényege, hogy a hangszalagok rezgésének folyamata öngerjesztett rezgés, amelynek során a rezgő hangszalagok szabályozzák az energiát (levegőáramlást), és ez biztosítja rezgésüket (Stevens 2000). Normál ejtésben zöngétlen mássalhangzók képzésekor a hangszalagok szűk lélegző állásban, magánhangzók és zöngés mássalhangzók képzésekor zöngeállásban vannak. Suttogáskor ún. suttogóállásban helyezkednek el.

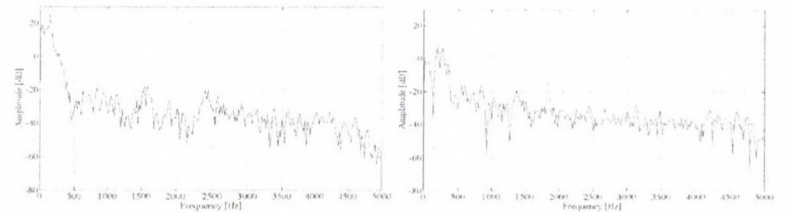
Suttogóállásban a kannaporcok nyitottak, a hangszalagok ugyanakkor majdnem teljesen zártak, ezért a levegő csak a kannaporcok közötti résen tud áthaladni, amit súrlódási zörej kísér. Az így létrehozott beszédet ún. „préselés" jellemzi. „Az álhangszalagok ilyenkor lefelé és befelé nyomódnak" (Gósy 2004: 32). A hangszalagok megváltozott feszülése miatt a hangrezgés aperiodikussá válik, ami a beszédet zörejessé teszi, illetve a hangerő is csökken (1. ábra). A suttogott beszédben a zöngés gerjesztés hiánya miatt megváltozik a zöngés hangok frekvenciaszerkezete, ugyanakkor a formánsoknak

megfeleltetett zörejgócok megmaradnak. A zöngétlen réshangok intenzitása jellemzően nagyobb, mint a magánhangzóké (Németh–Olaszy 2010: 166).



1. ábra
A *vidámpark* szó rezgésképe és hangszínképe normál (balra) és suttogott ejtésben (jobbra) (női beszélő)

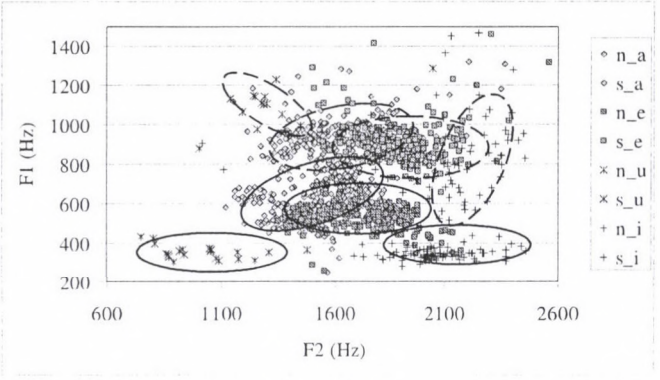
Suttogáskor a glottális szűkületen turbulens levegőáramlás keletkezik, amely a beszédhang képzéséhez alapvető gerjesztési forrásfajta. A turbulens levegőáram kiáramlásakor a szájüreg konfigurációjától függően változik a zörej akusztikai realizációja, amely létrehozza az adott pl. magánhangzóra vonatkozó tipikus akusztikai kulcsokat. Ezen kulcsok a magánhangzók esetében a formánsok. A suttogott magánhangzókkal foglalkozó kutatások többsége kiemelte, hogy suttogáskor hasonló spektruma van a beszédhangoknak, mint a normál ejtésűeknek. A legmarkánsabb különbség az intenzitásban van: a suttogott ejtéskor az intenzitás értéke jelentősen lecsökken. Különbség még az, hogy a spektrumban megjelenő csúcsok magasabb frekvencián jelennek meg (2. ábra).



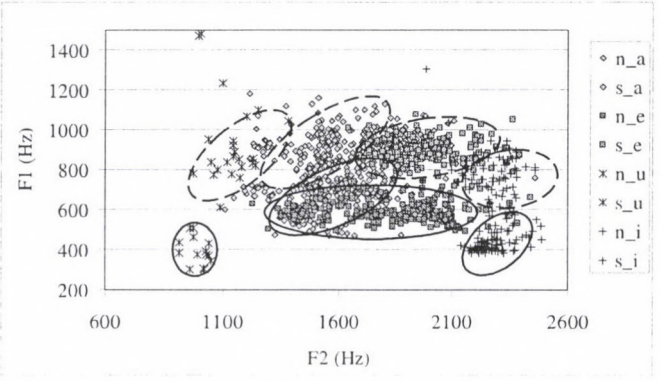
2. ábra
A normál (jobbra) és a suttogott (balra) [a:] magánhangzó spektruma

Továbbá, ellentétben a zöngével képzett magánhangzókkal, ahol az amplitúdó értéke a magasabb frekvencián megjelenő formánsnak általában alacso-

nyabb, mint az alacsonyabb frekvenciájú formánsnak, a második formáns a suttogott magánhangzóknál általában olyan intenzív, mint az első formáns. Kallail és Emanuel (1984) kutatásukban kimutatták, hogy a suttogott ejtésű magánhangzók első formánsa tipikusan magasabb frekvenciákon realizálódik, mint a normál ejtésűeké. Gósy (2002) kutatási eredményei szintén megerősítették, hogy a suttogott magánhangzók formánsai – jellegzetesen az F_1 – magasabb frekvencián jelennek meg. Ennek szemléltetésére a jelenlegi adatokból nyertünk ki formánsokat, és ábrázoltuk az F_2/F_1 térben [vö. 3. *a)* és *b)* ábra].



3. *a)* ábra
Férfi beszédben a normál (folytonos vonal) és suttogott (szaggatott vonal) ejtésű magánhangzók F_2/F_1 térben (a jelen kutatás adatai alapján)



3. *b)* ábra
Női beszédben a normál (folytonos vonal) és suttogott (szaggatott vonal) ejtésű magánhangzók F_2/F_1 térben (a jelen kutatás adatai alapján)

A zöngétlen réshangok és a zöngétlen explozívák esetében Gósy (2002) nem talált eltérést az akusztikai szerkezetben, az approximánsok és a nazálisok pedig a magánhangzókra jellemző eltéréseket mutatták.

A suttogott beszéd jellegzetességei (a beszéd zörejessége, az intenzitáskülönbség, a magánhangzók formánsainak változása, a zönge hiánya) észlelési nehézségeket okozhatnak. Gráczi (2005, 2006) kutatásában a suttogott mássalhangzók és magánhangzók észlelését vizsgálta. A mássalhangzók esetében a dekódolási nehézségek ellenére a véletlennél biztosabb volt a felismerés (a szavaknál átlagosan 74%, a logatomoknál 66%), ezért Gráczi az obstruensek esetében feltételezett „olyan invariáns akusztikai paraméter(eke)t, mely(ek) a fonéma zöngésségének függvényében jellemzi(k) a hangot" (2005: 89). A suttogott ejtés okozta módosulás a magánhangzók azonosításában nem okozott nehézséget. Tartter (1991) kutatásában logatomokon tesztelte a normál és a suttogott ejtésű magánhangzók észlelési arányát. Az eredményei azt mutatták, hogy a suttogott magánhangzókat csupán 10%-kal észlelték nagyobb pontossággal a normál ejtésűekhez képest.

Velde és van Heuven (2011) suttogott ejtésű (szókezdő és szó belseji pozíciójú), zöngésség szerint csoportosított zár- és réshangpároknál vizsgálták a felismerhetőséget biztosító akusztikai kulcsokat. Eredményeik szerint a zöngés-zöngétlen hangok elkülönítésben a másodlagos akusztikai kulcsok, vagyis a zárszakasz időtartama, a zárfelpattanás intenzitása és időtartama, illetve a zöngekezdési idő játszanak szerepet.

A intonációs szerkezetek akusztikai realizációját vizsgálta Heeren és van Heuven (2011) a suttogott ejtésben, kijelentő és kérdő mondatok felhasználásával, négy utolsó szótagi helyzetű magánhangzó segítségével. A magánhangzók vizsgált akusztikai jellemzői közül a spektrális energia és az intenzitás nem, a második formáns (F_2) mozgása és az intonációs szerkezet között ugyanakkor összefüggést találtak a suttogott ejtésben. Ez a jellemző a suttogáskor kompenzációs kulcsként működhet. Magyar mondatok esetében ugyanakkor Fónagy (1968) és Markó (2008) nem talált a formánsmozgások és az intonáció észlelési eredményessége között egyértelmű kapcsolatot: bizonyos esetekben a formánsmozgások másképp alakultak a kérdésben, mint a kijelentésben, rendszerszerű eltérések azonban igazolódtak.

Szerb kutatók (Jovičić–Šarić 2006) az idő és az intenzitás függvényében vizsgáltak suttogott mássalhangzókat. Kísérletükben igazolták, hogy a suttogott beszéd intenzitásában elmarad a normáltól, a zönge megszűnik, a prozódiai tulajdonságok ugyanakkor időben jól megőrződnek. A mássalhangzók időtartamát meghatározta a mondatban elfoglalt helyzetük, a suttogott mássalhangzók időtartama meghosszabbodott a normál ejtésűekéhez képest (a zöngéseké 15%-kal). A legnagyobb különbség a palatális helyen képzett hangokon volt megfigyelhető. A kutatók szerint a suttogás magasabb fokú precizitást igényel a hangképző szervek (gége) működésében, ezért igaznak tűnhet

az az elmélet, amely szerint a suttogás káros hatású a hangszalagokra nézve (vö. Rubin et al. 2006).

A suttogott beszéd felismerése nemcsak a humán percepció szempontjából lehet érdekes, hanem az automatikus gépi beszédfelismerés szempontjából is. A gépi felismerő rendszerek alapvető célja az ember és a technikai eszközök, gépek (pl. mobiltelefon, számítógép) közötti párbeszéd létrehozása. Egyes beszédfelismerő rendszerekkel szemben támasztott egyik elvárás lehet, hogy különböző beszédhelyzetekben és ejtésmódtól (pl. normál ejtésű vagy suttogott beszéd) függetlenül használhatóak legyenek. A lehető legpontosabb felismerés érdekében a felismerők létrehozásakor számos tényezőt figyelembe kell venni. Ezek közé tartozhat a beszédmódfüggetlen modellek létrehozása, amelyek nemcsak a normál ejtésű, de a suttogott beszédet is képesek azonosítani.

Mint láttuk, a suttogott beszéd akusztikai paraméterei eltérnek a normál beszédtől, ezért az automatikus felismerés romlik, ha a normál beszéden betanított modellt alkalmazzák. Ezt bizonyították Itoh és munkatársai (2005). A suttogott beszéd (kísérletükben felolvasott mondatokat használtak, az eredmények fonémákra vonatkoznak) automatikus felismerése normál rendszeren 52%-os, suttogott beszédre kidolgozott modellen 76%-os eredményt hozott. A megalkotott HMM-modell (Hidden Markov-model, azaz rejtett Markov-modell) a legtöbb hangot a [h]-val tévesztette össze. A magánhangzók között a legjobb eredményt az [ɔ], a legrosszabbat az [o] esetében adta. A modell a zárhangokat és réshangokat is felcserélte néhány esetben, főleg a képzési hely szerint. A téves azonosításokat a zönge hiányával és az akusztikai szerkezet megváltozásával hozták összefüggésbe a szerzők. A kutatásban a folyamatos beszédfelismerés vizsgálatát is elvégezték, suttogott, normál és beszédmódfüggetlen (azaz suttogott és normál együttesen) HMM-modelleket használva. Eredményeik azt mutatták, hogy a beszédmódfüggetlen modell mindkét ejtéstípus felismerésére alkalmas (Itoh et al. 2005).

További kutatások bizonyították, hogy más akusztikai jellemzőkinyerési technikákkal növelni lehet a suttogott beszéd felismerési eredményeit. Carlin és munkatársai (2006) normál és suttogott beszédrészek különválasztására osztályozót hoztak létre nem ellenőrzött tanuló algorimussal. Kísérletükben az energiaaránnyal, módosított lineáris predikciós eljárással mint akusztikai jellemzőkkel klasztererezték a normál és a suttogott beszédrészeket 97,5%-os eredménnyel. Mathur és Hegde (2011) egy új LP-n (lineáris predikciós eljárás) alapuló spektrális arányt számoló eljárást alkalmazott, mint akusztikai jellemzőt a suttogott és normál beszédrészek osztályozásához. Az eredmények azt mutatták, hogy az új jellemzőkinyerési technikával csökkenteni tudták a felismerési hiba arányát.

Kifejezetten a suttogott beszéd beszédhangjainak automatikus osztályozására magyar nyelven még nem voltak próbálkozások. Kutatásunk ezt a hiányt igyekszik pótolni. Célunk, hogy MFC-kel előfeldolgozva (mel-frequency cepstrum) rejtett Markov-modellekkel automatikusan osztályozzuk a sutto-

gott beszéd beszédhangjait. Feltételeztük, hogy a suttogott beszéden betanított modellek segítségével jobb osztályozási eredményeket érhetünk el suttogott beszéden, mint normálbeszéd-alapú modellek használatával. A kutatásban megvizsgáltuk azon beszédhangokat, amelyeket a gépi osztályozó nem vagy csak rossz eredménnyel tudott osztályozni.

Anyag és módszer

A suttogottbeszéd-osztályozó tanításához és teszteléséhez korpuszt építettünk. Három nő és egy férfi (átlagéletkoruk 27 év) suttogott beszédét vettük fel csendesített szobában. Az adatközlőknek két mondatot kellett felolvasniuk 40-szer (összesen 80 normál és 80 suttogott ejtésű mondat): *Országszerte nagy érdeklődéssel figyelték az új vidámpark megnyitását. Különösen a gyerekek várták, hogy kipróbálhassák a dodzsemet és a céllövöldét.* A beszédüket 44,1 kHz-en mintavételeztük 16 biten, majd újramintavételeztük 16 kHz-en.

A mondathatárokat automatikusan bejelöltük a beszédben a Praat (Boersma–Weenink 2010) programban írt scripttel, majd a MAUS szoftverrel (Schiel–Beringer 1995) automatikusan annotáltuk őket hangszinten. Az így kapott hanghatárokat manuálisan javítottuk.

A jelen kutatásban egy beszédhangmodell-alapú felismerőt készítettünk. A felismerésre kerülő egységeket (beszédhangok) MFC-kkel előfeldolgozva három állapotú HMM-ekkel modelleztük. A rendszert különbözőképpen tanítottuk és teszteltük (1. táblázat).

1. táblázat: A rendszer tanítása és tesztelése

	Tanítás	Tesztelés	Rövidítés
1.	normál	normál	NN
2.	normál	suttogott	NS
3.	suttogott	suttogott	SuS
4.	normál és suttogott	normál és suttogott	BF

A HMM-ek tanításához a teljes adathalmaz használatakor (BF esetében) 120 mondatot, a teszteléshez 40-et használtunk. Az NN, SuS, NS eseteiben a tanításhoz 60, a teszteléshez 20 mondatot használtunk.

Mind a négy modellt 2, 4, 8, 16 gaussos kibocsátási valószínűséget leíró függvénnyel tanítottuk. A nyelvtanban minden hangmodellt egyenlő súllyal – azaz egyenlő valószínűséggel – tanítottunk.

Az eredmények kiértékelése érdekében a következő jellemzőket határoztuk meg (vö. Mihajlik 2009): a felismerés aránya (Corr.), a felismerés pontossága (Acc.), a felismerési hiba (ER.) és a relatív különbség (RI. improvement). Ezek kiszámításához a rendszer találati és tévesztési értékeit használtuk fel: a korrekten felismert modellek száma (C), a helyettesítettek száma (S), a töröltek száma (D) és a beszúrások száma (I). (N) az összes felismerési egység

száma. A kiértékelésénél az ún. felismerési arány és felismerési pontosság az egyik meghatározó. Kiszámítási módjuk a következő:

Felismerési arány:

$$"Corr" = \frac{N - S - D}{N} * 100\% ,$$

Felismerési pontosság:

$$"Acc" = \frac{N - S - D - I}{N} * 100\% ,$$

Felismerési hiba:

$$"ER" = 100\% - "Acc" .$$

Több felismerési eredmény összehasonlítására a relatív különbség mértéke adhat választ, mely a következőképp határozható meg:

$$\text{Relatív különbség} = \frac{ER_{referencia} - ER_{új}}{ER_{referencia}} * 100\% .$$

Mind a négy HMM-modell esetében meghatároztuk a felismerés arányát, pontosságát, a felismerési hiba értékét. A modelleket összehasonlítottuk, meghatároztuk a relatív különbség értékét.

A beszédhangok elemzését a Praat szoftver (spektrogram- és rezgéskép-elemzések), a statisztikai elemzést az SPSS 13. szoftver (ANOVA 95%-os konfidenciaszinten) segítségével végeztük el.

A rejtett Markov-modell

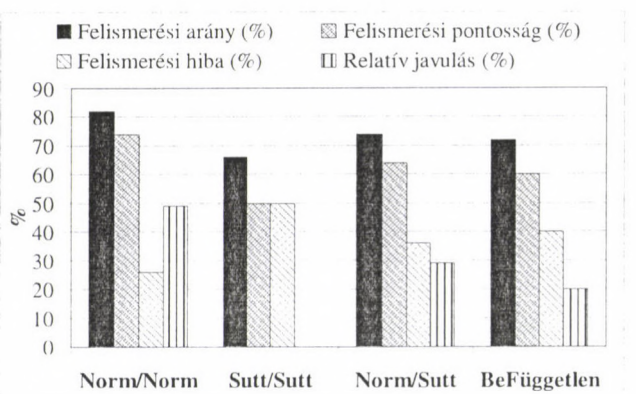
Az automatikus beszédfelismerésben számos algoritmus létezik: ANN (artificial neutral network), SVM (support vector machine), GMM (gauss mixture machine). Az egyik legelterjedtebb ezek közül a rejtett Markov-modell (HMM).

A rejtett Markov-modellezéskor az akusztikai előfeldolgozás tekintetében a beszédfelismerésben a beszédhangok akusztikai modellezéséhez használt előfeldolgozási láncot meghagyjuk, és a gyakran alkalmazott MFC (mel frekvenciás kepsztrális) együtthatókat számítjuk 39 elemű jellemzővektorokat létrehozva. Ezek a jellemzővektorok delta és delta-delta együtthatókat is tartalmaznak. A modellkomplexitást legfeljebb 16 komponenst tartalmazó Gauss-keverék (GMM) sűrűségfüggvényig növeljük. Az alkalmazott modellek topológiájuk tekintetében minden esetben 3 állapotú balról jobbra modellek. A Markov-modellek a beszédfelismerőben valójában kettős feladatot látnak el, a jellemzővektorok osztályozása mellett a beszédjelet illesztik is a neki megfelelő beszédhangsorozatra, azaz meghatározzák az egyes beszédhangok kezdő- és végidőpontjait. A Markov-modellek úgy is használhatók,

hogy előre elkülönített osztályokra betanított modellek alapján osztályozzanak. Ilyen esetben a beszédfelismerőben használatos szótár szerepét az egyes osztályok listája, a nyelvtan szerepét pedig az osztályozás szabályai veszik át. A beszédhangok osztályozása esetén a lista az osztályozni kívánt beszédhangokból áll, az osztályozás szabályai pedig megadják, milyen beszédhangokat milyen sorrendben lehet illeszteni.

Eredmények

A négy megalkotott HMM-modell közül a legrosszabb osztályozási arányú (66,13%) a SuS volt (4. ábra). A modellben az osztályozási pontosság mindössze 49,63%, közel azonos a hiba értékével. A négy modell közül ez adta a legrosszabb eredményt, ezért a relatív különbség értékét SuS-hez hasonlítva adjuk meg. A suttogott beszéd modellezésére kialakított három HMM közül a NS modell érte el a legjobb eredményt, 28,65%-os a javulás mértéke SuS-hez képest. Az osztályozási arány értéke 73,94%, az osztályozási pontosságé 64,06% volt, vagyis a felolvasott suttogott beszédet ez a modell 64%-ban helyesen ismerte fel.



4. ábra
A négy modell osztályozási eredményei

BF modell esetében a suttogott és a normál ejtésű adatokon együttesen végeztük el a betanítást, vagyis egy adott beszédhang modellje annak suttogott és normál ejtésű akusztikai paramétereit is tartalmazza. Kérdéses, hogy ha egy beszédmodell ilyen jellemzőkkel van ellátva, mennyire lesz sikeres a felismerés a hangok modelljei közti esetleges hasonlóság miatt (l. suttogott magánhangzók közötti átfedések). Feltételeztük, hogy a két ejtéstípus együttes használata javítja a felismerést, mert az így kialakított beszédhangmodellek jobban lefedik az egyes hangok lehetséges realizációit. A modell osztályozási

aránya 72,31% volt, közel azonos NS-sel. Az osztályozási pontosság azonban 4%-ot romlott, emiatt az osztályozási hiba értéke is nőtt, 40,25%-ra. BF eredményéből arra következtethetünk, hogy a modell több beszédhangot törölt a felismerésben, mint NS. Ugyanakkor a relatív különbség BF esetében is jelentős, 20,1%.

A legjobb osztályozási arányt a normál beszéden betanított és tesztelt modell (NN) adta, osztályozási aránya 82,15%, pontossága 74,25%. Ennek megfelelően NN-ben a legmagasabb a relatív különbség értéke is, közel 49%.

A négy megalkotott modell – SuS kivételével – a beszédhangokat átlagosan 66%-os pontossággal tudta felismerni. SuS eredményei ettől átlagosan 10%-kal maradtak el.

A beszédhangokat képzési jegyeik alapján csoportokba soroltuk, és vizsgáltuk átlagos osztályozási arányukat a négy modellben. A magánhangzók esetében a nyelv vízszintes mozgásától függően (palatális-veláris elkülönítés) nem találtunk jelentős különbséget (2. táblázat). A hangok felismerése nyelvállásfoktól függetlenül három modellnél 90% fölötti volt. SuS esetében az osztályozási arány átlagosan 80%-os volt. A felső és középső nyelvállású magánhangzók felismerése sikeresebb volt, kivéve a SuS modellt. Az ajakműködés szerint osztályozva a hangokat a labiális magánhangzók felismerése volt biztosabb (kivétel SuS modell).

2. táblázat: A magánhangzók átlagos osztályozási aránya modellek szerint

		SuS (%)	NS (%)	BF (%)	NN (%)
Palatális	**Felső**	81,5	92,2	93,8	99,3
	Középső	83,9	89,3	87,5	93,5
	Alsó	93,8	78,9	94,8	97,5
Veláris	**Felső**	76,2	100,0	100,0	100,0
	Középső	63,9	97,4	89,9	96,8
	Alsó	75,8	83,9	84,8	84,0
	Legalsó	88,0	73,0	69,8	79,7

A mássalhangzók esetében a zöngés zárhangok átlagos felismerése pontosabb volt zöngétlen társaikénál, a legpontosabb eredményt NS és NN adta. A zöngétlen zárhangokat ismerte fel minden modell – NN kivételével – a legkevésbé. A nazálisok felismerése modelltől függetlenül magas, 90% fölötti volt. A réshangok esetében a zöngétlen hangok felismerése bizonyult könnyebbnek zöngés párjukkal szemben, ugyanakkor az osztályozási arányok mindkét esetben 85–100%-osak voltak. A tremuláns felismerése közel azonos mértékben sikerült minden modellnek, az osztályozási arányok 80% körüliek voltak. A két közelítőhang eredményei már nagyobb eltérést mutattak: [j] esetében a felismerés aránya viszonylag magas volt, BF-ben a legalacso-

nyabb, 87,5%. A [l] felismerése BF-ben kevésbé volt sikeres (55,5%), ugyanakkor SuS és NN modellekben 80-90%-os osztályozási arányú volt. A mássalhangzócsoportok átlagos osztályozási arányait a 3. táblázat mutatja.

3. táblázat: A mássalhangzók átlagos osztályozási aránya modellek szerint

		SuS (%)	NS (%)	BF (%)	NN (%)
Zárhang	**Zöngés**	74,3	90,2	83,7	90,4
	Zöngétlen	69,0	79,0	78,9	94,9
Nazális		94,1	97,2	94,2	98,1
Réshang	**Zöngés**	85,3	86,7	91,2	98,0
	Zöngétlen	92,9	96,7	97,9	95,9
[r]		84,5	81,7	80,4	82,7
[j]		95,2	95,0	87,5	100,0
[l]		82,2	63,7	55,5	89,7

A magánhangzó-tévesztések NN, NS és BF modellekben leggyakrabban az [a:] (75%), [e:] (79%) és [ɔ] (84%) hangokat érintették. A három hang egymással való felcserélésének oka a nem tiszta artikuláció lehet, ami az akusztikai szerkezet torzulásához vezet(het). A spontán beszédben megfigyelhető a magánhangzók semlegesedése, ekkor a hang a normál realizációtól előtérően jelenik meg, és F_2 értékében szignifikáns eltérést mutat (Beke–Gráczi 2010). Bóna (2004) igazolta, hogy a téves hangpercepció legnagyobb mértékben az [ɔ], [a:], [e:] hangokat érinti.

A SuS modellben a leggyakrabban tévesztett magánhangzók az [o:] (52,4%), [ɔ] (75,8%), [ø] (75,8%) és [i] (72,1%) voltak. Ezeket a modell a leggyakrabban más magánhangzókkal, nazálisokkal ([m], [n]), közelítőhangokkal [l] és zárhangokkal ([t], [k], [g]) helyettesítette. A magánhangzókkal való helyettesítéskor jellemző volt a felső nyelvállásfokú hangok alsóbb nyelvállásfokúakkal való helyettesítése, hasonlóan a humán percepció tévesztéseihez (vö. Gráczi 2006), vagyis az F_1 emelkedése zavart okozhatott a felismerésben. Bár a nyelvállás foka szerint nem mutatható ki szignifikáns eltérés, a 80%-os felismerés alatti magánhangzók jellemzően felső és középső nyelvállásúak (kivétel az [ɔ]). A nyelv vízszintes mozgása szerint a veláris hangok téves azonosítása volt a jellemzőbb, az ajakműködés szempontjából pedig a labiális hangok felismerése okozott problémát, hasonlóan a humán percepciós eredményekhez (l. Gráczi 2006). A nazálisok és közelítőhangok suttogott ejtésben a magánhangzókra jellemző eltéréseket mutatták, ez magyarázhatja helyettesítésüket. Az explozívák esetében a téves azonosítást a zörejessé vált zárszakasz és a zárfelpattanás intenzitásának csökkenése okozhatta.

A mássalhangzók közül a téves azonosítások a leggyakrabban a zárhangokat érintették. A magánhangzókkal való helyettesítésekről a magánhangzók

kapcsán már ejtettünk szót. A zárhangokra való felcserélést a hasonló akusztikai szerkezet magyarázhatja. A tremuláns felismerése szintén problematikus volt, ennek oka az [r] hang változatos akusztikai szerkezete (Gósy 2008). A leggyakrabban helyettesített hangok osztályozási arányait és helyettesítő hangjait modellenkénti bontásban a 4. táblázat mutatja.

4. táblázat: A leggyakrabban helyettesített beszédhangok

		[d]	[t]	[g]	[k]	[r]
SuS	**Osztályozási arány**	62,7%	54,8%	79,1%	71,6%	
	Helyettesítő hang	[n], [t]	[r], [e:], [ɔ], [a:]	[d], [v]	[e:], [o]	
NS	**Osztályozási arány**	79,4%	81,7%		54,2%	80,0%
	Helyettesítő hang	[ɔ], [n]	[l]		[g], [i]	[ʃ], [t]
BF	**Osztályozási arány**	79,2%	75,8%	74,8%	64,8%	
	Helyettesítő hang	[n], [ɔ]	[s], [p]	[ɛ]	[n]	

A beszédképzés módjának változása a hangok spektrális jellemzőinek módosulásához vezetett: a magánhangzók formánsai emelkedtek [ANOVA F_1: $F(1, 7734) = 8914{,}400$; $p < 0{,}001$; F_2: $F(1, 7734) = 343{,}235$; $p < 0{,}001$], a közelítőhangok és nazálisok a magánhangzókra jellemző eltéréseket mutatták, az explozíváknál a zárfelpattanás kevésbé volt intenzív, a hang zörejessé vált. A spektrális jellemzők változása miatt a suttogott ejtésű hangok akusztikai megvalósulásai kevésbé voltak homogének, ezért a beszédhangok modellezése nehezebbé vált (pl. a felismerő a tanítás során a zörejességnek nagy jelentőséget tulajdonít[hat]ott a hang szerkezetében). A SuS modellben ezért nőtt a tévesztések száma.

Következtetések

A jelen kutatásban a suttogott beszéd beszédhangjainak automatikus osztályozását elvégző rendszer kiépítésére vállalkoztunk. A magyar nyelven elért eredményeink közel azonosak, mint az Itoh és munkatársai (2005) által megalkotott rendszeré. A jelen feladat fontos lehet egy beszédmódfüggetlen felismerő kiépítésében. Az akusztikus modelleket különböző beszédmódokban tanítottuk: normál, suttogott beszédmód, beszédmódfüggetlen.

A suttogott beszéd beszédhangjai MFC-vel előfeldolgozva HMM-ekkel jól modellezhetők felolvasásban. A suttogott beszéd hangjait a normál beszéden tanított modellel és a beszédmódfüggetlen modellel lehetett a legjobban felismertetni.

A normál beszédből kinyert jellemzővektorok alkalmasabbak voltak a suttogott beszédhangok modellezésére. Ebből arra következtethetünk, hogy a beszédhangok spektrális jellemzői ejtésmódtól függetlenül részben megőrződhetnek, illetve csak olyan mértékben torzulnak, hogy azok még felismerhetők (l. alveoláris és palatális explozívák). A beszédmódfüggetlen modell

használata javította a felismerést, mert a kinyert jellemzővektorok jobban modellezték a hangokat, és lefedték azok lehetséges realizációit.

A jövőben tervezzük, hogy nagyobb, a hangkörnyezet változatos modellezését biztosító korpuszon végezzük el a kísérleteinket. Feltételezzük, hogy nagyobb beszédadatbázison az eredmények javíthatók: az adatközlők számának növelése lehet a korpusz megnövelésének dimenziója, mivel az egyes bemondások hosszát nem lehet növelni, hiszen csak egy bizonyos ideig vagyunk képesek suttogni.

Irodalom

Mathur, Arpit – Hegde, Rajesh M. 2011. Significance of the LP-MVDR spectral ratio method in whisper detection. *2011 National Conference on Communications (NCC)*. Bangalore, 126–130.

Beke András – Gárczi Tekla Etelka 2010. A magánhangzók semlegesedése a spontán beszédben. In Navracsics Judit (szerk.): *Nyelv, beszéd, írás. Pszicholingvisztikai tanulmányok I.* Tinta Könyvkiadó, Budapest, 57–65.

Boersma, Paul – Weenink, David 2010. *Praat: doing phonetics by computer.* (v 5.1.) http://www.fon.hum.uva.nl/praat/

Bóna Judit 2004. A beszédészlelési folyamat megakadásai. *Beszédkutatás 2004.* 223–231.

Carlin, Michael A. – Smolenski, Brett Y. – Wenndt, Stanley J. 2006. Unsupervised detection of whispering speech in he presence of normal phonation. Presentation. In: *Proceedings of INTERSPEECH 2006. Pittsburgh, PA, USA. September 17-21, 2006.*

Fónagy Iván 1968. Suttogott dallam? *Magyar Nyelvőr* 92. 253–262.

Gósy Mária 2002. Beszédképzés zönge nélkül. *Beszédkutatás 2002.* 18–37.

Gósy Mária 2004. *Fonetika, a beszéd tudománya.* Osiris Kiadó, Budapest.

Gósy Mária 2008. „R" hangok: kiejtés, hangzás, funkció. *Magyar Nyelvőr* 132. 1–17.

Grászi Tekla Etelka 2005. Suttogott mássalhangzók zöngésségének észlelése. *Beszédkutatás 2005.* 78–91.

Grászi Tekla Etelka 2006. Suttogott ejtésű magánhangzók észlelése. *Beszédkutatás 2006.* 86–100.

Heeren, Willemijn – van Heuven, Vincent J. 2011. Acoustics of whispered boundary tones: effects of vowel type and tonal crowding. In: *International Congress of Phonetic Sciences (ICPhS XVII).* Hong Kong, 851–854.

Itoh, Taisuke – Takeda, Kazuya – Itakura, Fumitada 2005. Analysis and recognition of whispered speech. *Speech Communication* 45. 139–152.

Jovičić, Slobodan T. – Šarić, Zoran 2006. Acoustic analysis of consonants in whispered speech. *Journal of Voice* 22/3. 263–274.

Kallail, Ken J. – Emanuel F. W. 1984. Formant-frequency differences between isolated whispered and phonated vowel samples produced by adult female subjects. *Journal of Speech, Language, and Hearing Research* 27/2. 245–251.

Markó Alexandra 2008. A magyar eldöntendő kérdő mondat megvalósulásai különböző beszédmódokban. *Beszédkutatás 2008.* 77–92.

Mihajlik Péter 2009. Spontán magyar nyelvű beszéd gépi felismerése nyelvspecifikus szabályok nélkül. Doktori értekezés. BME, Budapest.

Németh Géza – Olaszy Gábor 2010. *A magyar beszéd. Beszédkutatás, beszédtechnológia, beszédinformációs rendszerek.* Akadémiai Kiadó, Budapest.

Reetz, Henning – Jongman, Allard 2009. *Phonetics: Transcription, production, acoustics and perception.* Wiley–Blackwell, Oxford.

Rubin, Adam D. – Praneetvatakul, Veeraphol – Gherson, Shirley – Moyer, Cheryl A. – Sataloff, Robert T. 2006. Laryngeal hyperfunction during whispering: reality or myth. *Journal of Voice* 20/1. 121–127.

Schiel, Florian – Beringer, Nicole 1995. The Munich Automatic Segmentation System – MAUS. http://www.phonetik.uni-muenchen.de/forschung/Verbmobil/VM14.7eng.html

Tóth Bálint – Németh Géza 2008. Rejtett Markov-modell alkalmazása magyar nyelvű gépi szövegfelolvasáshoz. *Beszédkutatás 2008.* 182–194.

Tartter Vivien C. 1991. Identifiability of vowels and speakers from whispered syllables. *Attention, Perception, and Psychophysics* 49. 365–372.

van de Velde, Daan J. – van Heuven, Vincent J. 2011. Compensatory strategies for voicing initial and medial plosives and fricatives in whispered speech in Dutch. In: *International Congress of Phonetic Sciences (ICPhS XVII).* Hong Kong, 2058–2061.

SOROZATMEGAKADÁSOK MINTÁZATA A SPONTÁN BESZÉDBEN

Gósy Mária

Bevezetés

A spontán beszéd sajátosságaiból adódik, hogy diszharmóniás jelenségek fordulnak benne elő, illetve hogy a beszélő igyekszik javítani a hibás vagy hibásnak vélt beszédrészletet. Noha ilyenkor hibákról beszélünk, ezek a jelenségek a spontán beszéd természetes velejárói, koherens részei a beszédprodukciós mechanizmusnak (Levelt 1989; Shriberg 2001; Pouplier–Hardcastle 2005; Local 2007; Corley et al. 2007). Nemegyszer a felszíni közlés a hallgató szempontjából nem is tartalmaz hibát, a beszélő azonban – számos ok miatt – mégis korrigálja az elhangzottakat, módosít, változtat. Mindeközben a beszélő különféle stratégiákat használ, hogy az egyidejűleg zajló tervezési és kivitelezési folyamatokból esetlegesen adódó nehézségeit megoldja. Ha a spontán beszéd problémamentes, akkor hajlamosak vagyunk alábecsülni a beszédprodukciós mechanizmus komplexségét. A megakadások – amelyeket a pszicholingvisztika kis ékszereinek is neveznek (Perez et al. 2007) – egyben azt is jelzik, hogy nagyon bonyolult folyamatok mennek végbe a „háttérben", amelyek sikeres együttes működtetése az idő töredéke alatt kell, hogy megtörténjen.

A beszédtervezési nehézségek felszíni következményei több kutató szerint azért jönnek létre, mert a beszélő önmonitorozása, illetőleg a korrekció nem kielégítően működött (Blackmer–Mitton 1991; Postma–Kolk 1992; Kolk–Postma 1997). A megakadásjelenségek sok információt nyújtanak a beszédtervezési folyamatokról, az önmonitorozásról, a javítási műveletekről, a beszélő stratégiáiról. Osztályozásukat a közvetlen kontextus rendszerint probléma nélkül lehetővé teszi, azokat az okokat azonban, amelyek a létrejöttükért felelősek, nem lehet minden esetben pontosan megmondani (pl. Cutler 1988). A diszharmónia kialakulása a beszédprodukcióban jól definiált; azt jelenti, hogy a beszélő néha nem képes a csaknem vagy teljesen egyidejűleg zajló folyamatok hibátlan működtetésére. A beszédszándék bizonytalanságai, a tartalom kívánt kialakítása, avagy a rejtetten működő tervezési műveletek optimális megvalósításai ugyanakkor mégsem feleltethetők meg minden esetben a felszíni jelenségeknek. A tapasztalat és a célzott, konkrét vizsgálatok is azt igazolták, hogy a megakadások gyakoribbak a hosszabb és/vagy bonyolultabb közlésekben, illetve akkor, ha a közölni szándékozott gondola-

tok között a beszélőnek választási lehetősége van (Maclay–Osgood 1959; Levelt 1989; Clark–Wasow 1998; Poulisse 1999).

A beszédprodukciós folyamat működésére vonatkozó egyik jól ismert hipotézisben, Dell aktivációterjedéses produkciós elméletében négy olyan tényező van, amelyek működési hibát eredményezhetnek: a kivitelezési eltérés, a hasonlósági hatás, a beszédsebesség hatása és a távolsági hatás (1986: 292). A kivitelezési eltérés a mentális lexikon hibás aktiválásával kapcsolatos. A hasonlósági hatás lehet formai vagy szemantikai, és ezek következtében okozhat hibát. A gyors beszéd nyilvánvalóan több hibát eredményezhet, mint a lassabb, hiszen a lexikális egységek hosszabb ideig maradhatnak aktiváltak, illetve lassabban deaktiválódnak a gyors tervezés és kivitelezés miatt. A távolsági hatás abban jelentkezik, hogy a szomszédos vagy egymáshoz közel lévő nyelvi egységek nagyobb mértékben sérülékenyek, mint az egymástól távolabb lévők. Ha elfogadjuk az aktivációterjedés elméletét (amely évtizedek óta éppúgy kapott megerősítést, mint számos kritikát), akkor feltételezhetjük, hogy az aktiváció terjedése nem csupán a hibátlanul működő szintekre igaz, hanem valamiféle aktivációterjedés jellemezheti a diszharmóniás jelenségeket is. Ha pedig ez így van, akkor ennek a felszíni következményei egymást követő megakadásokat eredményezhetnek.

A hazai és a nemzetközi szakirodalom is jórészt egy vagy néhány jelenség vizsgálatával foglalkozott (Cutler 1988; Dell 1990; Mowrer–Mackay 1990; House 2003; Horváth 2004; Markó 2004; Huszár 2005; Nooteboom 2005; Curl et al. 2006; Markó 2006; Mahon et al. 2007; Bóna 2008; Dér–Markó 2010; Gósy–Horváth 2010; Gyarmathy 2010; Horváth–Gyarmathy 2010; stb.). A spontán beszédben azonban számos olyan megakadás található, amelyek mintegy sorozatot alkotnak. Úgy tűnik, a megakadás is terjedhet, a beszélőnek olykor nem csupán nem sikerül időben a korrekció elvégzése, hanem újabb diszharmóniás jelenségekkel kell szembesülnie, amelyek következtében újabb megakadások tapasztalhatók a felszínen. Felmerül a kérdés, hogy miként működik vagy nem működik ilyen esetekben az önmonitorozás, illetve a javítás. Dell négyféle működési hibát eredményező tényezői hatást gyakorolhatnak-e a beszédtervezésre egymást követően vagy akár egyidejűleg is? A hibákat előidéző hatások komplexen jelennek meg, vagy egymást követően lépnek fel egy adott közlési helyzetben? Boomer és Laver már több mint négy évtizeddel ezelőtt kijelentették (1968), hogy a spontán beszéd egységeit (frázisoknak nevezték) az is meghatározza, hogy azok határain nem fordulnak elő megakadások. Vajon a sorozatban megjelenő diszharmóniás jelenségek is megmaradnak-e a beszédegységek keretein belül, avagy átléphetik azokat? Dell négy tényezője közül három, a kivitelezési eltérés, valamint a hasonlósági és a távolsági hatás azt erősítik meg, hogy a létrejött megakadások a beszédegységen belül valósulnak meg. Kérdés azonban, hogy a beszédsebesség hatása felülírhatja-e ezt a tényt? Megválaszolandó továbbá az is, hogy ez a négyféle tényező valóban felelőssé tehető-e az összes definiált hibajelenségért, avagy

továbbiakat is számításba kell venni. Úgy gondoljuk, hogy a beszélő (leggyakrabban gondolati szintű) bizonytalanságai további okokra vezethetők vissza, amelyekben a delli tényezők közül kettőnek egyértelműen lehet szerepe, a távolsági hatásnak és a beszédsebességnek. A kivitelezési eltérés tartalmát bizonyos mértékig kibővítve, ez is eredményezhet bizonytalansági megakadást. Ekkor a beszélő úgy érzi, hogy az elhangzott közlés(részlet) nem adekvátan fejezte ki a szándékozott gondolatot. Ezt diszharmóniának éli meg, aminek következményei vannak a kivitelezésben, sőt a beszélő emiatt módosíthatja, „javíthatja" az elhangzottakat. Tényező lehet továbbá az, amikor a beszélő a kivitelezés során vagy közvetlenül utána meg kívánja változtatni az eredeti gondolatot, „másként dönt" vagy másként szelektál a gondolatok között, és ez új nyelvi tervezést igényel. Ezt a helyzetet újratervezésnek nevezzük. Valószínűsítjük, hogy mindezen hatások többé-kevésbé egyidejű fennállása befolyásolja a beszédprodukciós mechanizmust, és ezen a módon magyarázhatja az egymást követően megjelenő megakadásokat.

Megállapíthatjuk, hogy a spontán beszéd olyan felszíni megakadásjelenségeket (is) tartalmaz, amelyek bizonyos tekintetben eltérnek a diszharmónia következtében létrejövő, szokásos jelenségektől. Ezek a megakadások ugyanis egymást követően, mintegy sorozatban fordulnak elő a spontán közlésekben, és a hallgató számára a megszokottal szemben a beszélő nagyobb mértékű beszédtervezési nehézségeit jelzik. Két példa ezekre a sorozatmegakadásokra (zárójelben az adott megakadásjelenség időtartama látható ms-ban, a négyzet [□] a néma szünetet, az ***öööö*** és a ***mmm*** a kitöltött szünetet [hezitálást], a kettőzött betű a nyújtást jelöli, míg a félkövér betűkkel magukat a megakadásokat szemléltetjük):

a) ***éss*** (136) ***hát*** □ (162) *járhattamööö* (150) *volna még* ***térítés*** □ (450) ***ess*** (511) ***öööö*** (363) ***mmm*** (338) ***tát*** □ (168) ***aa*** (133) □ (104) ***a költségtérít*** □ (132) ***té*** □ (184) ***tést***

b) erkölcsileg ***beta*** (259) ***betargya*** (560) ***öööö*** (105) *be is* ***tu*** □ (399) *betartja és* □ (239) *be is tudja tartani*

Az *a)* példa egy időben hosszabb jelenséget szemléltet, amely összesen 14 darab bizonytalanságra utaló megakadást tartalmaz a következő sorrendben: nyújtás, töltelékszó, néma szünet, hezitálás, szünet a szóban, nyújtás, hezitálások, töltelékszó (a *tehát* rövidített formában: *tát*), nyújtás, néma szünet, ismétlés, szünet a szóban és (újabb) szünet a szóban. A *b)* példa rövidebb sorozatmegakadás, itt 6 egymást követő szándéktalan jelenség (bizonytalanságra utalók és hiba típusúak) tapasztalható: újraindítás, egyszerű nyelvbotlás, hezitálás, anticipáció és néma szünet.

A sorozatmegakadások **definícióját** a következőképpen fogalmazzuk meg: a felszínen kettőnél több megakadásjelenség előfordulása egymást követően, illetőleg oly módon, hogy a megakadásjelenségek között legfeljebb két tartalmas szó hangzik el olyan kontextusban, amely tartalmilag összefügg. (Megjegyezzük, hogy az esetek egy részében a dallamív egyfajta prozódiai

egységet is tükröz.) Kutatásunk célja a sorozatmegakadások minőségi és temporális jellemzőinek meghatározása, a mintázatuk sajátosságainak leírása. Kutatási kérdéseink a következők. Megállapítható-e, hogy mi idéz elő sorozatmegakadást a spontán beszédben? Vannak-e típusai a sorozatmegakadásoknak? Milyen beszédprodukciós szintekkel hozható kapcsolatba a sorozatmegakadás? Milyen diszharmóniás és temporális sajátosságokat mutatnak a sorozatmegakadások?

Feltételezésünk szerint *a)* a sorozatmegakadások nem véletlenszerűek abban az értelemben, hogy meghatározható okra vezethetők vissza, és csoportszinten igazolhatók; és *b)* sajátos mintázatot mutatnak. Ez a mintázat nagy valószínűséggel behatárolja a diszharmónia eredetét. Feltételezzük továbbá, hogy a sorozatmegakadások jellemzőinek megismerése tovább pontosítja ismereteinket a beszédtervezésről és a beszélő tervezési stratégiáiról.

Kísérleti személyek, anyag, módszer

Kutatásunkban tizenkét felnőtt beszélő (6 nő, 6 férfi) spontánbeszédanyagát (interjúkat) elemeztük a BEA adatbázisból (Gósy 2008). Életkoruk 22 és 60 év közötti volt. Összesen 117,98 percnyi anyagot elemeztünk, adatközlőnként átlagosan 9,83 percnyit. Összesen 76 sorozatmegakadást azonosítottunk; adatoltuk az egyes diszharmóniajelenségek felszíni előfordulásait, és megmértük a releváns időtartamokat a Praat 5.0 szoftver segítségével (Boersma–Weenink 2009). Az időtartammérések során meghatároztuk az egyes sorozatmegakadások teljes időtartamát, a bennük található néma szünetek, hezitálások (kitöltött szünetek), nyújtások és az ismétlések (az első és a második szó teljes) időtartamát. (Az időtartammérések során a fonetikai szakirodalomban szokásos kritériumoknak megfelelően jártunk el.) A nyújtásokat percepciós alapon határoztuk meg az előfordulás kontextusában. Néhány esetben a biztos azonosítás érdekében két fonetikus kolléga véleményét figyelembe véve döntöttünk. Kategorizáltuk továbbá az összes olyan megakadásjelenséget, amelyek a sorozatokban előfordultak (a „Nyelvbotlás"-korpusz 2009 szerint, vö. Gósy et al. szerk. 2009); például téves kezdés, grammatikai hiba vagy a lexikális hozzáférés problémája (téves szó találása, „nyelvem hegyén van" jelenség, téves kezdés).

A sorozatmegakadás teljes időtartamának a benne előforduló összes megakadás időtartamának összegét tekintettük (az időtartammérésekben nem vettük figyelembe a töltelékszavak időtartamát). A következő példában – ***hanem*** (327) ***hanem*** (398) ***öööö*** (504) □ (496) ***ilyen*** szigorításokkal meg ***mmm*** (687) belátásra ***hogy*** (186) □ (458) ***hogy*** (283) ***ugye öööö*** (309) *elmagyarázni* – a megakadások teljes időtartama 3648 ms-ot tesz ki. Megmértük a nyújtottan ejtett magán- és mássalhangzók időtartamát, a feldolgozásban azonban nem különítettük el a beszédhangokat sem kvalitatív, sem kvantitatív szempontból. A nyújtott beszédhangokat egyetlen kategóriaként kezeltük. A hezitálásokat fonetikai jellemzőik szerint kategorizáltuk (vö. Horváth 2010), továbbá

elkülönítettük a szóhoz (az első vagy az utolsó szótaghoz) tapadó hezitálásokat és a két szünet között megvalósuló, izolált ejtésű kitöltött szüneteket. A jelen kutatásban néma szünetként adatoltunk minden olyan jelkimaradást, amely 100 ms-nál hosszabb időtartamú volt. Ez az időérték önkényes, de a tapasztalat és a szakirodalmi adatok is arra utalnak, hogy a grammatikailag és szemantikailag összetartozó beszédegységek esetében a lélegzetvételt szolgáló néma szünet időtartama rendszerint 100 ms alatti. Az ennél rövidebb néma szüneteket a hallgatók alig vagy egyáltalán nem észlelik a spontán beszédben (pl. Duez 1993; Váradi 2008). Jelöltük azokat a jelkimaradásokat, amelyeknél sóhajtás volt hallható a felvételen. Az ismétlések esetében a szó első és második megjelenésének teljes időtartamát határoztuk meg. A statisztikai elemzéseket az SPSS 19.0 verziójával végeztük (egytényezős ANOVA, Tukey post hoc tesztek, párosított *t*-próba) 95%-os megbízhatósági szinten.

Eredmények

1. Elemzett anyagunkban átlagosan 1,55 percenként jelentek meg sorozatmegakadások, előfordulásuk átlaga beszélőnként 5,58 darab. Az időarányosan legkevesebb sorozatmegakadást produkáló beszélőnél hármat, a legtöbbet produkálónál pedig 13-at találtunk. Az egyes megakadások átlagos száma 6 volt (a leghosszabb sorozatmegakadás 15 megakadást tartalmazott, és 8,85 másodpercet vett igénybe). Meghatároztuk, hogy az egyes beszélők hány percenként ejtenek sorozatmegakadást. Adatközlőink éppen fele 1,5 percnél ritkábban, míg a másik fele ennél gyakrabban produkált sorozatmegakadást. A leggyakrabban 0,66 percenként (férfi beszélő), a legritkábban pedig 2,02 percenként (női beszélő) fordult elő. Mintegy 2 percenként tapasztaltunk sorozatmegakadást egy másik férfibeszélőnél, illetve viszonylag gyakran (1,14 percenként) fordultak elő ezek egy másik női beszélőnél.

A sorozatmegakadásokat a kontextuselemzés alapján három típusba soroltuk. Az első típusra az volt a jellemző, hogy a beszélőnek nyilvánvalóan a közlésre szánt gondolatok kiválasztása (a mondanivaló meghatározása) okozott nehézséget. Valószínűsíthetően nem tudta eldönteni, hogy hogyan folytassa a beszédét, illetve hogy mit is mondjon, avagy nem tudott szelektálni a versengő gondolatok között, hogy melyiket is hangosítsa meg. (a továbbiakban: „gondolati" típus). Például: ***dee*** (397) □ (169) ***öööö*** (427) ***én*** (158) ***én*** (133) *azt hiszem* [*asszíszem*] *hogy* □ (242) ***engemm*** (232) ***öööö*** (114) ***hát*** *nem érint*. Az ilyen okokra visszavezethető sorozatmegakadások azonosítását segíti, hogy a beszélő gyakran nyelvileg is megfogalmazza az aktuális problémáját. A jelen példában erre utal az *azt hiszem* közlésrészlet, avagy ilyenek a sajátos önreflexiók, mint *hogy is mondjam*, *na most leblokkoltam*, *na hogy fogalmazzam* stb. Azt nem tudhatjuk biztosan a fenti példa esetében, hogy a gondolati problémát eredendően mi okozta a beszélőnek, de a hosszabb kontextus alapján feltételezhetjük, hogy el kellett döntenie, hogy az adott kérdésben miként foglal állást. Időnyerésül hezitált, az első rövidebb néma szünet

után hosszabb néma szünetet tartott, miközben a kettő közé egy relatíve hosszú időtartamú hezitálás ékelődött. A személyes névmás megismétlése alatt további időt szánhatott a gondolat kiválasztására, az *aszíszem* (*azt hiszem*) a jelen esetben nagy valószínűséggel töltelékelemként funkcionált. A *hogy* kötőszót követő néma szünet, az *engem* szó szóvégi mássalhangzójának nyújtása, az azt követő hezitálás és az újabb (ezúttal egyértelmű) töltelékszó (*hát*) arra utalnak, hogy a véleményformálás nehezen jött létre a gondolati szinten. A példában csak bizonytalanságra utaló, illetve időnyerést biztosító megakadások fordultak elő. Hangsúlyozzuk, hogy a gondolati típus mint kategória nem általában véve az emberi gondolkodásra kíván utalni, hanem a mindenkori beszélő mondandójára.

A második típusba azok a sorozatmegakadások kerültek, amelyeknél egyértelműen adott volt a közölni szándékolt gondolat, azonban egy vagy több tervezési szinten is nehézségbe ütközött a beszélő, vagyis a nyelvi átalakítás folyamatában adódtak diszharmóniák (a továbbiakban: „formai"). Például: *hogy* ***tanít*** □ (293) ***hassak*** /sóhaj/ (409) *meg hogy egy* □ (180) ***ii*** (293) ***gyööövööö*** (551) ***mmm*** (206) *fergeteges órát tartsak*. A kontextuselemzés alapján állítható, hogy a beszélő pontosan tudta, mit akar mondani, milyen gondolatot akar hangosan megfogalmazni. A nyelvi átalakítás során azonban valószínűsíthetően többféle nehézségbe ütközött. A *hogy* kötőszót követő igető kiejtése után elbizonytalanodik (talán, hogy a megfelelő szót aktiváltae), majd a szünetet követően befejezi a lexéma kiejtését, ezt egy jól hallható sóhaj követi. Ez akár értelmezhető oly módon, hogy sikerült megoldania a lexikális hozzáférés pillanatnyi problémáját. A folytatásban azonban ismét nehézsége támad a lexikális aktiválásban, ezt jelzi a téves kezdés. Ezt követően egy nyelvbotlássorozatot produkál (amely többszöri meghallgatás után sem volt kapcsolatba hozható semmilyen szóval). Ezt kitöltött szünet (hezitálás) követi, amely után a közlésrészletet nehézség nélkül fejezi be. A példában hiba típusú megakadásokat látunk (téves kezdés, nyelvbotlás), de bizonytalanságra utalók is megjelennek (szünet a szóban, hezitálás).

A harmadik típus olyan sorozatmegakadásokat tartalmaz, amelyeknél mind a gondolati tervezés, mind a nyelvi átalakítás szintjén feltételezhető volt a zavar, valamiféle diszharmónia, és ennek jellegzetes következményeit tapasztaltuk a felszínen (a továbbiakban „kevert" típus). Például: *nehéz pálya és nem is* □ (117) ***ööö*** (524) □ (1030) ***szóval*** *nem fizetik* ***valame*** □ (114) *semennyire meg*. A szövegkörnyezet ismeretében kijelenthető, hogy a beszélőnek nehézséget jelentett az anyagiakra vonatkozó gondolatnak a legmegfelelőbb, esetleg eufemisztikus megfogalmazása. Erre utal a rövid néma szünet, majd a hosszabb hezitálás, amit egy igen hosszú néma szünet követ. A *szóval* töltelékszó (*szal* formában) ejtése még mindig a mondanivaló pontos kiválasztásának problémát igazolja. A mondandóról hozott döntést követően a nyelvi átalakításba lexikális aktiválási hiba csúszik, amit a *valame* téves kezdés és az azt követő rövid néma szünet jelez. A szünet után elhangzik a meg-

felelő szó (*semennyire*). A példában mind bizonytalanságra utaló, mind hiba típusú megakadás megjelenik.

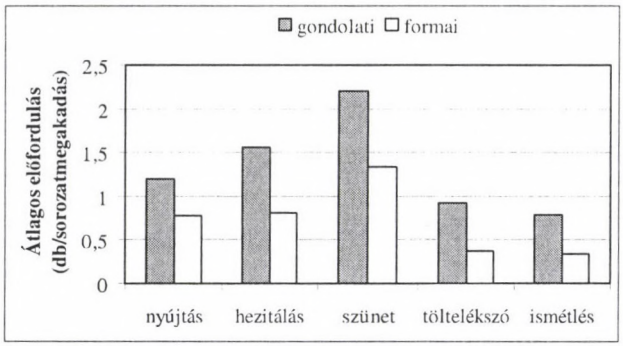
Valamennyi sorozatmegakadás kontextuselemzése, illetve a megakadások kategorizálása alapján megállapítható, hogy azok a jelenségek, amelyek a gondolatok kialakításával függenek össze, csak bizonytalanságra utaló megakadásokat tartalmaznak. Azokban a sorozatmegakadásokban, amelyek a nyelvi átalakítás nehézségeit tükrözik, jellegzetes hiba típusú megakadásokat tapasztalunk, de a bizonytalanságra utalók is megjelenhetnek bennük. Ennek a típusnak 9,3%-a kizárólag hiba típusú megakadásokat tartalmazott, például: *és a* ***modort*** (330) *a motort* ***mem*** (140) ***mek*** (96) *megmutatják*. Itt egy egyszerű nyelvbotlást a felszínen kétszeri téves kezdés követ, amelyek közül az első perszeverációval, a második anticipációval is magyarázható. Az ebbe a típusba tartozó sorozatmegakadásokban a bizonytalanságra utaló jelenségek közül a leggyakoribbak a hezitálások, a néma szünetek, ritkábbak a nyújtások, és még ritkábban fordulnak elő töltelékszók, szóban tartott szünetek és ismétlések. A mind a mondanivaló tekintetében, mind a nyelvi átalakításban diszharmóniát mutató sorozatmegakadások mindkét típusú megakadásjelenséget tartalmazzák.

A 76 sorozatmegakadás közül 51,3% (39 db) a gondolati tervezés, míg 42,1% (32 db) a nyelvi átalakítás szintjeihez volt köthető. A kevert típusúak aránya a legkisebb, mindössze 6,6% (5 db). Ez utóbbiakat – kis számuk miatt – a további összehasonlító elemzésekben nem vesszük figyelembe, de a típust magát jellemezzük. A gondolati típusú sorozatmegakadásokban található összes megakadásjelenség száma 261, a formaiakban előfordulóké 157, az öt kevert típusúban pedig 29-et találtunk (az összes elemzett megakadás 447 volt).

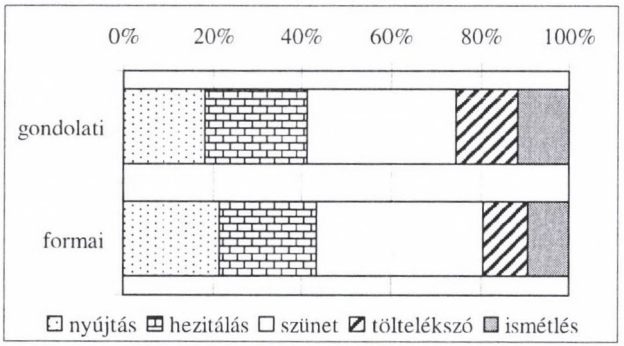
A bizonytalanságra utaló megakadások – mint említettük – jellegzetesen a gondolati típusú sorozatmegakadások részei. A nyújtások, hezitálások, néma szünetek, töltelékszók és a (szándéktalan) szóismétlések nagyobb arányban jelennek meg ezekben, mint a nyelvi átalakítás nehézségét mutató sorozatmegakadásokban (1. ábra).

A gondolati és a formai sorozatmegakadásokban külön-külön elemeztük a bizonytalansági megakadások arányait. A nyújtások és a néma szünetek valamivel nagyobb arányban jelennek meg a formai nehézségek esetén, mint a gondolati szinten előálló problémáknál. A töltelékszók és az ismétlések ugyanakkor a gondolati szintű problémák esetén fordulnak elő nagyobb arányban a formai nehézségek sorozatmegakadásaihoz képest (2. ábra). Néhányszor adatoltunk újraindítást és szóban tartott szünetet is.

A nyelvi átalakítás nehézségei különféle beszédtervezési szinteken álltak elő, ezeknek megfelelően különféle felszíni megakadásokban öltöttek testet. Anyagunkban 40 hiba típusú megakadás fordult elő a formai és 8 a kevert típusúakban. A hét gyakoribb típus arányait szemlélteti a 3. ábra. A szótalálási nehézség kétféle jelenséget foglal magában, téves szót és a nyelvem hegyén van problémát.



1. ábra
A bizonytalanságra utaló megakadások előfordulása
a gondolati és a formai sorozatmegakadásokban



2. ábra
Az egyes bizonytalansági megakadások aránya
a gondolati és a formai típusú sorozatmegakadásokban

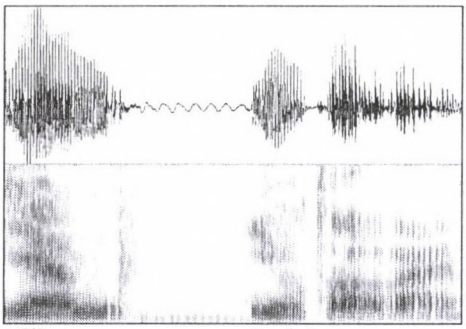
Legnagyobb arányban a felszínen a mentális lexikon aktiválásával kapcsolatos problémákra utaltak az adatok, elsősorban téves kezdések és szótalálási nehézségek formájában, ezeket követték a grammatikai hibák, az egyszerű nyelvbotlások, majd az anticipációk. A lexikális hozzáférés felszíni jelenségei azonban nemegyszer egyéb diszharmóniára is utalhatnak a rejtett beszédtervezés bármely szintjén (vö. Gósy 2012). Téves szó ejtésére két adatot, perszeverációra mindössze egyet találtunk, ez utóbbi a következő volt: *vagy ingyen adni* □ (294) ***öööegy*** (253) ***i***[j]***en*** □ (290) ***jutalványt tát*** *utalványt* (vö. 4. ábra hangszínképe). A példában a beszélő nem tudja a kívánt pillanatban aktiválni az *utalvány* szót, ezt jelzi a hosszabb néma szünet és a névelőt (*egy*)

megelőző, ahhoz tapadó hezitálás. Minthogy még mindig nem sikerül a hozzáférés, ezért töltelékszót ejt (*ilyen*), amelynek az intervokális helyzetű közelítőhangja aktív marad a 290 ms időtartamú néma szünetben is. Ennek következtében a beszélő a végre sikeresen aktivált szó első hangjaként ki is ejti. Észre is veszi a hibát és a *tát* (*tehát*) töltelékszó kiejtését követően javít. A javítás ténye egyértelművé teszi a perszeverációs hiba tényét. A töltelékszó jelezheti az önmonitorozás tényét, illetőleg a hibadetektálást, de azt is, hogy a mechanizmus felkészült a korrekcióra, sőt jelezheti akár mindkettőt is. Az [i] magánhangzót (az *ilyen* szóban) követi az approximáns ejtése, amely a rá következő két beszédhang és a jelkimaradás (ez a szakasz halk, mormogásszerű hangot tartalmaz) ellenére csaknem teljesen megőrzi az előző artikulációs gesztust. Ennek lenyomata látható a *jutalvány* szó [u] magánhangzóját megelőző [j] ejtésében. A két közelítőhang akusztikai szerkezetében a tiszta fázis alig mutat különbséget, csupán a koartikuláció következtében módosul kissé a második ejtésben az átmeneti szakasz (mivel az [u] második formánsa alacsonyabb frekvenciaértéken van, mint az [ɛ] F_2-je).

3. ábra
A nyelvi átalakításban jelentkező hiba típusú megakadások előfordulásai

Az öt darab kevert típusú sorozatmegakadásban összesen 29 diszharmóniajelenséget adatoltunk, közülük 8 volt hiba, és 21 utalt bizonytalanságra. A leggyakrabban a néma szünet fordult elő (az összes bizonytalanságra utaló megakadást tekintve 100%-nak 38,1%-ban), ennél jóval ritkábbak voltak a hezitálások (23,8%), még ritkábbak a töltelékszók (14,3%) és az ismétlések (14,3%), s a legritkábbak a nyújtások (9,5%). A kevert sorozatmegakadásokban négyféle hiba típusú megakadást találtunk: téves kezdéseket (55,4%-ban), grammatikai hibákat (22,2%-ban), nyelvbotlást (11,2%-ban) és kontaminációt (11,52%-ban). Az egyik leghosszabb sorozatmegakadás kevert típusú volt: *tanulmányait* ***tehát*** *ha ilyenre* ***öööö*** (425) ***elő*** □ (362) ***ffordult*** (280) *vagy sor kerül még*

egyszer ***tát*** *semmi* ***ki*** *mindig is* ***aa*** (189) □ (303) ***aa*** (231) □ (560) ***aa*** (383) ***mmm*** (182) *igazság* ***mgü*** (415) *felől.* A megakadások nélküli (szélesebb) kontextus a következőket tartalmazza: „akár olyat megmondani a szülőnek, hogy ha még egyszer ilyen előfordul, akkor ugye eltiltatják a gyereket, nem folytathatja a tanulmányait, tehát ha ilyenre előfordul vagy sor kerül még egyszer, tehát semmi mindig is az igazság felől kellene eljárni".

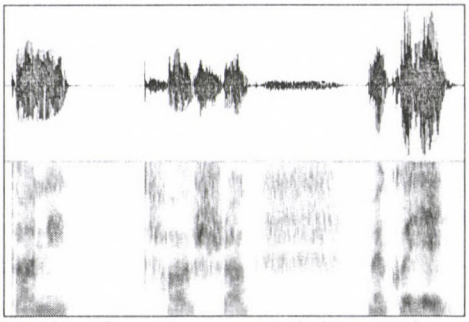


4. ábra
A [j] közelítőhang perszeverációjának ejtéséről készült hangszínkép (férfi beszélő)

A sorozatmegakadások nemegyszer tartalmaznak utalásokat a beszélő produkciós nehézségére; ezek segítik az osztályozást, a típusokba sorolást. Egy *na hogy is mondják* közbevetés esetében a beszélő az *antipátia* szót kereste, egy másik adatközlő a *na hogy mondjam* beékelődéssel pedig arra utalt, hogy nem tudja nyelvileg megfogalmazni a gondolatot. Ezek a közbevetések rendszerint a legtermészetesebb módon illeszkednek a szövegbe, mint például a következő sorozatmegakadásban, amelyet valószínűsíthetően egy szótalálási probléma indít el: *hogy vannak* ***öööö*** (315) *pozitív és negatív* □ (150) ***öö*** (600) □ (2600) ***öööö*** (82) □ (1200) ***na most leblokkoltam*** (1238) ***táát*** (227) *a* □ (2488) ***öööö*** (583) *a jó és* ***hi*** *rossszz* (215) ***öööö*** (360) □ (340) *következményei.* A szöveg folytatásában a beszélőnek a megjegyzése, hogy *nem tudom, mit is mondjak,* arról tanúskodik, hogy a szótalálás problémája a mondanivaló kialakítására is hatással volt. A következő példa a nyelvi átalakítás problémáját szemlélteti. A lexikális előhívás nehézségeit jelzi az első esetben egy közbevetés (ekkor a *forgalmi* szót nem tudja a beszélő gyorsan lehívni), a másodikban pedig nyújtott névelő, néma szünet és töltelékszó (*kvázi*) révén sikerül a beszélőnek a *kalauz* lexéma elérése: *a kocsidnak* ***aa*** (176) □ (930) *mi az már* (697) □ (687) *forgalmijátööö* (224) □ *bemutattad* ***aa*** (183) □ (179) ***kvázi*** *a kalauznak.* A töltelékszó funkciója itt összetettebb, mivel arra is utal,

hogy a beszélő tudatában van a nem megfelelő szó ejtésének (a *rendőr* szót kellett volna mondania).

A néma és a kitöltött szünetek mellett számos ún. egyéb, nem nyelvi hangadás jelenik meg a sorozatmegakadásokban, amelyek mind-mind stratégiaként szolgálhatnak az átmeneti produkciós nehézségek feloldására, időnyerésre (5. ábra). Ilyenek a krákogások, a glottális zár, a jól hallható, sokszor hosszú sóhaj (aminek többféle funkciója lehet, a példában egyértelműen az időnyerést szolgálja), a cuppogások, nyammogások, (kényszerű) nevetés, kuncogás (vö. Neuberger 2012).



5. ábra
A sóhaj mint időnyerés akusztikai lenyomata

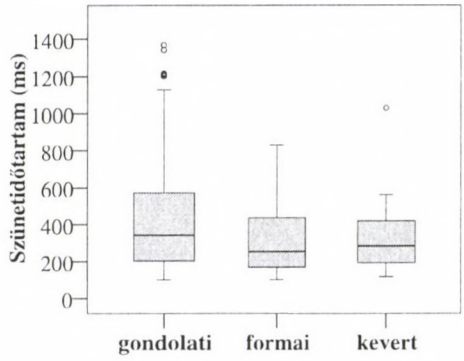
2. A sorozatmegakadások teljes időtartamára vonatkozó adatokat az 1. táblázat tartalmazza a gondolati és a formai típusokban. A megakadások összes időtartama a sorozatokban statisztikailag igazolható különbséget mutat a típustól függően [egytényezős ANOVA: $F(1, 69) = 10{,}206$, $p = 0{,}002$]. A gondolati nehézségre visszavezethető sorozatmegakadások hosszabb összidőtartamúak, mint a formai nehézségre visszavezethetők. Tájékoztatásul közöljük, hogy a kevert típusba tartozó sorozatmegakadások teljes időtartamára vonatkozó átlag 3042 ms, a legkisebb érték 647 ms, míg a legnagyobb 4594 ms volt. Ennek alapján azt mondhatjuk, hogy többségük hosszabb ugyan, mint a másik két típusba tartozók jó része, azonban az igen kis elemszám nem tesz lehetővé reális összevetést.

1. táblázat: A sorozatmegakadások összesített időadatai

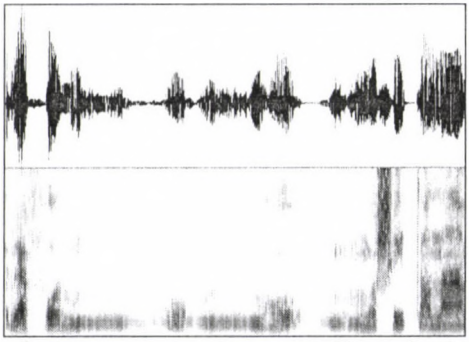
Típus	A sorozatmegakadások teljes időtartama (ms)			
	Átlag	Átlagos eltérés	Minimumérték	Maximumérték
Gondolati	2675	1811	610	9173
Formai	1550	913	258	3712

A néma szünetek felvételét a megakadások közé az indokolta, hogy a beszélők egyértelműen felhasználták időnyerésre, vagyis hasonló volt a funkciójuk, mint néhány megakadásjelenségé. Ezeknek az egymást követő néma szüneteknek a száma, illetve a relatív gyakoriságuk is inkább a megakadásjelenség funkciót támasztja alá, és nem az oxigéncsere szükségletét. A sorozatmegakadásokban adatolt (100 ms-nál hosszabb) néma szünetek csupán egy-két esetben fordulnak elő szintagmahatáron. A gondolati típusú sorozatmegakadásokban mért összes szünetet tekintve 100%-nak, 17,4%-uk időtartama esik 100 és 150 ms közé, 19,7%-uké 157 és 250 ms közötti. A többi 255 ms és 2600 ms között szóródik (a 250 ms-nál hosszabb szünetek döntő többségét már képesek a hallgatók észlelni, vö. Gósy 2000). Ez azt jelenti, hogy a sorozatmegakadásokban előforduló összes szünet mintegy kétharmadát a hallgatók 80%-osan vagy annál nagyobb mértékben felismerik. A formai típusban megjelenő néma szünetek közel fele 250 ms-nál rövidebb (100 ms és 150 ms között: 18,6%, 156 ms és 230 ms között 30,2%). A kevert típusúak néma szüneteinek mintegy 40%-a szintén rövidebb 250 ms-nál, az elemszámuk relatíve alacsony (12 db), de itt ezeket is tekintetbe vettük a statisztikai elemzésekben. Az eredmények szerint szignifikáns különbség van a néma szünetek időtartamában a sorozatmegakadás típusának függvényében [egytényezős ANOVA: $F(2, 141) = 3{,}996$, $p = 0{,}021$]. A Tukey post hoc teszt szerint a gondolati és a formai típusúak között szignifikáns a néma szünetek időtartamának különbsége ($p = 0{,}020$), a kevert típusokban előfordulók azonban egyiktől sem különböznek szignifikánsan. A gondolati típusú sorozatmegakadások néma szüneteinek időátlaga 536 ms (átl. eltérés: 524 ms), a formai típusúaké 320 ms (átl. eltérés: 197), a kevert típusúaké 348 ms (átl. eltérés: 254 ms). A legnagyobb a szóródás a gondolati típusúakban mért néma szünetek időtartamában (6. ábra). (Ez utóbbiban 3 extrém hosszúságút – 1778 ms, 2496 ms, 2600 ms – az ábra nem tartalmaz.)

Láttuk, hogy a nyújtások gyakoribbak a gondolati típusú sorozatmegakadásokban, mint a formai típusúakban, ugyanakkor az utóbbiakban nagyobb arányban jelennek meg mint időnyerő stratégiák. A nyújtások minden fonetikai helyzetben előfordulhatnak; valamivel gyakoribbak a szavak utolsó szótagjában, és jellegzetesen tapasztalhatók a határozott névelőkön (vö. Bóna 2008). A 7. ábra példája egy helyzetét tekintve ritkább előfordulású nyújtást szemléltet, egy szó közepi bilabiális nazális nyújtott ejtését, valamint annak az artikulációjába ékelődve egy nyelvbotlásszerű artikulációs váltást. Ez utóbbi a *poopöö* hangsor érzetét kelti a hallgatóban (amiben hezitálás is van). A szélesebb kontextus tartalma: „hoznék egy olyan döntést, reformosítanám azt, hogy ilyen helyzetekben". A *reformosítás* szót megelőző több izolált hezitálás, néma szünet és nyújtás jelzi, hogy a beszélő igyekszik a legmegfelelőbb szót megtalálni, de még a második szótag után is bizonytalan a lexikális hozzáférés eredményében, magában a szóban, ezért a szó artikulációs tervezése egyértelműen sérül.



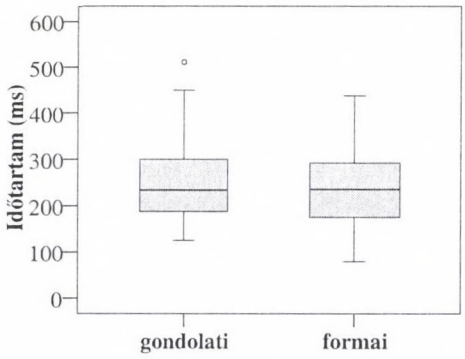
6. ábra
A néma szünetek időtartama a sorozatmegakadások három típusában (medián és szórás)



7. ábra
A szó közepén előforduló nazális mássalhangzó nyújtott ejtése és a közbeékelődő nyelvbotlásszerű hangsor (női beszélő)

Gyakrabban hallható a magánhangzók (főként az alsóbb nyelvállásúak), mint a mássalhangzók (közöttük főként a résképzésűek) nyújtott artikulációja. Anyagunkban a sorozatmegakadás típusát tekintve nem volt különbség a nyújtott ejtésű beszédhangok minőségében vagy a fonetikai pozíciójukban. A gondolati típusúakban megjelenő nyújtások időtartamának átlaga 257 ms (átl. eltérés: 95 ms), a formai típusban előfordulóké 233 ms (átl. eltérés: 88 ms). Az e két típusban adatolt nyújtások időtartama nem mutat szignifikáns különbséget, ami a nagy szórástartománnyal (is) magyarázható. A tendencia azonban

az, hogy a gondolati típusú sorozatmegakadásokban előforduló nyújtott beszédhangok hosszabbak, mint a formai típusúakban adatoltak (8. ábra). Az előbbiek esetében a legrövidebb nyújtott beszédhang 125 ms-os, a leghosszabb 510 ms-os volt, míg az utóbbiban a legrövidebb nyújtás 80 ms-osnak, a leghosszabb pedig 439 ms-osnak adódott. A kevert típusba tartozó sorozatmegakadások nyújtásainak átlaga 274 ms.



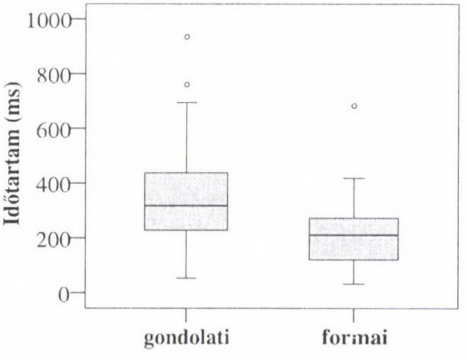
8. ábra
A nyújtások időtartama a sorozatmegakadások két típusában (medián és szórás)

A kitöltött szünetek többsége az *ö*-féle hanggal realizálódott (76,2% a gondolati és 83,9% az átalakítás típusban), jóval ritkábban fordultak elő *mm*-típusú hezitálások (23,8%, ill. 16,1%), a *höö* vagy *ööh* típusra pedig egy-két példa adódott. Ez a tendencia megfelel a spontán beszédben tapasztalt jellemzőknek (Horváth 2010). A mormogásra emlékeztető (*mm*-mel kitöltött) szünetek átlagidőtartama a gondolati típusúaknál 216 ms, a formai típusúaknál 282 ms; nem volt szignifikáns különbség a sorozatmegakadás típusának függvényében. Az *öö*-vel jellemzett hezitálások időtartamai között ugyanakkor szignifikáns különbséget találtunk a gondolati és a formai típusú sorozatmegakadások között [egytényezős ANOVA: $F(1, 85) = 11{,}863$, $p = 0{,}001$]. Ezeknek a hezitálásoknak az időtartamátlaga a gondolati sorozatmegakadásokban 359 ms (átl. eltérés: 202 ms), a formai típusúakban előfordulók átlaga 261 ms (átl. eltérés: 160 ms), vö. 9. ábra. A legrövidebb hezitálás az előbbiben 54 ms-os, az utóbbiban 37 ms-os volt, míg a leghosszabbak 934 ms és 683 ms tartamúnak adódtak. A kevert típusú sorozatmegakadásokban mindössze 6 hezitálást adatoltunk, ezek átlagos időtartama 432 ms volt (a legrövidebb 313 ms, a leghosszabb 524 ms).

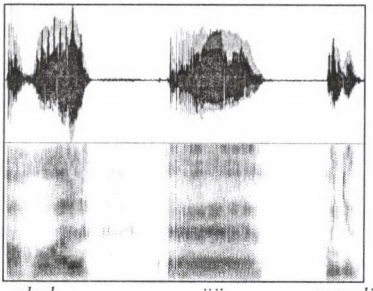
A sorozatmegakadásokban kétféle helyzetben fordultak elő a hezitálások, egyrészt két néma szünet között, másrészt egy szóhoz tapadva, mintegy an-

nak részeként. Ezek a tapadó hezitálások vagy a szókezdő beszédhanghoz vagy a szó utolsó beszédhangjához tapadva jelentek meg (10. ábra). Az artikuláció mindkét esetben folyamatos volt.

Elemeztük, hogy a tapadó és a nem tapadó hezitálások időtartamában van-e matematikailag igazolható különbség a sorozatmegakadásokban. Feltételeztük, hogy *a)* az izolált hezitálások hosszabbak lesznek, mint a tapadók; és *b)* minél magasabb tervezési szinten következik be a diszharmónia, annál több időre van szüksége a beszélőnek a korrekcióhoz. Ez pedig azt kell, hogy eredményezze, hogy a hezitálások időtartama a gondolati sorozatmegakadásokban lesz nagyobb. Az eredmények ugyan nem igazoltak szignifikáns különbséget az izolált és a tapadó hezitálások között, de a tendencia azt mutatja, hogy az izolált ejtés hosszabb időtartamban realizálódik. A gondolati típusú sorozatmegakadásokban az izolált hezitálások időtartamának átlaga 372 ms, míg a tapadóké 312 ms. A formai típusban előforduló hezitálások izolált ejtésben átlagosan 258 ms-osak, míg a tapadók átlagidőtartama 181 ms. Elemeztük továbbá, hogy az izolált és a tapadó hezitálások mutatnak-e függőséget a sorozatmegakadás típusától. Az izolált ejtésűek esetében nem kaptunk matematikailag igazolható különbséget, noha a tendencia arra utal, hogy a gondolati típusúak hosszabbak, mint a formai típusúak (11. ábra). Feltehetően a kis elemszám következtében nem sikerült szignifikáns eltérést kimutatni. A tapadó hezitálások időtartama ugyanakkor szignifikánsan különbözik attól függően, hogy melyik típusú sorozatmegakadásban fordul elő [egytényezős ANOVA: $F(1, 39) = 9{,}667$, $p = 0{,}003$]. A gondolati típusúakban az átlagos időtartam 312 ms (átl. eltérés: 142 ms), a formai típusúakban az átlag 181 ms (átl. eltérés: 112 ms), vö. 12. ábra.

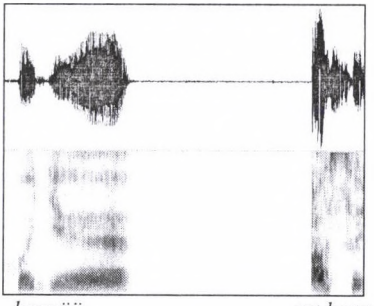


9. ábra
A hezitálások időtartama a sorozatos megakadások két típusában (medián és szórás)



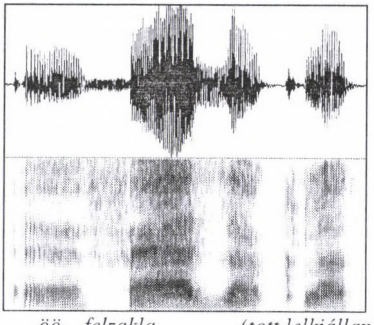
ahol *öö* *mondju(k)*

10. *a)* ábra



hogyöö *azt hogy*

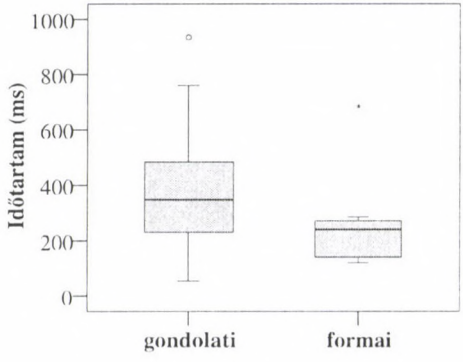
10. *b)* ábra



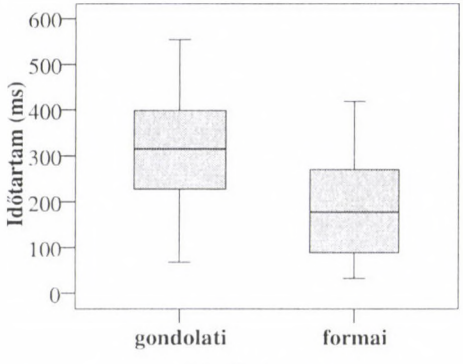
öö felzakla *(tott lelkiállapotában)*

10. *c)* ábra

a) Izolált ejtésű (két néma szünet közötti), valamint a *b)* szó elejéhez és *c)* végéhez tapadó hezitálások hangszínképe ugyanazon nő ejtésében (a zárójelek közötti beszédrészek a hangszínképen nem jelennek meg)



11. ábra
A nem tapadó hezitálások időtartama a sorozatmegakadások két fő típusában (medián és szórás)

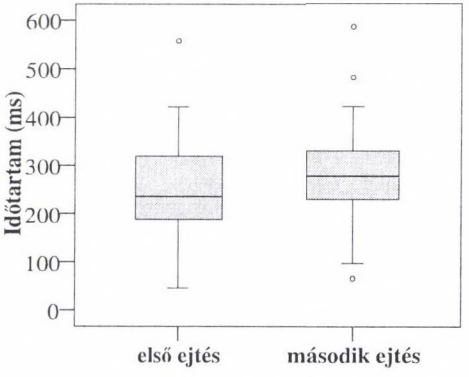


12. ábra
A tapadó hezitálások időtartama a sorozatmegakadások két fő típusában (medián és szórás)

A spontán beszédben végzett elemzések – nyelvtől függetlenül – azt igazolták, hogy a beszélők főként funkciószavakat ismételnek, és a másodikként kiejtett szó szignifikánsan rövidebb, mint az első (vö. Gyarmathy 2009). A sorozatmegakadásokban található ismétlésekre ezek a megállapítások csak részben igazak. Itt is elsősorban funkciószavak ismétlődnek; mindössze 5,4%-ban fordultak elő tartalmas szavak. Az időtartamok azonban éppen ellentétes tendenciát mutatnak (13. ábra).

A másodikként kiejtett szó időtartama a típustól függetlenül az esetek 57,77%-ában hosszabb volt. Az első és az ismételt ejtés között ugyanakkor

nem volt kimutatható szignifikáns különbség. Az első ejtés átlagos időtartama 258 ms (átl. elt.: 106 ms), az ismételt szavak időtartamának átlaga pedig 279 ms (átl. elt.: 98 ms). A nyelvi átalakítás nehézségét mutató sorozatmegakadások ismétlései nagyobb mértékben mutatják a második ejtés rövidülését.



13. ábra
Az elsőként és másodikként kiejtett (ismételt) szavak időtartama (medián és szórás)

Következtetések

Kutatásunkban egy eddig még nem elemzett jelenséget vizsgáltunk, amely a spontán beszéd egyik jellemzője. Sorozatmegakadásnak neveztük, és a terminust úgy definiáltuk, hogy az a felszínen kettőnél több megakadásjelenség folyamatos előfordulása egymást követően. A kritérium az volt, hogy az egyes megakadások között legfeljebb két tartalmas szó hangozhat el, és a kontextus tartalmilag összefüggő beszédrészlet legyen. A sorozatmegakadások mind a 12 beszélő spontán beszédében megjelentek, gyakoriságuk azonban különböző volt. A sorozatmegakadások sajátosságai, valamint a kontextus vizsgálata alapján három típust különítettünk el. A gondolatinak nevezett típusba azok a sorozatmegakadások kerültek, amelyeknél a nyelvi (és pszicholingvisztikai) elemzés eredményeképpen azt feltételeztük, hogy a diszharmónia oka a mondanivaló kialakításában, a gondolatok kiválasztásában (versengő gondolatok), egyfajta döntési helyzet problémájában, avagy a beszédszándék kismértékű módosulásában, illetve megváltoztatásában keresendő. A formainak nevezett típusba azokat a jelenségeket soroltuk, amelyekre a gondolat nyelvi formába öntésének nehézsége volt a jellemző. Ekkor a nyelvi átalakítás egy vagy több szintjén jött létre diszharmónia. A kevert típusba azok a sorozatmegakadások kerültek, amelyeknél a beszédprodukciós tervezés valamennyi szintjén jelentkezett diszharmónia.

A 76 adatolt sorozatmegakadás közül 39 az első, 32 a második típusba került, és mindössze 5-öt sorolhattunk a harmadik típusba. Ezek az adatok azt mutatják, hogy a beszélőknek a legnagyobb problémát a spontán megnyilatkozásokban a beszédszándék kialakítása, a mondanivaló, a gondolatok megformálása jelenti. Valamivel kisebb mértékben állnak elő a nyelvi tervezés nehézségei. Meglepően ritkák azok a sorozatmegakadások, amelyek a teljes tervezési mechanizmust érintik, ami a jól működő önellenőrzési folyamatokkal van összefüggésben. Úgy gondoljuk, hogy az a tény, hogy a sorozatmegakadások relatíve ritkák, a monitor működésével magyarázható. Ez gyorsan jelzi a problémát a beszélőnek, aki az első jelzéstől kezdve fokozottan igyekszik figyelni a tervezési folyamatokra, és így az esetek többségében sikeresen elkerüli a nyelvi tervezés további diszharmóniáinak kialakulását.

A sorozatmegakadások átlagosan mintegy másfél percenként fordulnak elő a narratívaszerű spontán közlésekben; és nagy eltéréseket mutatnak beszélőnként. Minthogy a mindennapokban az átlagos beszélők ritkán kerülnek olyan helyzetbe, hogy egy adott témáról hosszan kell beszélniük, ezért kimondható, hogy az önellenőrzés és a beszédtervezési stratégiák még az ilyen, kevésbé rutinszerű helyzetekben is igen jól működnek.

A sorozatmegakadások sajátos mintázatot mutatnak a bennük szereplő megakadások arányaiban általában, az egyes jelenségek sajátos felszíni mutatóiban specifikusan, valamint a temporális jellemzőikben. A gondolati típusúak kizárólag bizonytalanságra utaló megakadásokat tartalmaznak (pl. hezitálások, ismétlések, nyújtások), átlagosan 6,7-et jelenségenként. A formai típusúakban ezek átlaga 3,7. A nyelvi átalakítás nehézségét mutatókban és a kevertekben hiba típusú megakadások és bizonytalanságra utalók egyaránt megjelennek. A mintázatukat a kevert típusúak csekély elemszáma miatt nem tudtuk összevetően elemezni. A sorozatmegakadások egyes diszharmóniajelenségeinek gyakorisága és arányai eltérők az egyes típusokban. Az a tény, hogy a megakadások hatványozottan fordulnak elő a gondolati típusúakban, a gondolatok meghatározásának, a meghangosítani szándékozott tartalom kialakításának problémáját támasztja alá. A nyelvi átalakítás nehézségei legnagyobb mértékben a lexikális előhívás folyamatait érintették az elemzett sorozatmegakadásokban. Meglehetősen sérülékenynek mutatkozott az artikulációs tervezés, illetve a grammatikai műveletekben adódtak nem normatív jelenségek. Az előre tervezésből adódó hiba (anticipáció) jóval nagyobb mértékben jelent meg, mint a túlaktiválásból adódó (perszeveráció). A sorozatmegakadások ebben a tekintetben hasonló tendenciákat mutatnak, mint amit a spontán beszédre vonatkozóan általában már megállapítottak (Gósy et al. 2008; Gyarmathy 2009).

Az időviszonyok elemzése objektív értékekkel igazolta a feltételezéseket. A gondolati és a formai típusba került sorozatmegakadások teljes időtartamainak átlagai mintegy 1,12 mp-es különbséget mutatnak. Az e két típusba tartozók legrövidebb időtartamú előfordulásai között 340 ms, míg a leghosszab-

bak esetében 5461 ms, azaz közel 5,5 másodperc az eltérés. A temporális viszonyok is igazolják azt, hogy a beszélőnek a gondolati diszharmónia feloldásához sokkal több időre van szüksége, mint a nyelvi tervezési szinteken előálló hiba javításához. Az összesített időtartamok különbségei azt támasztják alá, hogy a gondolati szinten történő tervezés nagyobb nehézséget okoz(hat) a beszélőnek, mint a nyelvi átalakítás, másfelől, hogy ezeken a problémákon nehezebben jut túl. Az önellenőrzési és a korrekciós stratégia tehát feltehetően másként működik a gondolati és a nyelvi átalakítás szintjein (vö. Postma 2000).

A néma szünetek funkciója – a felszínen – gyakran nehezen állapítható meg. A jelkimaradás önmagában még nem igazolja azt, hogy a beszélőnek lélegzetvételre volt szüksége, avagy hogy ez a szünet csupán időnyerésre szolgált volna, és az sem zárható ki, hogy ún. hatásszünet okozta az artikuláció hiányát. Ez utóbbi tipikus beszélő esetén, az elemzett spontán közlésekben, különösen egy sorozatmegakadásban kizárható. Az mindig kérdéses, hogy ha történt is levegőpótlás, az kizárólagos célja volt-e a jelkimaradásnak, avagy a beszélő felhasználta az időnyerésként tartott néma szünetet lélegzetvételre. A szükséges levegőcsere is adhat lehetőséget beszédtervezési műveletek végzésére. A hangfelvételek e tekintetben nem nyújtanak teljesen megbízható információt. Feltételezzük azonban, hogy a kizárólag levegővétel funkciójú szüneteket a beszélők többnyire szintagmahatáron, közlésegység végén, tartalmi egység lezárásakor stb. tartják. Továbbá, a csak levegővételt szolgáló jelkimaradások rendszerint relatíve rövidebb ideig tartanak, mint az egyéb funkcióban megjelenők. Mindezek tudatában adatoltuk a néma szünetek időtartamát mindhárom sorozatmegakadás-típusban (azzal az ismételt megjegyzéssel, hogy a számuk a kevert típusúakban relatíve kevés volt). Az eredmények megerősítik annak a valószínűségét, hogy a sorozatmegakadásokban adatolt néma szünetek nagyon sokszor megakadásjelenség funkcióban fordulnak elő, még akkor is, ha a beszélő ezeket (néha) felhasználja lélegzetvételre. Ez utóbbi egyfajta beszélői stratégia, ami leplezi, hogy időnyerésre volt szükség. Az időtartamértékek eltérései alátámasztották azt a megállapítást, hogy a hosszabb néma szüneteknek valóban az a funkciójuk, hogy időt biztosítsanak a beszélőnek az önellenőrzésre, illetőleg a korrekcióra. A gondolati típusú sorozatmegakadások néma szünetei szignifikánsan hosszabbak, mint a nyelvi átalakítás szintjein bekövetkező diszharmóniás jelenségeket tartalmazó sorozatmegakadásokéi. Ez utóbbiak esetében kisebb a szórástartomány is.

A gondolati szinten és a nyelvi átalakításban bekövetkező problémák elkülönülését igazolják a nyújtások és a hezitálások időtartam-különbségei is. Mind a nyújtottan ejtett beszédhangok, mind a kitöltött szünetek időtartamai a gondolati sorozatmegakadásokban hosszabbak. Kimondható, hogy a nyelvi átalakítással kapcsolatos sorozatmegakadások esetén a beszélő igyekszik minél kevesebb feltűnés árán túljutni a problémán, míg a gondolati típusúak esetében valószínűleg képtelen erre, ezért ilyenkor a probléma a hallgató

számára is sokkal nyilvánvalóbb (vö. Cutler 1983). A kitöltött szünetek részben izoláltan fordultak elő, részben tapadtak egy szó elejéhez vagy végéhez. A gondolati típusú sorozatmegakadásokban a tapadó hezitálások aránya 41,0%, a nyelvi átalakítás diszharmóniáját mutatókban 61,5%. Feltételezhető, hogy a tapadó hezitálások mind az önellenőrzés, mind a korrekció folyamatainak gyorsabb, sikeresebb működését jelzik, mivel a mentális lexikon aktiválása és az artikulációs tervezés folyamatában egyidejűleg jelennek meg. Az izolált hezitálások arra utalnak, hogy a beszélőnek vagy az önellenőrzést, vagy a javítást, esetleg egyiket sem sikerült a szükséges időben véghezvinnie. Mindebből annak kell következnie, hogy a tapadó hezitálások rövidebbek az izolált ejtésűeknél. A mért adatok igazolják is ezt, típusonként és az összes hezitálást figyelembe véve is.

A sorozatmegakadásokban adatolt ismétlések temporális viszonyai bizonyos fokig ellentmondanak a spontán beszédre megállapított adatoknak. Az a tény, hogy itt az esetek többségében a második szó, vagyis az ismételt ejtés a hosszabb, illetőleg, hogy nincs szignifikáns különbség a két ejtés időtartamában, arra utal, hogy a sorozatmegakadások esetében az ismétlések nem jelentenek feltétlenül jó időnyerési stratégiát. Többszörös diszharmóniajelenség, illetve nehezebben megoldható gondolati szintű probléma esetén a szavak megismétlése, úgy tűnik, nem is annyira automatikus, mint egyéb esetekben.

Valószínűsítettük, hogy a kevert típusú sorozatmegakadások lesznek a leghosszabbak, hiszen itt biztosan egynél több tervezési szinten következett be diszharmónia. Az ilyen típusú példák csekély száma azonban nem teszi lehetővé a válaszadást. Ezeknek a sorozatmegakadásoknak a mintázata meglehetősen változatos; a teljes időtartamuk és a nyújtásoknak az ideje átlagosan hosszabb, mint a másik két típusban mértek esetében. A néma szüneteik időtartamának átlaga mintegy a két másik típuséi között foglal helyet; a hezitálások átlagos időtartama pedig ebben a típusban a legrövidebb.

Az elemzett sorozatmegakadásokban nem volt javítatlan megakadás. Ez arra enged következtetni, hogy noha az önellenőrzés és a korrekció egyaránt működik ezekben a jelenségekben, mintha éppen a javítás ténye is hozzájárulna a létrejöttükhöz. A beszélő az esetek egy részében nehezen korrigál, nehezen találja meg a kívánt tartalmat vagy formát, ez pedig negatívan hat vissza a tervezési folyamatokra; ráadásul szorongást okozhat, s a szorongás újabb diszharmónia kialakulásához vezethet, illetve nehezíti a korrekciót.

Adataink, illetve elemzési eredményeink alapján kijelenthetjük, hogy a Dell (1986) meghatározta tényezők, valamint az általunk ezek alapján módosítottak valószínűsíthetően egymást követően és egyidejűleg is hatást gyakorolnak a beszédtervezésre. A kivitelezési eltérés, a hasonlósági és a távolsági hatás összefügg egymással és a beszédsebességgel is, akármelyik sorozatmegakadás-típus adatait vesszük tekintetbe. Ezt pedig egyfajta aktivációterjedésként értelmezhetjük. Az egyes tényezők hatásának részletes elemzése azonban további vizsgálatokat igényel.

Kutatási kérdésünk volt, hogy milyen beszédprodukciós szintekkel hozható kapcsolatba a sorozatmegakadás. A sorozatmegakadások mintázatuk szerint a beszédprodukciós folyamat legkülönfélébb tervezési szintjeihez köthetők, és ezek a szintek azonosíthatók a megakadásjelenségek alapján. Választ kerestünk arra, hogy megállapítható-e, vajon mi idéz elő sorozatmegakadást a spontán beszédben. A meglévő adatok alapján erre csak valószínű választ adhatunk. A mondanivaló kialakításának problémája egyértelműen vezet sorozatmegakadáshoz akkor, ha a beszélő relatíve rövid időn belül nem tudja eldönteni, hogy mit, mennyit, miként mondjon, milyen véleményt alakítson ki valamiről, illetve ha az eredetileg elgondolt tartalmat kismértékben vagy teljesen meg akarja változatni, de bizonytalan abban, hogy ezt miként tegye. A nyelvi átalakítás szintjeit érintő sorozatmegakadások okaként szintén hipotéziseket tudunk megfogalmazni. Egy részüket nyilvánvalóan maga az első megakadás indukálja, mert ennek következtében az önellenőrzés fokozottabbá válik, ettől a beszélő elbizonytalanodhat, és a megfelelő formában is kételkedik. Ennek a következménye egyrészt a bizonytalanságra jellemző megakadások megjelenése, másrészt újabb hibák létrejötte. A fokozottabb önellenőrzés kiválthat szorongást a beszélőből; mind tudatosabban igyekszik elkerülni egy újabb diszharmóniás helyzetet, annál nagyobb a lehetősége egy további kialakulásának. Azt gondoljuk, hogy a belső és a külső monitor (Levelt 1989; Gyarmathy et al. 2009; Huettig–Hartsuiker 2010) talán kevésbé együttműködő a sorozatmegakadásokban. A temporális adatok arra engednek ugyanis következtetni, hogy a külső monitorozás nagyobb szerepet kap ilyenkor, ami egyben hozzájárul a sorozatmegakadások tartamának növekedéséhez is.

A kutatást megelőzően feltételeztük, hogy a sorozatmegakadások nem véletlenszerűek. Ez a feltételezés igazolódott, noha a jelen kutatás eredményei nem szolgáltathattak minden esetben kizárólagos magyarázatot a létrejöttükre vonatkozóan. Hipotézisünk volt, hogy a sorozatmegakadások sajátos mintázatot mutatnak, amit a nyelvi és a temporális elemzések egyértelműen alátámasztottak. Hangsúlyozzuk, hogy míg számos megakadásjelenség mind rögzített spontán beszéd, mind hallott és lejegyzett megakadáskorpuszok alapján vizsgálható, addig a sorozatmegakadások megismerése – a percepciós korlátok miatt – csak rögzített spontán beszéd elemzésével lehetséges. A sorozatmegakadások jellemzői tovább pontosítják ismereteinket a beszédtervezésről és a beszélő tervezési stratégiáiról, és ennek számos közvetlen gyakorlati felhasználása lehet, például a beszédtechnikában, a retorikában vagy a dadogás terápiájában.

Irodalom

Blackmer, Elizabeth – Mitton, Janet 1991. Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition* 39. 173–194.

Boersma, Paul – Weenink, David 2005. *Praat: Doing phonetics by computer.* http://www.praat.org/ [Letöltés: 2005. március 14.]

Boomer, Donald S. – Laver, John D. M. 1968. Slips of the tongue. *British Journal of Disorders of Communication* 3. 2–11.
Bóna Judit 2008. A nyújtás sajátosságai a dadogó és az ép beszédben. *Beszédkutatás 2008.* 148–156.
Clark, Herbert H. – Wasow, Thomas 1998. Repeating words in spontaneous speech. *Cognitive Psychology* 37. 201–242.
Corley, Martin – MacGregor, Lucy J. – Donaldson, David I. 2007. *It's the way that your, er, say it:* Hesitations in speech affect language comprehension. *Cognition* 105. 658–668.
Curl, Traci S. – Local, John – Walker, Gareth 2006. Repetition and the prosody–pragmatics interface. *Journal of Pragmatics* 38. 1721–1751.
Cutler, Anne 1983. Speakers' conception of the functions of prosody. In Cutler, Anne – Ladd, Robert D. (eds): *Prosody: Models and measurements.* Springer-Verlag, Berlin, 79–92.
Cutler, Anne 1988. The perfect speech error. In Hyman, Larry M. – Li, Charles N. (eds.): *Language, speech and mind: Studies in honor of Victoria A. Fromkin.* Croom Helm, London, 209–223.
Dell, Gary S. 1986. A spreading-activation theory of retrieval in sentence production. *Psychological Review* 93. 283–321.
Dell, Gary S. 1990. Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Process* 5. 313–349.
Dér, Csilla Ilona – Markó, Alexandra 2010. A pilot study of Hungarian discourse markers. *Language and Speech* 53. 135–152.
Duez, Danielle 1993. Acoustic correlates of subjective pauses. *Journal of Psycholinguistic Research* 22. 21–39.
Gósy Mária 2000. A beszédszünetek kettős funkciója. *Beszédkutatás 2000.* 1–14.
Gósy Mária 2008. Magyar spontánbeszéd-adatbázis – BEA. *Beszédkutatás 2008.* 194–207.
Gósy Mária 2012. Az artikuláció leállása a spontán beszédben. In Navracsics Judit (szerk.): *Pszicholingvisztikai tanulmányok III.* Tinta Kiadó, Budapest, 52–65.
Gósy Mária – Horváth Viktória – Bata Sarolta 2008. Szabálytalan „előzés" a beszédprodukcióban. *Magyar Nyelvőr* 132. 442–461.
Gósy, Mária – Horváth, Viktória 2010. Changes in articulation accompanying functional changes in word usage. *Journal of the International Phonetic Association* 40/2. 135–161.
Gósy Mária – Bóna Judit – Gráczi Tekla Etelka – Gyarmathy Dorottya – Horváth Viktória – Imre Angéla – Markó Alexandra – Neuberger Tilda (szerk.) „Nyelvbotlás"-korpusz 6. rész. *Beszédkutatás 2009.* 257–267.
Gyarmathy Dorottya 2009. A beszélő bizonytalanságának jelzései: ismétlések és újraindítások. *Beszédkutatás 2009.* 196–216.
Gyarmathy Dorottya 2010. A spontán beszéd időzítési zavara: a perszeveráció. *Beszédkutatás 2010.* 139–158.
Gyarmathy Dorottya – Gósy Mária – Horváth Viktória 2009. A rejtett és a felszíni önmonitorozás temporális jellemzői. In Keszler Borbála – Tátrai Szilárd (szerk.): *Diskurzus a grammatikában – grammatika a diskurzusban.* Tinta Kiadó, Budapest, 46–55.
Horváth Viktória 2004. Megakadásjelenségek a párbeszédekben. *Beszédkutatás 2004.* 187–199.

Horváth, Viktória 2010. Filled pauses in Hungarian: their phonetic form and function. *Acta Linguistica Hungarica* 57/2–3. 288–306.

Horváth Viktória – Gyarmathy Dorottya 2010. „A lónak is négy nyelve van, mégis megbotlik." A mentális lexikon útvesztői. *Beszédkutatás 2010.* 171–183.

House, David 2003. Hesitation and interrogative Swedish intonation. *Phonum* 9. 185–188.

Huettig, Falk – Hartsuiker, Robert J. 2010. Listening to yourself is like listening to others: External, but not internal, verbal self-monitoring is based on speech perception. *Language and Cognitive Processes* 25/3. 347–374.

Huszár Ágnes 2005. *A gondolattól a szóig. A beszéd folyamata a nyelvbotlások tükrében.* Tinta Kiadó, Budapest.

Kolk, Herman – Postma, Albert 1997. Stuttering as a covert repair phenomena. In Curlee, Richard F. – Siegel, Gerald M. (eds.): *Nature and treatment of stuttering: New directions.* Allye and Bacon, Needham Heights, Massachusetts, 182–203.

Levelt, Willem J. M. 1989. *Speaking. From intention to articulation.* A Bradford Book. Cambridge, Massachusetts.

Local, John 2007. Phonetic detail in talk-in-interaction: on the deployment and interplay of sequential context and phonetic resources. *Nouveaux cahiers de linguistique française* 2. 67–86.

Maclay, Howard – Osgood, Charles E. 1959. Hesitation phenomena in spontaneous English speech. *Word* 15. 19–44.

Mahon, Bradford Z. – Costa, Albert – Peterson, Robin – Vargas, Kimberly A. – Caramazza, Alfonso 2007. Lexical selection is not by competition: a reinterpretation of semantic interference and facilitation effects in the picture–word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33. 503–535.

Markó Alexandra 2004. Megakadások vizsgálata különféle monologikus szövegekben. *Beszédkutatás 2004.* 209–222.

Markó Alexandra 2006. A megakadásjelenségek hatása a beszédészlelésre. *Alkalmazott Nyelvtudomány* VI/1–2. 103–117.

Mowrer, Robert – Mackay, Ian 1990. Phonological primitives: electromyographic speech error evidence. *Journal of the Acoustical Society of America* 88. 1299–1312.

Neuberger Tilda 2012. Nonverbális vokális elemek a spontán beszédben. (közlésre elfogadva)

Nooteboom, Sieb 2005. Lexical bias revisited: Detecting, rejecting and repairing speech errors in inner speech. *Speech Communication* 47. 43–58.

Pérez, Elvira – Santiago, Julio – Palma, Alfonso – O'Seaghdha, Padraig G. 2007. Perceptual bias in speech error data collection: Insights from Spanish speech errors. *Journal of Psycholinguistic Research* 36. 207–235.

Postma, Albert 2000. Detection of errors during speech production: a review of speech monitoring models. *Cognition* 77. 97–131.

Postma, Albert – Kolk, Herman 1992. The effects of noise masking and required accuracy on speech errors, disfluencies, and self-repairs. *Journal of Speech and Hearing Research* 35. 537–544.

Poulisse, Nanda 1999. *Slips of the tongue. Speech errors in first and second language producton.* John Benjamins, Amsterdam.

Pouplier, Marianne – Hardcastle, William J. 2005. A re-evaluation of the nature of speech errors innormal and disordered speakers. *Phonetica* 62. 227–244.

Shriberg, Elizabeth. 2001. To 'errrr' is human: Ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association* 31. 153–169.
Váradi Viola 2008. A virtuális mondatok műfaji meghatározottsága. *Beszédkutatás 2008*. 134–147.

A kutatás a 78315 sz. OTKA-pályázat támogatásával készült.

BESZÉLŐVÁLTÁSOK SORÁN HASZNÁLT DISKURZUSJELÖLŐK A MAGYAR SPONTÁN BESZÉDBEN

Dér Csilla Ilona

Bevezetés

A nyelvészeti konverzációelemzés, azaz a természetes, kötetlen társalgásokat vizsgáló pragmatikai tudományterület elsősorban a beszélgetésekben megfigyelhető szabályszerűségeket, konvenciókat igyekszik feltárni (Iványi 2001: 77–78). A társalgások során az egyik alapvető feladat, hogy a beszélők eldöntsék, milyen sorrendben követik egymást, melyikük meddig beszél, hogyan adja át a szót, illetve veszi vissza, milyen verbális és nonverbális megnyilvánulásokkal él ezek kivitelezése során.

A társalgások legkisebb egységének a **beszédlépés**t (turn, [beszéd]forduló) szokás tekinteni. Beszédlépésnek az minősül, amit egy beszélő a szó átadásáig mond, illetve a szó átadása nélkül fejezi be a mondandóját (Sacks et al. 1974: 706). A beszédlépések változó hosszúságúak lehetnek, állhatnak akár egyetlen megnyilatkozásból, vagy több megnyilatkozás egymásutánjából; a prototipikus társalgás azonban kevés hosszú monológszerű részt tartalmaz. A beszélgetőpartnerek rendszerint folyamatos verbális és nonverbális visszajelzéseket (pl. hümmögés, igenlés, bólintás) nyújtanak egymásnak arról, hogy követik az elhangzottakat, illetve várják, hogy mikor kerül(het) hozzájuk a szó.

A **beszélőváltás** és a **beszédlépésváltás** (turn-taking) lényegében azonos értelemben használatos a nemzetközi és a hazai szakirodalomban, mindkettő azt a jelenséget takarja, amikor a beszédlépések váltakoznak, mivel megváltozik annak a személye, akinek éppen joga van az adott társalgásban beszélnie. A beszélő váltakozása azonban nem okvetlenül elégséges szempont a beszédlépések meghatározásában, mivel egyes résztvevők egyszerre is beszélhetnek; biztosabb fogódzó, ha a **beszédjog**ra támaszkodunk a fordulók mennyiségének meghatározásánál (ennek részleteit l. alább). A **beszédhez való jog** alatt azt értjük, hogy a társalgásnak az a résztvevője veheti át a szót az előző beszélőtől, aki a leggyorsabban reagál, esetleg az éppen beszélő megszólította, vagy kérdezett tőle valamit. Lényeges, hogy a szóátvétel nem azonos azzal, hogy valaki más kezd el beszélni, mint a korábbi beszélő, mivel az előbbi ettől még nem okvetlenül vette át a szót (részletesebben l. alább).

A beszélőváltás mechanizmusát, a beszédforduló lezárási módjait és a beszédjog átadását a szakirodalom különböző szempontokból már vizsgálta, így szintaktikai, fonetikai, szemantikai és pragmatikai tényezőket soroltak fel ezek

megvalósítására vonatkozóan (pl. Sacks et al. 1974; Stephens–Beattie 1986; a magyar nyelvet illetően lásd például Markó 2006; Németh 2007–2008; Bata 2009; Lerch 2011). A jelen tanulmány ezek közül a pragmatikai, azon belül pedig a beszélőváltáskor használt diskurzusjelölők (discourse markers) jellemzőivel foglalkozik. Korábbi spontánbeszéd-vizsgálatok már utaltak rá (Dér 2010), hogy számos magyar diskurzusjelölő tipikusan az általa bevezetett diskurzusszegmens elején fordul elő, a kötőszói eredetű jelölők (pl. *hát, tehát, és, de*) szinte mind ilyenek; kivételt azok képeznek, amelyek kötőszóként sem vagy nem mindig tagmondatkezdő helyzetűek (pl. *meg, pedig, bár*). Ebből következően a társalgások beszédlépéseit bevezető elemként is gyakorinak kell lenniük. A kérdés az, hogy a társalgásokban használatos diskurzusjelölők közül melyek, milyen gyakran és milyen funkciókban jelennek meg a fordulók élén vagy eleje felé.

Hipotéziseink a következők:

1. Az egyébként is gyakori jelölők (mint pl. a *hát, és*) fordulót felvezető egységként is számottevőek lesznek mennyiségileg.

2. A diskurzusjelölők elsősorban diskurzusszegmenseket kapcsoló elemek (Fraser 1999: 938), így természetszerűleg használatosak szóátvételt vagy szóátvételi szándékot jelző elemként is. E funkciójuk azonban egyidejűleg él a többi mellett, csak szóátvételre használatos, vagy aktuálisan kizárólag szóátvételt jelölő egység nincsen közöttük.

3. A szóátadáskor használt jelölők elsősorban nem a kötőszói eredetű diskurzusjelölők, hanem más szófajok köréből kerülnek ki, mivel a kapcsolódást az adott beszédforduló után a szót átvett beszélő saját beszédfordulója elején megjelenő elemek fogják jelezni.

A spontán beszélgetések jellemzően nem úgy valósulnak meg, hogy egyetlen beszélő magánál tartja a szót, majd elhallgat, és némi szünet után egy másik beszélő szólal meg; ez annál kevésbé van így, minél inkább informális társalgásról van szó, és minél jobban ismerik egymást a társalgó felek (Bata 2009: 117–118). Ennek következményeképpen a beszélők egyes megnyilatkozásai átfedhetnek, ami a társalgás természetes jelensége, ugyanakkor megfigyelhető, hogy a társalgók rendszerint kerülni igyekeznek mind a beszédfordulók közti túlságosan hosszú szüneteket, mind a hosszan átfedő beszédet, az előbbi ugyanis megakasztja a beszédet („kínos csönd"), az utóbbi pedig udvariatlan megnyilvánulásnak, erőszakos szóátvételnek minősül.

Mivel a jelen vizsgálat esetében kulcsfontosságú volt a beszédfordulók határainak kijelölése és megszámlálása, ezért az **egyszerre beszélések** megnyilvánulási formáinak tartható jelenségeket típusokba szedtük, és egyenként döntöttük el, hogy mely eseteket számítjuk egy, illetve több beszédlépésnek. Természetesen csak az olyan esetekről beszélünk itt, amikor a közbeszóló beszédpartner nem ragadja magához tartósan a szót, tehát jelentős szempont volt, hogy a „belebeszélés" nem haladhatta meg az egy mondatnyi szintet, ellenkező esetben új beszédlépést vettünk fel. Bár az egy tagmondatos meg-

nyilvánulások önálló beszédlépésként való felvétele mellett is hozhatók fel érvek, jobbnak éreztük a határt itt – és nem a közbeszólás puszta ténye alapján – megvonni (ez esetben ugyanis a pusztán a hallgatói/befogadói figyelmet kifejező megnyilvánulásokat is önálló beszédlépésnek kellett volna vennünk). Valójában minden közbeszólás (legalább egy) mondatszintű egységnek tartható, akár tagolt, akár tagolatlan szerkezetű.

Szempont lehetett volna az elhatárolásnál, hogy miként változtatja meg a másik fél közbevágása a beszédjoggal bíró fél megnyilatkozásának szintaktikai struktúráját, ez azonban – épp a spontán szövegek sajátosságai miatt – elméleti szempontból nehezen határolható el azoktól az esetektől, amikor a beszélő, noha senki nem vág a szavába, maga változtat a szintaktikai megformáláson; vagyis nem lehet tudni, hogy egy közbevágás után a beszélő amiatt változtatott-e a szerkezeten, vagy egyébként is változtatott volna rajta.

Az egyszerre beszélések jellemző megnyilvánulási formái a következők voltak:

a) **A figyelem verbális visszajelzése(i) a hallgatótól** (*ühüm, igen, ja, aha*), miközben az egyik beszédpartner beszél. Lényeges, hogy ebben az esetben az igenlő kifejezések és az ilyen funkciójú hümmögések nem kérdésekre vagy felvetésekre adott igenlő, egyetértő válaszok a hallgató részéről, hanem azt jelzik, hogy az figyeli, érti a másikat, és a további beszédre bátorítja. Az ilyen megnyilvánulásokat ezért nem vettük külön beszédfordulónak. Két példa [a konverzációelemzésben szokásosan használt transzkripciók helyett jelen esetben a BEA lejegyzési megoldásait alkalmazzuk, közülük is csak a következőket: ( ) = egyszerre beszélés; ! = a beszélő produkálta, nem a beszédhez tartozó hangok, pl. nevetés; A = fő adatközlő; T1 = első beszélgetőpartner, T2 = második beszélgetőpartner; a példákban mindig egyetlen bekezdésben hagytuk az egy beszédfordulóba sorolt megnyilatkozásokat]:

(1) (A: én úgy szeretem) egyébként hogy ott vannak körülettem (T2: **aha**) A: (és segítenek)

(2) A: egy kicsit még lesz valami négy négy nem két két kettő vagy három Varázsfuvola lesz Németországba T1: **ühüm** A: meg hát még júniusban lehet hogy lesz Hoffmann meséi is a Kovalikkal meg a Kovalik-növendékekkel itt az Operában

Az előbbi esetek különböznek attól a helyzettől, amikor az egyik beszélő szóra akarja bírni a másik felet, és nem magánál akarja tartani a szót, vagyis nem a beszélő szavába vágva szólal meg valaki. Ilyenkor új beszédfordulóval számoltunk, például:

(3) T1: ööö tehát ugye említette hogy most született ööö kis unokája A: **ühüm**
T1: és ööö arra sszeretném kérni hogy egyrészt hogy meséljen őróla illetve hogy ööö hogyan gondolja ön hogy ööö ugye nagyon sok sz nő

számára hát probléma nem probléma nem tudom hogy a karriert illetve a gyere gyermeknevelést hogyan egyezteti össze ugye

b) **Kommentár** (saját vélekedés, pl. egyetértés, csodálkozás rövid kifejezése) hangzik el, miközben az, akinél a szó van, vagy folyamatosan beszél, vagy rövid szünetet tart (pl. levegővétel, szókeresés). Abban az esetben, ha a szóátadás egyébként nem történik meg, tehát a rövid kommentár közben/után az eredeti beszélő folytatja a mondandóját (jellemzően ott, ahol abbahagyta), nem vettünk fel külön beszédlépést. Számított ugyanakkor a kommentár hossza, tehát ezek a megjegyzések egy tagmondatnál nem lehettek hosszabbak, ellenkező esetben új beszédlépésnek számítottuk őket. A legjellemzőbb kommentárok az egyszavas minősítések voltak, főként az egyetértést kifejező *persze; ja; igen,* illetve az olyan **érzelemnyilvánítások,** mint a *jaj!* vagy az *istenem!*. Néhány példa:

(4) (A: a kórussal mennénk igen) (T1: **azt a mindenit**) A: ööö de hát az egy jó kéthetes út repülővel tehát egy kicsit húzósabb mint egy ilyen ööö közelebbi tengerparti nyaralás

(5) A: hát nálunk általában családi körbe szoktuk ezt ünnepelni tehát meghívunk ! szűk rokonságot is ! hát ugye akkor ! mindenféle finomságokat csinálunk addigra ! sonka ugye ööö tojás és egyéb dolgok! gyerekek már nemm olyan kicsik de azér még mindig ööö most kertes házba lakunk és el szoktunk dugni ilyen ! (T1: **hú de jó**) (T2: **az olyan jó**) A: ajándékokat és akkor végigpásztázzák a kertet és akkor nagy örömmel megállapítják hogy megin megjött! a nyuszi

c) Ugyancsak nem képeznek új beszédlépést az **ismétlések,** amikor a beszédpartner megismétli a beszélő mondandójának egy részét, kifejezve figyelmét és/vagy egyetértését, például:

(6) A: (hháát) azelőtt az öregek ööö! ugye! mondták ezt a! aa probemát hogy ettő függetleü bármilyen diplomád van egy szakmát tanuljál ki (fiam!) T1: (igen igen) T2: (mhm) A: és ez ez a szemlélet ööö szabályosan a negyven év alatt (eltűnt) T2: (**el is tűnt**) T1: (igen) A: (eltűnt ! azelőtt!)

d) **A másik gondolatának befejezése/kitalálása** is okozhat egyszerre beszélést, ugyanakkor a beszédpartner mégsem veszi át a szót, például:

(7) T1: hát igen ez is olyan nehéz hogy igazából amikor már valakii hosszabb ideje együtt van valakivel hogy hogy mikor jön el el az a pont amikor már együtt is (karácsonyoznak) (T2: ühüm) T1: mer azér egy darabig mindig mindenki (vissz a saját) (T2: **otthon**) T1: családjával otthon és akkor maximum másnap találkoznak

e) **A szó átvételére tett sikertelen kísérlet** esetében a beszéd joga az eredeti beszélőnél marad, és nem keletkezik újabb beszédlépés (csak később, a sikeres szóátvételnél), például:

(8) A: tehát ami cirkusz ma az országban van most már ugye tizen iksz év óta hogy a a zegyik politikai erő meg a másik politikai erő kígyóbékát kiabál egymásra rondán beszélnek tóval ez ez egy olyan minősíthetetlen magatartás ott fönt ami óhatatlanul továbbgyűrűzik ennyi
T1: (**de igazából**) A1: (mindenki elégedetlen) mindennel
T1: **ühüm de igazából** van ennek egy másik aspektusa is hogy azért a nevelés is tehát így globálisan az úgy (nagyon megváltozott)

A (8)-as esetben látható, hogy T1 először nem jár sikerrel a szóátvételt illetően, majd miután végighallgatta A-t, rövid hümmögés után újból belekezd ugyanabba a megnyilatkozásba.

f) **Az előbbi esetek keveredve** is megjelenhetnek, a (9)-es példában ismétlés és egyetértő kommentár kapcsolódik össze:

(9) (A: hát kérdés hogy ki) tiltakozna meg itt nálunk nem is iigazán ilyen bevándorlókról van szó mer itt sa itt nem a bevándorlók nem dolgoznak hanem a magyar állampolgárok (nem dolgoznak tehát hogy ez egy ez) inkább belügynek (számít inkább) (T2: **belügy aha**) A: mint egy ilyen külügy bár biztosan rágyűrűzne ezekre az emberekre is

Anyag, módszer, kísérleti személyek

A vizsgálat anyagát a magyar beszélt nyelvi adatbázis (BEA, Gósy 2008) 30 darab, összesen 464 percnyi (7 óra 44 percnyi) spontán társalgása képezte. Egy beszélgetésben mindig 3 adatközlő vett részt, a fő adatközlő (A) és két beszélgetőtárs (T1, T2). A vizsgált társalgások esetében a fő adatközlők esetében a nemek aránya egyenlő volt (15 nő, 15 férfi), akik életkorukat tekintve 20 és 71 év közöttiek voltak (átlagéletkor: 42,73 év), az egyes korosztályokat egyenlő arányban képviselték. A társalgások hossza 5 perc 59 másodperc és 24 perc 34 másodperc között ingadozott, tehát meglehetősen nagy volt a szórás, az átlag 15 perc 27 másodperc volt.

A statisztikai elemzéseket az SPSS 17.0 verziójával végeztük.

Eredmények

A társalgások lejegyzése és a beszédfordulókra történő osztása után manuálisan kigyűjtöttük a szóátvételkor és -átadáskor használt diskurzusjelölőket, amelyek esetében gyakorisági vizsgálatokat végeztünk.

A **szóátvételkor** használt diskurzusjelölők példányszáma összesen 1102 darab volt, és az adatok 189-féle diskurzusjelölő-típust reprezentáltak. Ez azt jelzi, hogy nemcsak igen magas a szóátvételek során megjelenő diskurzusjelölők száma, de azt is, hogy nagyon sokfélék. Az egyes jelölők használati gyakorisága ugyanakkor élesen eltért egymástól, mivel a 189 típusból 177

(93,7%) esetében a példánygyakoriság nem haladta meg a 10 előfordulást, és mindössze egyetlenegyszer használták az adatközlők a szóátvételnél megjelenő jelölők 66,7%-át (126 db). Bár a többszavas diskurzusjelölők túlsúlya jellemezte a vizsgált korpuszt, mivel ezek tették ki a vizsgált típusok 81%-át (153-féle jelölő), és az egyszavasok csak 19%-ot (36-féle típus), a példánygyakoriság tekintetében órási volt az eltérés köztük: az előfordulások több mint felét (581 db, 52,7%) mindössze háromféle egyszavas diskurzusjelölő adta ki, a *hát*, a *de* és az *és*. Néhány példa a használatukra:

(10) A: (és hol tartjátok)
T2: **hát** benn a házban merhogy úri macskák ahogy mondád ki nem mennének ám a szabadba aludni
T1: **hát** nem hát ott hideg van
T2: (**hát** ja)
T1: (te kimennél) a sza!badba alu!dni most komolya!a!an!
T2: **há** nincs is akkora bundám
T1: **hát**! ! az egy dolog

(11) T1: azoknak e- elég bonyolult a gondozásuk (neem)
A: (! hát) igen igen
T2: **de** nem vonyítanak

(12) T1: de hogy nem is mászunk a nyakába hogy elnézést hány éves is tetszik lenni tehát ez most már így ki van adva
A: ühüm
T2: **és** ezt be is tartják

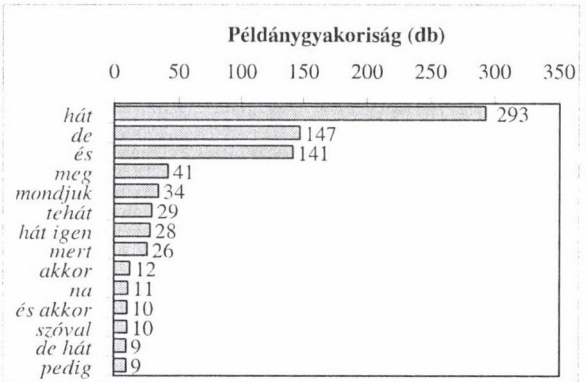
Mivel nagy számuk miatt a szóátvételkor megjelenő összes diskurzusjelölő-típust nincs módunk e tanulmány keretei között bemutatni, csak a leggyakoribb egy- és többszavas jelölőket soroljuk fel, ezek a következők:

Egyszavasok (10 vagy annál több előfordulás): *hát, de, és, meg, mondjuk, tehát, mert, na, szóval.*

Többszavasok (5 vagy annál több előfordulás): *hát igen, és akkor, de hát, na jó de, de mondjuk, jó mondjuk, na hát, meg hát, de egyébként, hát nem, jó akkor, jó hát, na de.*

Az 1. ábra a szóátvételek során leggyakrabban (9 előfordulás és afelett) használt diskurzusjelölőket mutatja. Mint látható, a *hát* kétszer gyakrabban fordult elő a második leggyakoribb *de* jelölőhöz képest, amely az *és*-sel körülbelül azonos arányban tűnt fel. A negyedik leggyakoribb jelölőt, a *meg*-et már nagy különbség választja el tőlük: a *de* és az *és* 3,5-szer gyakoribb volt nála.

Az 1. hipotézist illetően elmondható, hogy a vizsgálat megerősítette, tehát az általában gyakori diskurzusjelölők a szóátvételkor megjelenő jelölők között is a legtöbbször használtak között voltak, kivételt csak azok képeztek az előbbi csoportból, melyek jellemzően nem a megnyilatkozások (leg)elején foglalnak helyet, hanem mindig az őket tipikusan követő nyelvi egységek előtt (pl. *ilyen, így*).



1. ábra
A szóátvételkor használt diskurzusjelölők előfordulási gyakorisága

Korábban említettük, hogy bár a szóátvételkor használt jelölők a beszédfordulók elején jelennek meg, melynek magyarázatát kapcsoló szerepükben, és dominánsan kötőszói eredetükben találhatjuk meg, különböző szintaktikai pozíciókat vehetnek fel:

a) **Abszolút beszédforduló-kezdő helyzet:**

(13) T1: **pedig** az olyan jó szerintem kell mi még a mai napig szoktunk

b) **(Formálisan) egyetértést vagy egyet nem értést kifejező szó/kifejezés (*igen; így van; ja; aha; nem; nem tudom*) vagy hümmögés (*hm; ühüm*) utáni helyzet:**

(14) T1: igen **meg** ugye a médiának is nagyon nagy a szerepe

A *hát* mellett az *igen*, *ja* illetve *nem* megjelenése két változatot is mutatott, az *igen hát* és *nem hát*, illetőleg *hát igen* és *hát nem* – az utóbbiakat az előbbiektől elkülönítettük, mivel ott az igenlés jelentésének árnyalása a jelölő funkciója, és nem az *igen/nem* után induló megnyilatkozás tartalmára vonatkozik:

(15) T1: mennyire izé ilyen nnagy öööö hogy mondjam divatja volt ennek
A: **hát igen igen igen** ! s ez most egy kicsit visszaszorult

(16) T1: nemtom mennyi euró éjszakánként és akkor így mtuuk hogy tehát annyi mint egy pusz minthogyha egy plusz (embert) T2: (mhm) T1: vinnénk !! mgyhogy
A **ja hát** mi is mindenhova akarjuk vinni

Szintén a *b)* kategóriába soroltuk a nem mondatszókat vagy módosító mondatrészleteket, de igenlést, egyetértést kifejező ragozott igealakokat, például:

(17) T1: és milyen a jó tanár szerinted tehát milyen tanár akarsz t!e lenni vagy milyen példák vannak előtted
T2: azt nem mondta hogy jó tanár akar lenni
T1: ja igaz !
A: (**szeretnék**) **hát** öööö ami szerintem nagyon fontos hogy tehát öm azt mondják hogy ugye a megtanuljuk ezeket a pedagógiai marhaságokat meg pszichológiai hülyeségeket akko juj de jó lesz de nem

c) A legritkább eset az, amikor **a beszédforduló első traktusában (pl. a fókusz után) foglal helyet a jelölő** (pl. *egyébként, mondjuk, amúgy, meg, pedig*), de nem annak a legelején vagy az egyet(nem)értő kifejezés után:

(18) T2: mennyire van **egyébként** a Balaton

Az előbbi diskurzusjelölő változatos pozíciókat vehet fel: állhat a forduló legelején is, vagy a beszédfordulót nyitó megnyilatkozás végén. A *meg* és a *pedig* szintén lehet szegmenskezdő helyzetű [lásd *a)* eset], például:

(19) A: ezt nem tudom mert ha a metrón utazunk akkor azt látjuk hogy a metrókocsik üvege ömm televésve nem tudom késsel vagy (mivel ffeliratokat)
T1: (**meg** ilyen ragasztóval szokták)

A többszörös diskurzusjelölők, amelyek esetében két vagy több egy- vagy többszavas diskurzusjelölő jelenik meg egymás mellett, a megnyilatkozásbeli helyzet szempontjából nem képeznek külön esetet, mivel az első jelölő után álló további diskurzusjelölők önállóan ugyancsak a szegmens elején [*a)* eset] vagy az egyet(nem)értő kifejezés után [*b)* eset] foglalnának helyet, így ezekbe a kategóriákba sorolhatók. Például:

(20) T1: **de mondom** még az én kutyám is pedig az egy utcagyőztes

A diskurzusjelölők előfordulási gyakoriságát beszédfordulókra vetítve is megvizsgáltuk. A 30 társalgásban összesen 2540 darab beszédfordulót találtunk, ez átlagosan 5,47 darabot jelentett percenként, 1,64 és 11,28 darab beszédforduló/perc között mozgott a szórás, vagyis igen nagy volt. A társalgáshossz növekedésével jellemzően a beszédfordulószám is emelkedett, de az egyes társalgások esetében nagy különbségek mutatkoztak. Így például a bea046-os beszélgetés 23 perc 40 másodperce alatt mindössze csak 39 beszédforduló keletkezett, mivel az egyes adatközlők hosszabban tartották maguknál a szót, 19-szer a fő adatközlőnél (57 éves nő) volt beszédjog. Ezzel szemben a bea044-es társalgásban 17 perc 22 másodperc alatt 196 darab forduló született, itt a fő adatközlő (31 éves nő) 69-szer produkált saját beszédfordulót.

Ha a szóátvételkor megjelenő diskurzusjelölők mennyiségét összevetjük a beszédfordulók számával, azt látjuk, hogy az összes beszédforduló 43,38%-ában szerepelt a szóátvételkor diskurzusjelölő elem. Ez percenként átlagosan 2,34 darab ilyen jelölőt jelent, a szórás itt is ugyancsak nagy volt (0,79 db/perctől 4,31 db/percig). A fő adatközlők használta diskurzusjelölők száma összesen 416 darab volt, az összes ilyen elem 37,7%-át tették ki.

A fő adatközlők által használt diskurzusjelölők nemek szerinti megoszlását tekintve elmondható, hogy a női adatközlők átlagosan 0,78 darab jelölőt használtak társalgáspercenként, a férfiak pedig 0,96 darabot (férfiak: 0,31–2,06 jelölő/társalgásperc; nők: 0,27–1,2 db jelölő/társalgásperc), a különbség nem szignifikáns [$t(28) = 1{,}316$; $p = 0{,}199$).

A 2. hipotézis, amely a diskurzusjelölők funkciójára vonatkozott, az elsőhöz hasonlóan megerősítést nyert, ugyanis a vizsgált jelölők egyike sem szolgált pusztán a szó átvételére, a két fő diskurzusjelölői szerepkörnek megfelelően az előzményekhez, illetve a kontextus valamely eleméhez való kapcsolást fejezték ki, illetve attitűdjelölést. A kapcsoló funkció természetszerűleg adódik a szóátvétel mint funkció mellé, hiszen a kezdő pozícióban lévő diskurzusjelölő célszerűen számos információt hordoz: hogyan integrálják a hallgatók a következő megnyilatkozás jelentését a diskurzus során eddig elhangzott információk sorába; hogyan viszonyul a követő megnyilatkozás az előzőhöz; milyen a beszélői viszonyulás az elhangzó megnyilatkozásokkal kapcsolatban.

A többszörös diskurzusjelölők esetében a kép még differenciáltabb lehet, jellemzően az első elem a visszakapcsolásért és a folytonosságért felelős, a második pedig az általa felvezetett megnyilatkozás tartalmával kapcsolatban közöl valamit, például:

(21) T1: (**meg hát** ezek a kárpitok) ez! gye

Az előbbinek a fordítottja is lehetséges, a kiegészítést (*hozzá kell tenni*) felvezető *hát* a beszélő attitűdjét fejezi ki, ami mind a követő többszavas jelölőre, mind az azután következőkre is vonatkozhat, ugyanakkor a *hát* az előzményekhez való kapcsolódást is biztosíthatja:

(22) T2: hát nem de állítólag ennek az újnak ennek az új S***- S***nek egy kicsit már megerősítették a (karosszériáját) (T1: ühüm) T2: amennyire lehetett hát nem tudom (T1: ühüm)
A: **hát hozzá kell tenni** nekem van egy ismerősöm aki ott dolgozik a S*** gyárba ööö tolmácsként és ő mondta hogy soha ne jusson eszünkbe hogy S***t vegyünk

A harmadik eset az, amikor több jelölő ugyanazt a funkciót fejezi ki:

(23) T1: **de viszont** ebbe az a poén hogy ha ha meg ööö túl nagy nálad a bevétel akkor a többiek akik ugyanígy csalnak tehát mondják hogy akkor csak háromszáz akkor ők meg meglincselnek

A *de* diskurzusjelölő esetében egyfajta változásnak lehetünk a tanúi: ez az egység Bańczerowski (2005: 488–489) szerint új replikát vezethet be, vagy a replikán belül új információt, melynek során általában szembenállás, tiltakozást, tagadást fejez ki. Egy példa erre a korpuszból:

(24) T1: de hát mondjuk mindenkinek az iskolai végzettségének vagy hát a aa hogy mondjam tudásának megfelelően
A: **de** azt hiszem hogy most itt csak ilyen utcatakarításról van szó nem

Ugyanakkor a *de* a spontán társalgásokban már nem mindig hordoz ellentétes jelentést, van, hogy **pusztán csak kiemeli a bevezetett témát, hangsúlyosnak szánt új szempontot emel be tényleges ellentét nélkül** (vö. Bańczerowski uo.). A következő példában a beszélgető felek egyetértenek abban, hogy nem jó megoldás a bliccelők miatti sűrű BKV-s ellenőrzés, a kiemelt *de* helyében nyugodtan állhatna az *és* vagy az *és... is* is, mivel az A közlése annyi, hogy 'és még csak nem is hatékony az ellenőrzés':

(25) T2: de ez is nonszensz hogy egyszerűen nem kísérheted ki a hozzátartozódat a vonathoz hát mi ez rendőrállam komolyan mondom más máshol ilyen nincsen tehát most jó persze bliccelünk meg minden de ezt akkor máshogy próbáljuk már meg megoldani nem ilyen A **de:** T2 nem ilyen diktatorikus A: **de** T2 eszközökkel
A: **de** ez ezzel ezzel ezzel hogy ennyi a bliccelés és külön foglalkoztatni kell erre embereket ezzel automatikusan aláírták hogy hogy a kalauzok nem végzik a dolgukat

A **szóátadás** során használt diskurzusjelölőkből nagyságrendekkel kevesebb fordult elő a vizsgált anyagban, mint szóátvételekkor alkalmazott jelölőkből: összesen 44 előfordulást és 14-féle típust regisztráltunk, az utóbbiak a következők voltak (gyakorisági sorrendben): *úgyhogy, tehát, vagy, nem?, de hát, hát nem?, de, vagy hogy, vagy így hogy?, mondjad, tehát így, hát de, tehát hogy, úgyhogy hát nem tudom*. A korpuszt az *úgyhogy* fölénye jellemezte (23 darab), egy példa a használatára:

(26) T1: ühüm nekem még megvan az a születési anyakönyvi kivonatomba van így (belerakva igen) **úgyhogy**
A: (igen nahát) tehát ez olyan divat hogy az első fogat meg a (kis hajfürtöt őrizgeti)

Az *úgyhogy* alkalmazása ebben a szerepben ugyanakkor csak az adatközlők egy részénél (16 főnél a 90-ből, 17,7%) volt megfigyelhető, és közülük is a többség csak egyszer alkalmazta (16-ból 12 fő), tehát egyéni nyelvhasználati sajátosságnak tekinthető.

A második leggyakoribb diskurzusjelölő a *tehát* (5 darab) volt, amit a *vagy* (3 darab) követett, a többi jelölő csak egyszer vagy kétszer szerepelt. Ez tehát azt jelenti, hogy a szó átadásakor a beszélők minimálisan élnek a diskurzusjelölőkkel.

A 3. hipotézis az első kettővel ellentétben nem bizonyult tarthatónak, mivel a szóátadáskor alkalmazott jelölők többsége ugyanúgy kötőszói eredetű volt, mint a szóátvételek során használtak, más szófajra visszamenő jelölők (pl. *mondjad*) pedig elvétve tűntek fel. Kötőszói eredetű jelölők használata szóátadáskor első pillantásra logikátlannak tűnik, ugyanakkor valószínűleg éppen azért teszik lehetővé a szóátadást, mert a félbehagyottnak tűnő szerkezet folytatást kíván, és így a beszédpartnert megszólalásra készteti. Emellett akár egyértelmű is lehet, hogyan fejeződne be a megnyilatkozó közlése, s így a beszédpartner úgy érzi, hogy ezután övé a szó.

Mindez arra utal, hogy a diskurzusjelölők többsége azért a beszédfordulók elején jelenik meg, mert jellemző funkcióik (kapcsolás, visszautalás, attitűdkifejezés) ezt kívánják meg, ti. azokat a közlés elején kell vagy célszerű kifejezni. A beszédforduló végén ugyanis ilyen rövid jelölőkkel már nem lehetne egyértelműen visszautalni az előzményekre (legfeljebb hosszabb kifejtés révén), így feltételezhetően az utolsó lehetséges megjelenési helyük a fordulókezdő megnyilatkozás vége. Ennek tisztázása azonban további vizsgálatot kíván meg.

Elmondható tehát, hogy a diskurzusjelölők meghatározó pragmatikai elemnek tekinthetők szóátvételkor, szóátadáskor azonban nem. Funkciójuk nem pusztán a szóátvétel jelölése, hanem a korábban történtekhez vagy elhangzottakhoz való kapcsolódás, illetve a beszélői attitűd jelölése, így tisztán szóátvevő szerepben használt diskurzusjelölő valószínűsíthetően nincs.

Irodalom

Bańczerowski, Janusz 2005. A kötőszók mint metaszövegbeli operátorok. *Magyar Nyelvőr* 129. 488–495.

Bata Sarolta 2009. Beszélőváltások a beszédpartnerek személyes kapcsolatának függvényében. *Beszédkutatás 2009.* 107–120.

Dér Csilla Ilona 2010. „Töltelékelem" vagy új nyelvi változó? A *hát, úgyhogy, így* és *ilyen* újabb funkciójáról a spontán beszédben. *Beszédkutatás 2010.* 159–170.

Fraser, Bruce 1999. What are discourse markers? *Journal of Pragmatics* 91. 931–952.

Iványi Zsuzsanna 2001. A nyelvészeti konverzációelemzés. *Magyar Nyelvőr* 125. 74–93.

Gósy Mária 2008. Magyar spontánbeszéd-adatbázis – BEA. *Beszédkutatás 2008.* 194–207.

Lerch Ágnes 2011. A szintaxis, az intonáció és a pragmatika szerepe a beszélőváltás megvalósulásában magyar nyelvű konverzációkban. Előadás. *Grammatika és kontextus: új szempontok az uráli nyelvek kutatásában III.* Budapest, 2011. április 21.

Markó Alexandra 2006. Beszélőváltások a társalgásban. http://fonetika.nytud.hu/letolt/ma_2.pdf (Letöltve: 2011. október 1.)

Németh Zsuzsanna 2007–2008. A forduló (beszédlépés) kiterjesztésének grammatikája a magyarban. *Nyelvtudomány* III – IV. 149–184.

Sacks, Harvey – Schlegloff, Emanuel A. – Jefferson, Gail 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50. 696–735.

Stephens, Jane – Beattie, Geoffrey 1986. On judging the ends of speaker turns in conversation. *Journal of Language and Social Psychology* 5/2. 119–134.

A kutatás a Bolyai János kutatási ösztöndíj támogatásával készült.

A *HÁT* DISKURZUSJELÖLŐ A TELEVÍZIÓS VITAMŰSOROKBAN

Schirm Anita

Bevezetés

A diskurzusjelölők a nyelvi elemeknek egy nehezen meghatározható körét alkotják, a kutatók (pl. Dér 2010) azonban abban többnyire egyetértenek, hogy a diskurzusjelölők diskurzusszegmenseket kötnek össze és pragmatikai viszonyokat jelölnek. Az egyes diskurzusjelölők írott és beszélt nyelvi megvalósulásai között azonban különbség van, az élőbeszédben – annak spontaneitása és kevésbé tervezettsége miatt – ugyanis gyakrabban és többféle szerepkörben fordulnak elő ezek az elemek. A különféle beszélt nyelvi műfajok közül tanulmányomban a televíziós vitaműsorokban megjelenő diskurzusjelölő-előfordulások közül a *hát* szerepköreit vizsgálom. Ez ugyanis a magyarban az egyik leggyakoribb diskurzusjelölő (Dér–Markó 2007), amelyhez a beszédtervezési funkción kívül számos szerepkör tartozik. Korábbi kutatásaim során (pl. Schirm 2007–2008) a *hát* grammatikalizálódásával és mai funkcióival zömében írott anyagok elemzésével foglalkoztam, ezért kíváncsi voltam arra, hogy az akkor elhatárolt szerepkörökhöz képest a beszélt nyelvi anyag bevonása a vizsgálatba milyen új eredményeket hozhat. A kiinduló hipotézisem az volt, hogy a *hát* a televíziós vitaműsorokban a műfaj jellegzetességéből adódóan az arculatmunka szolgálatában áll.

Anyag, módszer, kísérleti személyek

A *hát* diskurzusjelölő funkcióinak a feltérképezéséhez televíziós vitaműsorokat használtam fel, s mindig az egyedi megjelenésekből kiindulva, adatközpontúan vizsgálódtam. A funkciók megadása azonban korántsem problémamentes. A diskurzusjelölőknek tulajdonított textuális funkciót a legkönnyebb megragadni, ám az interakciós és attitűdjelölő funkciók egzakt leírása már nehezebb. A funkciók vizsgálatakor azonban nem szabad elfelejteni, hogy a „funkció nem merev kategória, [...] hanem folyamat jellegű jelenség, amely a beszélői vagy hallgatói feldolgozás révén jön létre" (Tolcsvai 2005: 350). A szerepmegjelöléseknél tehát mindig figyelembe vettem a kontextuális tényezőket és a vizsgált anyag jellegzetességeit is. Az egymással néha ellentétes funkciók motiváltságának alapjaira és azok magyarázatára jelen tanulmányban nem térek ki, mivel ezt más munkában már megtettem (vö. Schirm 2011).

A felhasznált korpusz 3 órányi hangzó szövegből, 4 vitaműsorból állt: az ATV-n 2002-ben sugárzott *Pro és kontra* című műsor Médiaegyensúly részét és a *Közhang* című műsort elemeztem, illetve az MTV *A szólás szabadsága* című műsorának két 2010-es májusi adását. Szándékosan úgy választottam ki a vizsgált műsorokat, hogy egymástól kissé eltérő szerkezetű anyagok kerüljenek bele, a feltételezésem ugyanis az volt, hogy a műsorszerkezet is befolyásolja a *hát* elemnek az arculatmunkában betöltött szerepét. Nézzük röviden az egyes műsorok forgatókönyvét!

A *Pro és kontra* az ATV-n hetente jelentkező félórás vitaműsor volt, Elek István vezetésével. A műsorba mindig két eltérő álláspontot képviselő híres embert (politikust, újságírót, történészt) hívtak meg. A vitaműsor adásai közül a 2002. szeptember 5-i adást választottam elemzésem tárgyául, melyben Kerényi Imre és Kocsi Ilona vitatkoztak a médiaegyensúly kérdéséről.

A *Közhang* az ATV-n hetente jelentkező 55 perces vitaműsor volt, melyet Juszt László vezetett. A stúdióba meghívott közemberek (általában négy fő) beszélgetése volt a magánélet és a közügyek számukra fontos jelenségeiről. Interaktív műsor SMS-küldéssel, telefonos megszólalási lehetőséggel. A diskurzusjelölők és az arculatmunka vizsgálatára különösen alkalmas műsortípus, hiszen azon túl, hogy a 2-2 vitapartner ellentétes álláspontot képviselt, a betelefonálók is alakították megszólalásaikkal és SMS-eikkel a műsort.

A harmadikfajta vitaműsor az MTV-n hetente látható *A szólás szabadsága* című közéleti magazinműsor volt, melynek Krizsó Szilvia a műsorvezetője. A 2010-es májusi adások közül a 17-én sugárzott *A távozó kormányfő összegez – interjú Bajnai Gordonnal* című részt, illetve a május 31-én bemutatott *Megalakult az új kormány, megvan az első konfliktus* című műsort elemeztem. A két eddig bemutatott politikai vitaműsortól abban különbözik *A szólás szabadsága*, hogy itt mérsékelten konfrontatív légkörben zajlik a beszélgetés, s a műsorvezető egyszerre mindig csak egy vendéggel beszélget, így a kizárólagos kérdezői és irányító szerepéből adódóan a kialakuló verbális konfliktus sem olyan heves. Verbális konfliktuson Vuchinich (1990: 118) alapján a két vagy több résztvevő között kialakuló konfrontációt tartalmazó szóbeli eseményt értem. Továbbá *A szólás szabadságá*nak magazinműsor jellegét biztosítja az is, hogy a beszélgetések előtt mindig egy összefoglaló bejátszás látható a téma előzményeiről.

A vitaműsor a szemtől szembe interakció tipikus példája, s ez a műsortípus a hírműsoroktól eltérően egyszerre közöl információt és egyszerre szórakoztat is, vagyis alapvetően hallgatóorientált. A vitaműsoroknak általában többszörös hallgatóságuk van: egyrészt a közvetlenül megcímzett hallgatók, akik közelről érintettek a témában, másrészt a stúdióban lévő nézők, harmadrészt pedig azok a tévénézők, akik (esetleg csak véletlenül) bekapcsolták a készüléket.

A választott műsorok a média ún. félintézményes diskurzusai (Ilie 1999; 2001) közé tartoznak. Ez azt jelenti, hogy vannak bizonyos külső elvárások és szabályok, amelyeket be kell tartani a társalgás során, de a kontroll lénye-

gesen alacsonyabb, mint a teljesen intézményes diskurzusok esetén, ám magasabb, mint egy teljesen kötetlen beszélgetés során.

Ezeknek a műsoroknak kétféle szereplőtípusa van: a moderátor, azaz a műsorvezető, aki irányítja a vitát; és a vitapartnerek. A beszélgetés szabályai a félintézményes vitaműsorok esetén bizonyos fokig szerephez kötöttek, ugyanis vannak elvárások arra vonatkozóan, hogy ki és milyen jellegű megnyilatkozásokat tehet (vö. Schirm 2007: 133–134), vagyis megkülönböztethetjük a szerephez illő, illetve a szerephez nem illő megnyilatkozásokat. A moderátor kérdését csak az korlátozza, hogy ne térjen el az újságírói semlegességtől (Clayman 1992: 163), a vitapartnerekre pedig az a megszorítás érvényes, hogy nem tehetnek fel a műsorvezetőnek személyes kérdéseket, és nem kérdőjelezhetik meg annak a kérdezéshez való jogát és kérdésének a relevanciáját sem. A műsor intézményes jellemzőjéből adódó szabályai mellett a társalgási jelleg pedig abban nyilvánul meg, hogy a témaváltás, a kérdés feltevése és a fordulók hosszának a szabályozása nem kizárólag a műsorvezető joga, hanem bármely résztvevőé.

A *hát* diskurzusjelölő nem mutatott kizárólagos szerephez kötöttséget, ugyanis mind a műsorvezető, mind pedig a vitapartnerek használták ezt az elemet a megnyilatkozásaikban, a *hát* jellegzetes válaszjelölői funkciójából adódóan azonban sokkal gyakrabban fordult elő a vitapartnerek megnyilatkozásaiban, ahogy azt a következő fejezet példái is mutatják majd.

Eredmények

A korpuszban a *hát* elem összesen 61-szer jelent meg, s ezek az előfordulások a korábban az írott szövegekben megfigyelt használati körökhöz nagyon hasonlóan rendeződtek, vagyis ezek az adatok is a *hát* gazdag jelentéshálózatáról tettek tanúságot. A politikai vitaműsorok jellegzetes egységei a kérdés-válasz szomszédsági párok, így nem meglepő, hogy a *hát* gyakran válaszjelölői szerepben jelent meg. Ilyenkor a beszélő a saját véleményének a kifejtése előtt használta az elemet, ahogy azt az alábbi [(1) és (2)], a *Közhang* című műsorból származó példák is mutatják:

(1) C: Illetve nagyon furcsállom, hogy a hölgynek a kokárdáról, erről az áprilisi kokárdaviselésről az jutott eszébe, hogy itten Trianon revízióját akarják, nem is tudom, hogy akkor
R: Ugye nem alaptalanul jut ez az embereknek az eszébe?
C: **Hát** szerintem teljes mértékben alaptalanul jutott eszébe.

(2) A: De utána ami most volt, a Parlament előtt, arra viszont elmentem
B: De azt nem rendezte senki, az egy spontán megmozdulás volt.
A: Ugye ezt te most, ezt el is hiszed?
B: **Hát** asszem a rendőrség rendezte, de erre még rátérünk.

Az idézett két példában a kérdező saját állásfoglalást kért, s a válaszadó a mondanivaló szubjektivitását a diskurzusjelölő utáni elemekkel (*szerintem,*

asszem) jelezte, maga a *hát* pedig textuális szereppel bír, jelöli a megelőző kérdéshez való kapcsolódást, miközben az utána következő részre irányítja a figyelmet. A megnyilatkozások hangsúlya egyértelműen jelzi, hogy a példákban az elem nem megakadásjelenségként áll, és nem is a hezitációt jelöli, hanem nyomatékosítja a mondanivalót.

A válaszjelölői szerep egy másik alfunkciója is megjelent a korpuszbeli példák közt, amikor a *hát* azt jelölte, hogy a megkérdezett egy hosszabb válaszba, monológba vagy történetbe kezdett bele (vö. Dér 2010: 162). Ezt példázza a következő (3), *Közhang*beli részlet is:

(3) D: Illetve a kérdésem pontosan az Orbán beszédre irányult, nem is a tüntetés tényére, mer ez rendbe volt, a jobboldal ezt a sajtószabadság melletti tüntetésnek írja, az én olvasatomban ez az önkormányzati kampánynyitója volt a Fidesznek és a polgári köröknek és én kimondottan a kérdésem az Orbán beszédre irányul. Hogy erről mi a véleményetek, hogy ítélitek meg, mennyire tudtok vele azonosulni? És hogyha nem, akkor mi az, amiben nem?
C: **Hát** ezt próbáltam elmondani, hogy a miniszterelnök úrnak az volt a beszédének az volt a lényege, hogy egy olyan közszolgálati televízióra van szükség, ahol a baloldali értékrend és a jobboldali értékrend megfelelő egyensúlyban van. Tehát legalább a választási eredményeknek megfelelő egyensúlyt tud alkotni, ami azt jelenti, hogy 51-49, tehát a baloldal 51%-ot kapott áprilisban, a jobboldal 49-et.

A kérdésre adott választ indító *hát* sokszor a hezitálást és az időnyerés szándékát fejezi ki, ahogy azt a (4) példában is láthatjuk:

(4) A: De mitől fél?
B: **Hát**, sajnos Magyarországon sokaknak félniük van kell.

Ilyenkor a *hát* után jellemző a szünettartás és gyakran egyéb hezitálójelölő elemek (pl: *ööö*) is kísérik még a diskurzusjelölőt. Ez a fajta használati kör nem csupán kérdések után volt általános, ahogy azt az alábbi, (5) példa is illusztrálja:

(5) B: Menjünk tovább!
A: Menjünk, ezen lépjünk túl, mert mind a két oldalról mondom lehetne hozni ilyen-olyan példákat és **hát** ööö nem tudom.

A válasz szubjektivitásán, illetve a hezitálás kifejezésén túl a *hát*-ot a vitaműsorokban annak kifejezésére is használták, amikor a megkérdezett a véleményét mintegy tényként, evidenciaként mutatja be. Ilyenkor a diskurzusjelölő erős hangsúlya jelzi, hogy a beszélő az állítás nyomósítására használja az elemet, ahogy azt *A szólás szabadságá*ból (6) és a *Közhang*ból (7) származó példák is mutatják.

(6) R: De mondjuk a Malévet feltüntették, mint kockázati tényező? Hiszen a napokban Szijjártó Péter úgy fogalmazott, hogy bármikor hullhatnak még ki csontvázak a szekrényből. Ő például felhozta egyik példaként a Malévet.
B: **Hát** a Malév-tranzakcióról nagyon világos tájékoztatást adtunk, amikor az megtörtént, bekövetkezett ebben az évben.

(7) A: És mitől fél?
B: **Hát** nem fogja nyilvánosan kiadni magát!

Mindkét idézett példában a *hát* diskurzusjelölő azt fejezi ki, hogy a megkérdezett a választ magától értetődőnek tartja.

A korpusz a *hát* általános válaszjelölői szerepén túl sokféle egyéb használati módra is példát szolgáltatott. Az anyag tanúsága szerint a *hát* a televíziós vitaműsorokban jelezheti még a mondanivaló továbbvitelét, megakadásjelző elemként és önjavító szerepben is állhat, illetve lehet a magyarázkodás eszköze is. Állhat a 'vajon' jelentésben, kódolhatja az 'igen' értelmet, és következtető kötőszóként, 'tehát' értelemben is használatos. A beszélői attitűdök közül felfokozott érzelmi állapotot is képes jelölni: felindultságot, kételkedést, méltatlankodást és csodálkozást, erősítheti a kérdés retorikusságát, emiatt pedig arculatromboló elemként is funkcionálhat. Ez utóbbi aspektusra a későbbiekben még részletesen is ki fogok térni.

Nem csupán az elem funkcióját vizsgáltam meg, hanem a megnyilatkozásban betöltött helyét is. A pozíciók szerinti előfordulások arányai hasonlóak Dér Csilla Ilonának a BEA adatbázisból származó 4 órányi korpuszán elvégzett 2010-es kutatásának eredményeivel. Vagyis a pozíciók közül nálam is a leggyakoribb a mondatot, illetve tagmondatot megelőző használat volt. Ettől eltérő pozíciójú megjelenésre csupán egyetlen példa adódott a korpuszban: amikor a *hát* önmagában állt, s így többféleképpen volt értelmezhető. A mondat, illetve tagmondat előtti megjelenések pedig az alábbi, 1. ábrán látható arányban voltak jelen a mintában: a tagmondatot indító *hát*-ra 45 adat volt, a kötőszó utánira 12, míg más diskurzusjelölő után 3-szor fordult elő a *hát*.

1. ábra
A *hát* mondat, illetve tagmondat előtti előfordulásainak megoszlása

A funkciók és a pozíciók a korábban elvégzett vizsgálatokhoz (pl. Dér 2010; Schirm 2011) képest nem igazán szolgáltak új eredménnyel, azonban a beszélt nyelvi korpusz bevonása a vizsgálatba megerősítette egy korábbi hipotézisemet. Az 1660-as sárospataki hitvita elemzése kapcsán felvetettem, hogy a *hát* elem használata a hitvitában az udvariassággal, illetve az arculatmunkával is kapcsolatba hozható.

Az *arculat* (ang. *face*) nyelvészeti terminusként Goffmantól ered, s egy olyan tulajdonságegyüttest jelöl, amellyel mindenki rendelkezik. „Olyan énkép, amelyet a társadalmilag megerősített tulajdonságok körvonalaznak" (Goffman 1967/1995: 3). A beszélgetésben a résztvevők általában törekednek az arculatfenyegetés minimalizálására és saját arculatuk rendbe hozására, s legtöbbször olyan stratégiát alkalmaznak, amellyel el tudják kerülni az arculatfenyegető aktusokat, ideális esetben pedig arra is ügyelnek, hogy a beszélgetőpartnerük arculata se sérüljön az interakció során.

A sárospataki hitvitában a térítő pápisták az egész vita során nyíltan kifejezték a reformátusokkal szembeni ellentétüket, és ezt az attitűdöt a diskurzusjelölő elemekkel is erősítették. A térítő pápisták kérdései már a vita legelejétől kezdve erősen arculatrombolók voltak, s ezt a *hát* szócska is jelezte, amelynek használata a szerepekhez igazodott. Csupán a térítők használták kérdéseikben ezt az elemet, a térítendők nem, pedig ők is sokat kérdeztek a hitvita során. A *hát* elem a vitában nyomatékosító szerepéből adódóan képes volt az ellentét jelzésére, s a kérdés támadó jellegének a felerősítésére is (Schirm 2011: 37–38).

Azonban a sárospataki hitvita bármennyire is a beszédet szó szerint rögzítő, hitelesen adatolt dokumentum, mégiscsak történeti forrás, ezért a rá vonatkozó megállapítások hipotetikusak. Viszont napjaink vitaműsoraiban a *hát* adatainak egy része teljesen hasonlóan viselkedik, így a sárospataki hitvita diskurzusjelölőire vonatkozó megállapításokat a vitaműsorok elemzései igazolhatják. Így egyrészt a szinkrón nyelvhasználat segít a diakrón adatok értelmezésében, másrészt pedig az írott források adatairól tett állítások a beszélt nyelvi szövegek segítségével erősödnek.

Az általam vizsgált félintézményes vitaműsorok a sárospataki hitvitához hasonlóan arculatfenyegető aktusokat tartalmaztak. A vita lefolyása témától függetlenül általában azonos alapséma szerint zajlik, amihez meghatározott arculatmunka is köthető. A beszélgetést fordulókra bontva Muntigl és Turnbull (1998: 227) a vita szerkezetét a következőképpen adták meg: Tegyük fel, hogy van két beszélőnk, A és B. A beszélő F1-ben tesz egy kijelentést, amit a B beszélő az F2-ben vitat, mire A az F3-ban vitatkozik B-nek az F2-es kijelentésével, vagy úgy, hogy alátámasztja az eredeti F1-es kijelentését, vagy pedig direkt módon kifejti az F2-vel való egyet nem értését. Az arculatmunkát vizsgálva pedig a kutatók azt találták, hogy a verbális konfliktusban B beszélő F2-es megnyilatkozása rongálja A beszélő arcát. Erre A egy

arculatvisszaállító lépést tesz: F2 agresszivitásának a függvényében vagy F1-et védelmezi, vagy F2-t támadja.

Mivel az érvelés és a vitatkozás szükségszerű velejárója a politikai vitaműsoroknak, ezért nem meglepő, hogy az arculatmunka fontos szerephez jut bennük. A több résztvevős beszélgetésekben (*Pro és kontra, Közhang*) a moderátor feladata a vita elindítása, a vitapartnerek eltérő véleményének az előhívása, továbbá neki kell ügyelnie arra is, nehogy túlságosan eldurvuljon a vita. Szerepéből adódóan tehát felügyelnie és irányítania kell a vitapartnerek arculatromboló és arculatépítő tevékenységeit, úgy, hogy közben meg kell próbálnia semlegesnek maradni. A vitapartnerek tehát nem mondhatnak bármit és bárhogy, hiszen kontrollálják a megnyilatkozásaikat. Az általam vizsgált vitaműsorokban ezért a vitatkozó felek gyakran éltek azzal a stratégiával, hogy nem nyíltan, állítás formájában fejtették ki az ellentétes álláspontjukat, hanem szónoki kérdésbe burkolva közölték a véleményüket. Tipikusan a vitaparnerek érvelő monológjában fordult elő, s használatukkal a beszélők egyrészt óvták a saját arculatukat, másrészt pedig a megnyilatkozásaik nagyobb figyelemfelhívó erővel bírtak, hiszen a szónoki kérdések alakzatként rendkívül hatásosak. A beszélők ilyenkor kissé mást mondtak, mint amit gondoltak, ám a beszélői attitűdjüket diskurzusjelölő elemekkel, többek között a *hát*-tal tették nyilvánvalóvá. A szónoki kérdésben megjelenő *hát* diskurzusjelölő viselkedését szemléletesen mutatja az alábbi, a *Pro és kontrá*ból származó részlet (idézi Schirm 2008: 293–294).

(8) A: És szerintem maga is érti, hogy mi miről beszélünk, és maga segítsen nekünk, és mi okosak szépen bújjunk össze, nem így, a televízióban, meg a Parlamentben, hatan-nyolcan; hívjon meg embereket, a Hankiss tanár urat hívja meg, a Verebes Istvánt hívja meg. Egy csomó ember a ballibből érti ám, hogy mi mit beszélünk. **Hát** milyen alapon lehet megmagyarázni a magyarországi médiahelyzetet? Melyik isten hatalmazta fel a szoclib tábort, hogy ez így legyen? Mi ezt vitatjuk, és kérjük a maguk segítségét, és ezért ki kell mennünk az utcára. És akkor engem nem zavar, hogy a magyar ATV-n a Fási Ádám hetvenöt darab műsort csinált egyfolytában MSZP-s polgármesterjelöltekkel és képviselőjelöltekkel, mer' akkor majd a mi képviselőjelöltünk, őket bemutatjuk ott. De **hát** így hogy lehet? Hogy az MSZP-s jelöltek hetvenöt műsorban, a mi jelölteink meg sehol. Melyik isten hatalmazta fel magukat erre?

Az idézett példa négy szónoki kérdést is tartalmaz, s mindegyik egy ellentétes értelmű állításnak feleltethető meg ('Semmilyen alapon nem lehet megmagyarázni a magyarországi médiahelyzetet', 'Semelyik isten nem hatalmazta fel a szoclib tábort', 'Így sehogy sem lehet', illetve 'Semelyik isten nem hatalmazta fel magukat erre'). A beszélő egyik feltett kérdésre sem várt választ, hanem hatáskeltő erejük miatt használta őket, hiszen a kérdések el-

hangzásakor a beszélgetőpartnerek fejében megképződött egy mentális válasz, amely valójában nem az ő feleletük volt, hanem az a tartalom, amit a kérdező sugallni, illetve bizonyítani akart. Ha a két, *hát*-ot tartalmazó kérdésből elhagyjuk a diskurzusjelölőt, azok továbbra is szónokiak maradnak, de a *hát* alkalmazása az elem nyomatékosító szerepéből adódóan felerősíti a retorikusságukat, s közben jelzi a beszélői attitűdöt. Az idézet tágabb kontextusát és az intonációt is figyelembe véve egyértelműen megállapítható, hogy a *hát* ellenvetést, erős felindultságot és méltatlankodást fejez ki, vagyis arculatromboló elem.

A másik jellegzetes kérdéstípus a korpuszban a vitázó kérdés volt, amelyben szintén megjelent a *hát* elem. Erre a *Közhang* című műsorból következzen egy példa (idézi Schirm 2008: 294).

(9) A: Az az igazság, hogy mind a két oldalról lehetne mindenfajta példát felhozni, tehát a jobboldal részéről is lehetne ugyanilyen példákat felhozni.
B: Konkrétan gyerektáborról?
A: Most nem gyerektáborról beszéltem; lehetne felhozni példákat.
B: **Hát** akkor honnan?

A *Hát akkor honnan?* kérdés vitázó kérdésnek tekinthető, mivel választ is vár a feltett kérdésre a beszélő. Ugyanakkor a szónoki kérdéshez hasonlóan ebben a kérdésben is benne rejlik az ellentétes értelmű állítás ('Sehonnan'). Az A beszélő ugyanis az előzőleg elhangzott megnyilatkozás igazságértékét kérdőjelezi meg. Nem információt kér, hanem nyíltan vitatkozik B álláspontjával, s a kérdés hangsúlyából kiderül, hogy a *hát*-ot annak a nyomatékos kifejezésére használja, hogy szerinte semmilyen példát nem lehet felhozni, vagyis a diskurzusjelölő az ellenvetést erősíti, és egyben retorizálja is a kérdést. Azaz a *hát* itt is arculatromboló eszköz.

A televíziós vitaműsorokban a *hát* diskurzusjelölő gyakran egymással ellentétes funkciókban jelent meg, ezért a hozzá kötődő arculatmunka sem kezelhető egységesen. A korpusz a diskurzusjelölő általi arculatfenyegetésre és az arculatvédésre is szolgált példával. Az időhúzást, hezitálást és a beszélő bizonytalanságát jelölő *hát* arculatvédő mechanizmusként értékelhető [(4) és (5) példa], míg a szónoki és vitázó kérdésekben előforduló *hát* [(8) és (9) példa] arculatromboló elemként funkcionál.

Következtetések

A tanulmányban a *hát* diskurzusjelölő funkcióit félintézményes televízós vitaműsorok adatai alapján elemeztem. A korpusz vizsgálata azt mutatta, hogy e beszélt nyelvi műfajban igen sokféle szerepkörben jelenik meg a *hát*. Az általános válaszjelölői funkción kívül az elem képes a mondanivaló szubjektivitását, bizonytalanságát, illetve a hezitálást is jelölni, de ennek ellentétére, az evidenciaként való értelmezésre is szolgáltatott példát a korpusz. Az

elemzés a *hát* funkciógazdagságának a megmutatásán túl azt is igazolta, hogy erősen konfrontatív környezetben a kérdések közül a *hát* főként szónoki kérdésekben és vitázó kérdésekben fordul elő, s szerepe az interpretációs mechanizmus segítése: erősíti a kérdésnek az állításként való értelmezését, ezáltal jelzi a diskurzusszegmens és a beszélő közti viszonyt, vagyis mutatja a beszélő attitűdjét.

A vitaműsorok elemzése pedig nem csupán a mai kommunikációs stratégiákról szolgált információkkal, hanem a szinkrón beszélt nyelv vizsgálata megerősítette a diakrón adatokról tett, korábbi hipotetikus állításainkat is, miszerint a *hát* elem a vitabeli kérdésekben arculatfenyegető eszközként funkcionál.

Irodalom

Clayman, Steven E. 1992. Footing in the achievement of neutrality: the case of news-interview discourse. In Drew, Paul – Heritage, John (eds.): *Talk at work: Interaction in institutional settings.* Cambridge University Press, Cambridge, 163–198.

Dér Csilla Ilona – Markó Alexandra 2007. A magyar diskurzusjelölők szupraszegmentális jelöltsége. In Gecső Tamás – Sárdi Csilla (szerk.): *Nyelvelmélet – nyelvhasználat.* Kodolányi János Főiskola–Tinta Könyvkiadó, Székesfehérvár–Budapest, 61–67.

Dér Csilla Ilona 2010. „Töltelékelem" vagy új nyelvi változó? A *hát, úgyhogy, így* és *ilyen* újabb funkcióiról a spontán beszédben. *Beszédkutatás 2010.* 159–170.

Goffman, Erving 1967/1995. A homlokzatról: A rituális elemek analízise a társas interakcióban. In Síklaki István (szerk.): *A szóbeli befolyásolás alapjai II.* Nemzeti Tankönyvkiadó, Budapest, 3–31. [Az eredeti mű: On face-work. In Goffman, Erving: *Interaction ritual. Essays on face-to-face behavior.* Penguin, Harmondsworth, 1967, 5–45.]

Ilie, Cornelia 1999. Question–response argumentation in talk shows. *Journal of Pragmatics* 31. 975–999.

Ilie, Cornelia 2001. Semi-institutional discourse: The case of talk shows. *Journal of Pragmatics* 33. 209–254.

Muntigl, Peter – Turnbull, William 1998. Conversational structure and face-work in arguing. *Journal of Pragmatics* 29. 225–256.

Schirm Anita 2007. A komplex kérdés mint nyelvhasználati stratégia. In Sinkovics Balázs (szerk.): *LingDok 6. Nyelvész-doktoranduszok dolgozatai.* JATE Press, Szeged, 131–154.

Schirm Anita 2007–2008. A *hát* diskurzusjelölő története. *Nyelvtudomány* III–IV. 185–201.

Schirm Anita 2008. A *hát* diskurzusjelölő partikulához kötődő nyelvhasználati stratégia. In Gecső Tamás – Sárdi Csilla (szerk.): *Jel és jelentés.* Kodolányi János Főiskola–Tinta Könyvkiadó, Székesfehérvár–Budapest, 288–296.

Schirm Anita 2011. A diskurzusjelölők funkciói: a *hát*, az *-e* és a *vajon* elemek története és szinkrón státusa alapján. Doktori disszertáció, kézirat. SZTE, Szeged, http://doktori.bibl.u-szeged.hu/759/1/schirm_anita_doktori_disszertacio.pdf (A letöltés ideje: 2011. december 2.)

Tolcsvai Nagy Gábor 2005. Funkcionális nyelvtan: elmélet és gyakorlat. *Magyar Nyelvőr* 129. 348–362.

Vuchinich, Samuel 1990. The sequential organization of closing in verbal family conflict. In Grimshaw, Allen D. (ed.): *Conflict talk – Sociolinguistic investigations of arguments in conversations.* Cambridge University Press, Cambridge, 118–138.

AZ ISMÉTLÉS ÉS A CSERE INTERAKCIÓS FUNKCIÓI MAGYAR NYELVŰ SPONTÁN TÁRSALGÁSOKBAN

Németh Zsuzsanna

Bevezetés

Az ismétlést és a cserét mint az önjavítás eszközeit vizsgáló, a konverzációelemzés elméleti keretében végzett eddigi kutatások e javítási műveletek működésének leírását elsősorban a grammatika és a javítás viszonyának megvilágítására használták. Közülük néhányan (Fox et al. 1996; Rieger 2003; Lerch 2007; Fox et al. 2009; Fox et al. 2010) több különböző morfoszintaktikai szerkezetű nyelvet vetettek össze ebből a szempontból. E vizsgálatok közös megállapítása, hogy a javítási mechanizmus univerzális mintázatokon alapul még akkor is, ha a nyelvspecifikus jellemzők időnként elfedik e mintázatot (Fox et al. 2009), amikor az egyes nyelvek szintaxisa befolyásolja a társalgás során felmerülő interakciós célok megvalósulásának módját.

A javítás a spontán beszéd bármely problémaforrásának kezelését jelenti (Schegloff et al. 1977). Rieger (2003) e definíció alapján javításnak tekinti a következő jelenségeket: hibajavítás, szókeresés, hezitációs szünet, szünetkitöltők használata, azonnali lexikális változtatás, téves kezdés, azonnali ismétlés. A felsorolás két szempontból is magyarázatot érdemel. Egyrészt fontos megjegyezni, hogy nem minden javítás hibajavítás. A hibajavítás csupán a javítás egy altípusa, amely arra szolgál, hogy a beszélő egy téves nyelvi egységet a korrekt formával helyettesítsen (Schegloff et al. 1977). Másrészt az azonnali ismétlés is a kategóriák között szerepel (a hibajavítással egy szinten), ami azt jelenti, hogy ugyan javításnak számít, de nem hibajavításnak. Ez pedig a javítás mint átfogó terminus használatát tekintve is nagyon fontos üzenettel bír: a problémaforrás kezelését meg kell különböztetnünk a probléma kezelésétől. (Itt és a továbbiakban „problémán" mindig valamely, a felszínen bekövetkező, a beszélő által javítandónak ítélt jelenséget fogok érteni.) Ez megmagyarázza, miért tekinthető az ismétlés javításnak. Az ismétlés ugyanis ugyanazon kvázi-lexikális vagy lexikális elem(ek) egymást követő használata (Rieger 2003), így a beszélő időnyerési stratégiaként olyan problémaforrások kezelésére használhatja, mint a szóátadási szándék hiánya vagy a soron következő egység megtervezése (*de ez a szervofék* ***ez ez*** *nem veszélyes?*). Cserén valamely kvázi-lexikális vagy lexikális elem(sor) más kvázi-lexikális vagy lexikális elemmel (elemsorral) való olyan helyettesítését fogom érteni, amikor a javított és a javító szegmentum szintaktikai osztálya azonos: *és még hát a lámpánál le****fordu- fulladt.*** Rieger (2003) osztályozásán

kívül más kísérletek is történtek a javítás altípusainak megállapítására. Schegloff és munkatársai (1977) négy önjavítási funkciót különítenek el: a szókeresést, a szócserét, a személyreferencia-javítást, valamint a következő beszélő kiválasztásának javítását. Fox és Jasperson (1995) hét önjavítási típust határoz meg, melyek a következő javítási műveletek kombinációiból jönnek létre: ismétlés vagy újraindítás, csere vagy helyettesítés, hozzáadás vagy betoldás, ill. megszakítás és újrakezdés. Látható, hogy mind Schegloff és munkatársai (1977), mind Fox és Jasperson (1995) (és Fox et al. 2009 is) az ismétlést és a cserét javítási műveletnek tekintik, melyek a javítási mechanizmus végrehajtásában vesznek részt. Ez azt jelenti, hogy az ismétlés és a csere önmagukban nem tartoznak az önjavítás altípusai közé, de lehetnek azok komponensei. Noha Rieger (2003) definíciójában az azonnali lexikális változtatás, valamint az azonnali ismétlés a javítási jelenségek között szerepel, ezek nem lehetnek azonosak a javítási műveletként értelmezett cserével, ill. ismétléssel, hiszen az ugyanezen a szinten felsorolt hibajavítás is megvalósulhat a csere műveletével. Ezért a Rieger-féle azonnali lexikális változtatást, valamint azonnali ismétlést olyan javítási jelenségeknek kell tekintenünk, amelyek a csere, ill. az ismétlés műveletével valósulnak meg. Ezt az értelmezést alátámasztja, hogy Rieger ugyanebben a tanulmányában a cserét és az ismétlést javítási műveletként (repair operation) is elismeri. A jelen tanulmány az ismétlést és a cserét szintén javítási műveletekként értelmezi, melyek akár más műveletekkel együtt is létrehozhatják a javítási jelenséget.

A javítási mechanizmus nem homogén, hanem különböző részekből tevődik össze, azok szervezett együttese (Schegloff et al. 1977; Schegloff 1997, 2000). Három komponense van: a javított szegmentum (repaired segment), ez tartalmazza a javítandó elem(ek)et (repairable), a javításkezdeményezés (repair initiation) és a javító szegmentum (repairing segment) (Rieger 2003).

A pszicholingvisztikai és fonetikai szempontú megközelítések a javítást a megakadásjelenségek korrekciójának tekintik. A megakadásjelenségek a folyamatos beszédet megtörő hibajelenségek, fontosságukat a beszédprodukció szempontjából az jelenti, hogy vizsgálatukkal tanulmányozhatók a rejtett tudati működések (Fromkin 1973). Funkcionális szempontból két csoportra oszthatók (Postma et al. 1990; Gósy 2005). Az egyik csoportba a beszélő bizonytalanságából adódó megakadásjelenségek, a (nem lélegzetvételt biztosító és nem retorikai célú) néma szünetek, a hangos szünetek, a hezitálások, a nyújtások, az újrakezdések, az ismétlések és a töltelékszavak sorolandók. A másik csoportba pedig a téves kivitelezés jelenségei tartoznak, amelyek a nyelv szabályainak mondanak ellent, így ebben az értelemben hibák: az elszólások, a nyelvbotlások, a grammatikai és lexikai hibák, a változtatások és a „nyelvem hegyén van" jelenség (Gósy 2005). A spontán beszédet tarkító megakadások azért jönnek létre, mert a beszédtervezés és a kivitelezés szinte egyszerre zajlik. A megakadásjelenségek egy része javítatlan marad, de működik az önkorrekciós mechanizmus, amely a felismerést és a javítást is lehetővé teszi

(Gósy 2004). A javítás pszicholingvisztikai szempontú elemzésével számos kutatás foglalkozott az elmúlt évtizedekben (l. pl. Levelt 1983, 1989; Postma–Kolk–Povel 1990; Postma 2000; Nooteboom 1980, 2005). Levelt (1983) egy 959 javítást tartalmazó spontán, monologikus szövegeket tartalmazó korpusz elemzése után arra a következtetésre jutott, hogy a felszínen tapasztalható zavarok önjavítása három részre osztható. Az első szakasz a zavar észlelése és az artikuláció megszakítása, amelyet a folyamatos önmonitorozás tesz lehetővé. A folyamat második részét a szerkesztési szakasz alkotja, amelyet hezitáció, szünet, de legfőképpen az ún. szerkesztő kifejezések használata jellemez. Itt Levelt megjegyzi, hogy a hiba (speech error) más szerkesztő kifejezéseket aktiválhat, mint egy pusztán nem megfelelő szó. Ez azt jelenti, hogy az észlelhető zavarokon belül a hibák Leveltnél is külön kategóriát képeznek. Az önjavítás harmadik szakasza a megfelelő forma kivitelezése.

Az ismétlés spontán beszédben való gyakori előfordulását számos korábbi vizsgálat kimutatta (Clark–Wasow 1998; Gósy 2003; Gyarmathy 2006, 2007, 2009; Horváth 2007; Bóna 2006). Az ismétléssel és az újraindítással mint megakadásjelenségekkel foglalkozik Gyarmathy (2009), aki mindkét jelenséget a tervezési bizonytalanság jelzéseként fogja föl. Az ismétlés esetében „a teljes lexéma kiejtése után a beszélő megismétli azt", az újraindítás esetében pedig „az aktivált és félig kiejtett szót a teljes szó kiejtése követi" (Gyarmathy 2009: 198). A szerző – a magyarban elsőként – kimutatja a funkciószavak tartalmas szavaknál gyakoribb ismétlését. A kapott eredményt egyrészt gyakorisági tényezőkkel magyarázza, másrészt azzal, hogy a funkciószavak tárolása klisészerűen történik (Gósy 2005). Gyarmathy szerint továbbá a funkciószavak nem változtatják meg tartalmilag a közlést, nem okoznak értelmi zavart a hallgató számára, viszont fenntartják a folyamatos beszéd látszatát, így a beszélő nagy valószínűséggel fog funkciószót választani az időnyeréshez.

A jelen tanulmány az ismétlést és a cserét mint javítási műveleteket vizsgálja az azonos fordulóban megvalósuló önkezdeményezett önjavításban. Célja annak feltárása, hogy a két javítási művelet hogyan jelenik meg magyar nyelvű, spontán konverzációkban az eddig vizsgált nyelvekhez képest (Fox et al. 2009: bikol, finn, sochiapam chinantec, indonéz, angol, japán, mandarin; Fox et al. 2010: angol, német, héber). A kérdést négy változón keresztül közelíti meg: a javítási művelet típusa (ismétlés vagy csere), a javított szegmentum hossza (egy, két, ill. több szótagú), a javított szegmentum szintaktikai osztálya (tartalmas vagy funkciószó), valamint a javításkezdeményezés helye (a javított szegmentum felismerhető befejezettsége előtt vagy után). Kiinduló hipotézise a következő: valamennyi vizsgált tényező, ill. az ezek között várható esetleges összefüggések visszavezethetők a javítási műveletek interakciós funkcióira. E feltevés mögött az a nézet húzódik, hogy maga a javítási mechanizmus sem véletlenszerűen jön létre, hanem a beszélő interakciós céljait szolgálja. Emögött pedig a konverzációelemzés társalgásfogalma áll, miszerint a beszélgetés mindenkor egy interakcióba ágyazott, strukturálisan szer-

vezett folyamat (talk-in-interaction), melynek során a beszélgetőpartnerek szituációkat hoznak létre (vö. Iványi 2001).

Kísérleti személyek, anyag és módszer

Az elemzés egy összesen 145'4" terjedelmű korpuszon készült, mely 10 db 3 fős, egyenként kb. 15 perces természetes, magyar nyelvű konverzációt tartalmaz. A felvételek mindegyike kísérleti körülmények között készült összesen 17 felnőtt résztvevővel, akik között 11 nő és 6 férfi található. A korpusz 8 beszélgetése a BEA spontánbeszéd-adatbázisból származik (Gósy 2008), 2 pedig az SZTE Pszichológiai Intézetének Mérei Ferenc Viselkedéselemző Laboratóriumában készült. A statisztikai elemzést (khi-négyzet próba) az SPSS 15.0 programmal végeztem.

Az adatgyűjtést Fox et al. (2009) módszeréhez hasonlóan azokra az azonos fordulóban kezdeményezett önjavításokra szűkítettem, amelyeknél tiszta szintaktikai kapcsolat volt a problémaforrás és a javítás között (pl. *hát az ember húsz percig tud koncentráltan figyelni, és utána* ***vé- kikapcsol***). Ez azt jelenti, hogy azok az esetek, amelyeknél a beszélő elkezdett egy megnyilatkozást, megszakította azt, majd egy új megnyilatkozást indított, vagyis a törlés műveletét hajtotta végre, kimaradtak a felhasznált adatok közül. Azért szükséges említést tenni ezekről a jelenségekről, mert nagyon hasonlítanak a csere műveletére, itt azonban nem kritérium, hogy a javított szegmentum utolsó lexikális eleme és a javító szegmentum első lexikális eleme azonos szintaktikai osztályba tartozzon. A következő példában tehát nem csere, hanem törlés történik: *hogy annak hogy lenne ér- rö-* ***ren- tehát*** *hogy lenne könnyebb megtanulni*. Az azonos fordulóban kezdeményezett önjavítás azoknak a javítási stratégiáknak a gyűjtőneve, amelyekben a javítandó elem(ek), a javított szegmentum és a javításkezdeményezés is azonos fordulóban találhatók, vagyis mindhárom komponenst ugyanaz a beszélő hozza létre (Rieger 2003). Ez a javítás leggyakoribb, preferált típusa (Schegloff et al. 1977). Tovább szűkítve az adatgyűjtést, csupán az egyszerű ismétléseket és az egyszerű cseréket vontam be a vizsgálatba. Egyszerű ismétlésről, ill. egyszerű cseréről akkor beszélünk, ha az adott javításban az ismétlés, ill. a csere javítási műveletén kívül semmilyen más művelet nem vesz részt (Fox et al. 2009). Példa egyszerű ismétlésre: *most így gondolkodom,* ***hogy- hogy*** *ki*. Példa egyszerű cserére: *a legtöbbet nekünk* ***e- szörnyű*** *hallgatni*. Az egyszerű ismétlések közül csak azokat vettem figyelembe, amelyek nem a mondandó nyomatékosítását szolgálták, továbbá nem a forduló folytatásának eszközei voltak a beszédpartner közbelépése esetén. Utóbbi esetben ugyanis a beszédpartner kezdeményezte a javítást, vagyis külkezdeményezett javítás történt. Ez nem jelenti azt, hogy azok az ismétlések, amelyek arra szolgálnak, hogy a beszélő magánál tartsa a szót, soha nem tekinthetők önkezdeményezett önjavításnak. Ennek az a feltétele, hogy a beszélő célja a lehetséges váltásreleváns hely elhalasztása legyen úgy, hogy közben nem történik átfedés. Egy vagy több lexikális elem megis-

métlése tehát akkor tekinthető önkezdeményezett önjavításnak, ha a beszélő célja a lehetséges váltásreleváns hely elhalasztása átfedés megjelenése nélkül. A nyomatékosítást szolgáló, valamint az átfedésben megjelenő ismétlések nem elégítik ki ezeket a feltételeket.

Eredmények

Az elemzett korpuszban 415 ismétlés és 142 csere található. Az 1. *a)* és az 1. *b)* táblázat a korpuszban megjelenő ismétlések és cserék eloszlását mutatja a szintaktikai osztályok és a szóhossz-kategóriák tekintetében. Láthatjuk, hogy a beszélők elsősorban funkciószavakat és egy szótagú szavakat ismételnek, míg tartalmas szavakat és több szótagú szavakat cserélnek.

1. *a)* táblázat: A javítási műveletek eloszlása a szintaktikai osztályok tekintetében

	Ismétlés	Csere	Összesen
Funkciószavak	315 (76%)	48 (34%)	363
Tartalmas szavak	100 (24%)	94 (66%)	194
Összesen	415	142	557

1. *b)* táblázat: A javítási műveletek eloszlása a szóhossz-kategóriák tekintetében

	Ismétlés	Csere	Összesen
Egy szótagú szavak	304 (73%)	50 (35%)	354
Két szótagú szavak	75 (18%)	32 (23%)	107
Több szótagú szavak	36 (9%)	60 (42%)	96
Összesen	415	142	557

A szintaktikai osztályok és a javítási műveletek közötti kapcsolatot vizsgálva az elemzés azt mutatta, hogy a csoportok elemszáma nem véletlenszerű [khi-négyzet: $\chi^2(1, n = 557) = 82{,}61$, $p < 0{,}001$]. Ugyanez elmondható a szóhossz-kategóriák és a javítási műveletek közötti viszonyról is. A javítási műveletek típusait (ismétlés, csere) a szóhossz-kategóriák minden lehetséges kombinációjával összevetettem, és az elemszámok eloszlása egy esetben sem bizonyult véletlenszerűnek [khi-négyzet: $\chi^2(1, n = 461) = 13{,}99$, $p < 0{,}001$ (egy szótagú szavak, két szótagú szavak); $\chi^2(1, n = 203) = 21{,}69$, $p < 0{,}001$ (két szótagú szavak, több szótagú szavak); $\chi^2(1, n = 450) = 95{,}69$, $p < 0{,}001$ (egy szótagú szavak, több szótagú szavak); $\chi^2(2, n = 557) = 94{,}40$, $p < 0{,}001$ (egy szótagú szavak, két szótagú szavak, több szótagú szavak)].

Ismétlés

Az ismétlés gyakori interakciós funkciója a soron következő lexikális elem késleltetése (vö. Lerch 2007). Ekkor a beszélő tágabb interakciós célja a lehetséges váltásreleváns hely elhalasztása, amelyet az ismétlés interakciós funkcióján keresztül kíván elérni. A késleltetési szándék több okból is kialakulhat. Eredhet például abból, hogy a beszélő nyelvi és/vagy kognitív tervezési időt szeretne nyerni pl. szókeresés következtében, de jelezhet tekintetkérést, átfedéskezelést vagy nem preferált nyelvi cselekvést is (Fox et al. 1996). Az alábbiakban látni fogjuk, hogy a magyar korpusz eredményei alátámasztják a késleltető funkciót.

Noha az összes ismétlés 76%-a funkciószó, és csupán 24%-a tartalmas szó, csak úgy győződhetünk meg arról, hogy ez a különbség az ismétlés interakciós funkciójából adódik, és nem a korpuszban mérhető gyakoriságokból, ha ugyanezeket a kategóriákat a teljes korpuszban is megvizsgáljuk. A 2. táblázat a korpusz szavainak eloszlását mutatja a szintaktikai osztályok és a szóhossz-kategóriák tekintetében. A korpuszban 9581 funkciószó (43%) és 12 598 tartalmas szó (57%) található, ami azt jelenti, hogy az ismétlésben tapasztalt funkciószó-gyakoriság nem a kategória korpuszbeli gyakoriságának köszönhető.

A szóhosszt és a szintaktikai kategóriát egyszerre tekintve felmerül a kérdés, hogy a magyar beszélők azért ismételnek többnyire egy szótagú funkciószavakat, mert a magyarban a funkciószavak többsége egy szótagú, vagy azért, mert az egy szótagú szavak többsége funkciószó. A 2. táblázat mutatja, hogy míg a teljes korpuszban a funkciószavak 77%-a egy szótagú, az egy szótagú szavak 72%-át alkotják funkciószavak.

2. táblázat: A korpusz szavainak eloszlása
a szintaktikai osztályok és a szóhossz tekintetében

	Funkciószavak	Tartalmas	Összesen
Egy szótagú szavak	7377	2884	10261
Két szótagú szavak	1995	4815	6810
Több szótagú szavak	209	4899	5108
Összesen	9581	12598	22179

Tehát Jurafsky és munkatársai (1998) megfigyelése az angolban, miszerint a gyakori funkciószavakra jellemző a fonológiai redukáltság, a magyarban is igaznak bizonyult: a funkciószavak többsége egy szótagú, vagyis az egy szótagú funkciószavak ismétlésekor a magyarban a szintaktikai osztálynak van nagyobb szerepe.

A magyarban kapott eredmények támogatják Fox és munkatársai (2010) megállapítását, amely azt jósolja, hogy azokban a nyelvekben, amelyekben vannak a tartalmas szavakat megelőző funkciószavak (többnyire azok a nyel-

vek, amelyekben az ige elöl, ill. középen van), a beszélők inkább funkciószavakat ismételnek, mivel így elhalaszthatják a soron következő tartalmas szót. Példa erre a magyar korpuszból: *tehát például* ***a- a: a*** *külföldi dalok, ha megnézitek, tök jól hangzanak.* A háttérben tehát az ismétlés interakciós funkciója, a késleltetés áll. Az, hogy a tartalmas szavak késleltetése gyakoribb, mint a funkciószavaké, többféleképpen is magyarázható. Az egyik lehetséges ok, hogy a tartalmas szavak kategóriáját nyitott szóosztályok alkotják, így a tervezés során bármely kontextusban több alternatíva áll a beszélők rendelkezésére, és nagyobb kihívás megtalálni a megfelelőt. Az is indokolhatja továbbá a késleltetés szükségességét, hogy az egyes szavak előfordulása tekintetében a tartalmas szavak ritkábbak, mint a funkciószavak (Fox et al. 2009; vö. Clark–Wasow 1998). Mindez azt eredményezi, hogy a tartalmas szavak kognitív tervezése nagyobb erőfeszítést igényel a beszélő részéről, mint a funkciószavak kognitív tervezése, ami nemcsak az esetleges halasztás szükségességét indokolja, hanem azt is, hogy a tartalmas szavak produkciója közben is nagyobb valószínűséggel merül föl olyan felszíni probléma, amely cseréhez vezethet.

Csere

A fentiek alapján tehát azt várjuk, hogy a csere javítási műveletében a tartalmas szavak nagyobb megterheltséget fognak mutatni, mint a funkciószavak. A korábbi vizsgálatok igazolják a feltételezést: Fox és munkatársai (2009) kimutatták, hogy az angolban az egyszerű cserék többsége (61%-a) tartalmas szavakban történik; Fox és munkatársai (2010) pedig az angol, a német és a héber összehasonlítása során jutottak hasonló következtetésre.

Ha ismét megnézzük az 1. *a)* táblázatot, azt mondhatjuk, hogy a magyar vizsgálata alátámasztja ezeket az eredményeket, ahhoz azonban, hogy biztos állításokat tehessünk, a teljes korpuszban számba kell vennünk a funkció- és tartalmas szavak arányát. Míg a cserék 66%-a tartalmas szó és 34%-a funkciószó, a teljes korpuszban ez az arány 57% vs. 43%, vagyis a korpuszbeli értékek nem indokolják a javításban mért különbséget, a magyar beszélők valóban leginkább tartalmas szavakban alkalmazzák a csere műveletét. Például: *a hang* ***része- fononokon*** *keresztül terjed.*

Ahogyan az 1. *b)* táblázat mutatja, a szóhossz tekintetében már nem ilyen egyértelműek a preferenciák. Bár a cserében leggyakrabban részt vevő szavak több szótagúak, jelentős különbséget csupán a két és több szótagú szavak között tapasztalunk e tekintetben. Hogy erre magyarázatot találjunk, vizsgáljuk meg a szintaktikai osztályok alakulását szóhossz-kategóriákra lebontva a cserékben és a teljes korpuszban!

A 3. *a)* táblázatban látható, hogy noha a magyar beszélők gyakrabban alkalmaznak cserét tartalmas szavakban, mint funkciószavakban, ez a különbség nem jelenik meg az egy szótagú szavaknál. Sőt több egy szótagú funkciószó vesz részt cserében, mint egy szótagú tartalmas. Ez abból adódik, hogy amikor a beszélő a határozott névelő megismétlésével, ill. elnyújtásával kívánja elhalasztani a soron következő főnévi csoportot, ezt jellemzően (való-

színűleg gazdaságossági okokból) annak mássalhangzó nélküli alternánsával teszi. Ebből pedig az következik, hogy ha az időközben kiválasztott szó, amely a névelőt követi, magánhangzóval kezdődik, a beszélő az alternánsok közti cserére kényszerül.

A 3. *a)–c)* táblázatok szerint a csere leggyakoribb célpontjai a magyarban a több szótagú tartalmas szavak. Újra felmerül a kérdés: ennek az az oka, hogy a magyarban a legtöbb tartalmas szó több szótagú, vagy éppen ellenkezőleg, a legtöbb több szótagú szó tartalmas? Nos, ha ismét szemügyre vesszük a több szótagú, valamint a tartalmas szavak gyakoriságát a korpuszban (2. táblázat), látni fogjuk, hogy miközben a tartalmas szavak 41%-a több szótagú, a több szótagú szavak 96%-a tartalmas. Ez pedig azt jelenti, hogy amikor a magyar beszélők több szótagú tartalmas szóban hajtanak végre cserét, nagyobb valószínűséggel játszik ebben szerepet a szóhossz, mint a szintaktikai osztály.

3. *a)* táblázat: Az egy szótagú szavak eloszlása a csere műveletében és a teljes korpuszban a szintaktikai osztályok tekintetében

	Csere	Teljes korpusz
Funkciószavak	37 (74%)	7377 (72%)
Tartalmas szavak	13 (26%)	2884 (28%)

3. *b)* táblázat: A két szótagú szavak eloszlása a csere műveletében és a teljes korpuszban a szintaktikai osztályok tekintetében

	Csere	Teljes korpusz
Funkciószavak	9 (28%)	1995 (29%)
Tartalmas szavak	23 (72%)	4815 (71%)

3. *c)* táblázat: A több szótagú szavak eloszlása a csere műveletében és a teljes korpuszban a szintaktikai osztályok tekintetében

	Csere	Teljes korpusz
Funkciószavak	2 (3%)	209 (4%)
Tartalmas szavak	58 (97%)	4899 (96%)

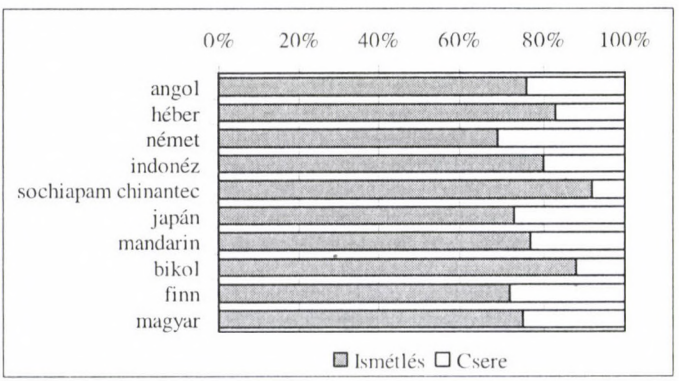
A korpusz cseréiben a két és a több szótagú szavak megterheltsége között mérhető a legnagyobb különbség. E két szóhossz-kategória abban különbözik az egy szótagú szavaktól a csere javítási művelete szempontjából, hogy a szintaktikai osztályok mindkettőben a várt eloszlást mutatják a tartalmas szavak javára. Ez azonban azt is jelenti, hogy a cserében való előfordulásuk különbsége (42% : 23% a több szótagú szavak javára) csak akkor magyarázható a szintaktikai osztállyal, ha a korpuszban több több szótagú tartalmas szó fordul elő, mint két szótagú. Ahogyan azonban a 3. *b)* és a 3. *c)* táblázat mutatja, nem ez a helyzet. A korpuszban 4815 két szótagú tartalmas és 4899 több szótagú

tartalmas szó található. Az értékek majdnem azonosak, ami csak egyet jelenthet: a két szótagú és a több szótagú szavak különbsége a cserében való megjelenésüket tekintve egyedül a szóhossz-különbségnek köszönhető. A hosszabb szavak tehát nagyobb valószínűséggel vesznek részt cserében, mint a rövidebbek. Mi lehet ennek az oka? Fox és munkatársai (2009) a csere lehetséges okaként szelekciós, ill. artikulációs nehézségeket említenek. Ez alapján azt mondhatjuk, hogy a csere interakciós funkciója egy szándékolatlan szegmentum és/vagy egy szándékolatlan artikuláció kiküszöbölése. Vagyis a beszélő akkor hajt végre cserét az interakció során, amikor egy nem szándékolt lexikális elem produkcióját indította el (ill. hajtotta végre), vagy egy szándékolt lexikális elem produkciója szándékolatlan artikulációval történt (ill. kezdődött el). A magyar korpusz tartalmas szavainak gyakori cseréje a szelekciós, a hosszabb szavak gyakori cseréje pedig az artikulációs okokat támasztja alá. A hosszabb szavak nyelvi tervezése tehát ugyanúgy nagyobb erőfeszítést igényel a beszélő részéről, mint a tartalmas szavak kognitív tervezése.

Az ismétlés és a csere viszonya az eddig vizsgált nyelvekben

Az eddigiekben láthattuk, hogy azokban a nyelvekben, amelyekben vannak a tartalmas szavakat megelőző funkciószavak, az ismétlésnek elsősorban a kognitív és/vagy nyelvi tervezésben, a cserének pedig a beszélő szándékának megfelelő kivitelezésben van szerepe. Mindez arra enged következtetni, hogy az ismétlés – funkciójából adódóan – ezekben a nyelvekben preferáltabb javítástípus lesz, mint a csere, hiszen a tervezésre szánt időt hosszabbítja meg. A csere ellenben akkor jelenik meg, amikor ez a tervezés nem volt megfelelő, azaz már elkezdődött egy nem szándékolt egység artikulációja, ill. a szándékoltnak nem megfelelő artikuláció. Az ismétlés és a csere mint javítástípusok ezért más-más szerepet töltenek be a javítási mechanizmusban (és ezáltal az interakcióban is): az ismétlés egy felszíni probléma megelőzését, a csere pedig egy már meglevő felszíni probléma orvoslását, kezelését szolgálja. Ha a javítási műveleteket aszerint jellemezzük, hogy a spontán beszéd milyen jellegű problémaforrásának kezelésére szolgálnak, akkor a kezelt problémaforrás interakcióban betöltött szerepe alapján megállapíthatjuk, hogy a csere „erősebb" javítási művelet, mint az ismétlés, hiszen a csere esetében egy, az interakcióban szándékolatlan szegmentum produkciója történik, az ismétlés viszont – késleltető interakciós funkciója révén – éppen egy esetleges szándékolatlan szegmentum kiküszöbölésére szolgál. Amíg tehát egy tartalmas szó előtti funkciószó ismétlése előkészíti a tartalmas szó szándékolt artikulációját, vagyis segíti a tervezést, egy tartalmas szó cseréje azt implikálja, hogy a tervezés a beszélői szándéknak nem volt megfelelő. Ha ez valóban így van, akkor azokban a nyelvekben, amelyekre jellemző, hogy a funkciószavak megelőzik a tartalmas szavakat, azonos nagyságú korpuszban több ismétlésnek kell előfordulnia, mint cserének. Ha megnézzük az eddigi kutatások ide vonatkozó eredményeit, azt tapasztaljuk, hogy nem csupán azokban a nyelvekben fordul elő azonos nagyságú korpuszban sokkal több ismétlés, mint

csere, amelyekben a funkciószavak és a tartalmas szavak a már említett viszonyban állnak egymással, hanem valamennyi vizsgált nyelvben (1. ábra). Ebből valószínűsíthető, hogy az ismétlés univerzálisan preferáltabb javítástípus, mint a csere.



1. ábra
Az ismétlés és a csere aránya az eddig vizsgált nyelvekben
[Az angol, a héber és a német adatok forrása Fox et al. (2010), az indonéz, a sochiapam chinantec, a japán, a mandarin, a bikol és a finn nyelvre vonatkozó adatok Fox et al. (2009) kutatásából származnak, a magyar adatok saját eredményeim.]

Felmerül azonban a kérdés: miért preferált minden eddig vizsgált nyelvben ugyanannyira az ismétlés? Hiszen nem mindegyikben használható a funkciószó-ismétlés a soron következő tartalmas szó elhalasztására, a japánban például a funkciószavak nem előzik meg a tartalmas szavakat. A választ azok az ismétlések adják meg, amelyeknél a javítókezdeményezés még a szó felismerhetőségi pontja előtt történik (*és annak a* ***rö- rögös*** *útjának a közepén tartok*). Ezeket az ismétlés alkategóriájának tekintem, és újraindításnak nevezem.

Mivel az újraindítást Fox és kutatócsoportja szintén az ismétlés alkategóriájának tekinti (vö. Fox et al. 2009; Fox et al. 2010), az 1. ábra ismétlései az újraindításokat is tartalmazzák. Ha megvizsgáljuk egy olyan nyelv ismétléseit, amelyben a funkciószavak nem előzik meg a tartalmas szavakat, vagyis nem áll a beszélők rendelkezésére a teljes szó megismétlése mint időnyerési stratégia (pl. japán), azt tapasztaljuk, hogy az ismétlések többségét a javított szegmentum felismerhető befejezettsége előtt indítják a beszélők (Fox et al. 2009). Ennek alapján felállíthatunk egy preferenciahierarchiát az ismétlés, az újraindítás és a csere között: a beszélő, ha késleltetni kívánja egy adott szó produkcióját, megvalósíthatja azt a célszót megelőző szegmentum megismétlésével. Ha ezt a halasztást az adott nyelv nem teszi lehetővé (pl. mert abban

a funkciószavak nem előzik meg a tartalmas szavakat), akkor a beszélő inkább újraindítja a célszót, csak hogy elkerülje egy olyan szándékolatlan szegmentum produkcióját, amely cserét tenne szükségessé. Célja tehát a csere elkerülése. Ez pedig megmagyarázza, hogy miért figyelhetünk meg minden eddig vizsgált nyelvben megközelítőleg azonos arányokat az ismétlések és cserék tekintetében. Ha a javítási műveleteket a kezelt problémaforrás interakciós erősségi foka alapján egy skálára helyezzük, akkor az újraindítást az ismétlés és a csere közé kell tennünk. Az újraindítás interakciós funkciója ugyanis az ismétléshez hasonlóan szintén a tervezési idő meghosszabbítása, ám itt a felszíni probléma lehetősége, amit a beszélő igyekszik kiküszöbölni, már a célszó elindítása után, a szó belsejében merül föl. Az újraindítás tehát interakciós funkcióját tekintve az ismétléssel, pozícióját tekintve a cserével állítható párhuzamba. Azért erősebb, mint az ismétlés, mert a célszó belsejében van, és azért gyengébb, mint a csere, mert újraindításkor a beszélő nem produkál szándékolatlan szegmentumot.

Az ismétlés tehát preferáltabb, mint az újraindítás. A megakadásjelenségek témakörén belül az újraindítások ismétlésekhez képest mért ritkább előfordulásáról számol be Gyarmathy (2009) is, aki lehetséges magyarázatként az önellenőrzési folyamat működését említi. A folyamatos önmonitorozásnak köszönhetően ugyanis az ismétlések esetében még a célszó artikulációjának megkezdése előtt mód van a rejtett korrekcióra. Gyarmathy azt is kimutatta, hogy az ismétlések szerkesztési szakasza (az első és a második kimondás között eltelt idő) átlagosan közel kétszerese az újraindítások szerkesztési szakaszának. Ez a jelen vizsgálat szempontjából azért érdekes, mert szintén utalhat arra, hogy az újraindítás a beszélő szempontjából kevésbé preferált, mint az ismétlés.

Tudjuk, hogy az ismétlés gyakran érint olyan funkciószavakat, amelyek tartalmas szavakat előznek meg. Mivel az újraindítás már a célszó belsejében van, azt feltételeztem, hogy az ismétléssel ellentétben inkább tartalmas szavakban fordul elő (már csak azért is, mert pozicionálisan a cserével állítottam párhuzamba). Eredményeim igazolták ezt: a korpusz újraindításainak 74%-a tartalmas szavakban történt. A cserével való (pozicionális) analógiát a szóhosszra vonatkozó eredmények is indokolják: a korpusz újraindításainak 83%-a két és több szótagú szavakban fordult elő.

Következtetések

Vizsgálatomban megkíséreltem feltárni az ismétlés és a csere működésének néhány aspektusát a magyarban. Úgy találtam, hogy mindkét javítási művelet esetében a javítási mechanizmus fő szervező elve a beszélő interakciós célja, amely a javítási műveletek interakciós funkcióiban realizálódik. A javítási műveletek működése ugyanakkor minden esetben adaptálódik az adott nyelv grammatikai lehetőségeihez.

A magyar társalgási korpusz elemzése támogatta azokat a korábbi eredményeket, amelyek szerint azokban a nyelvekben, amelyek rendelkeznek a tartalmas szavakat megelőző funkciószavakkal, a beszélők inkább funkciószavakat ismételnek, mint tartalmas szavakat. Így ugyanis elhalaszthatják a soron következő tartalmas szót (Fox et al. 2010). Az is megerősítést nyert, hogy csere leggyakrabban szelekciós, ill. artikulációs nehézségek miatt következik be (Fox et al. 2009): a tartalmas szavak gyakori cseréje az előbbi, a több szótagú szavak gyakori cseréje az utóbbi magyarázatot támasztotta alá a magyarban. Kiderült továbbá, hogy amikor a magyar beszélők több szótagú tartalmas szóban hajtanak végre cserét, nagyobb valószínűséggel játszik ebben szerepet a szóhossz, mint a szintaktikai osztály.

Végül az interakcióban betöltött szerepük alapján jellemeztem a javítási műveletek egymáshoz való viszonyát. Arra a következtetésre jutottam, hogy a csere az interakcióban erősebb javítási műveletnek számít, mint az ismétlés, ennélfogva kevésbé preferáltnak kell lennie azokban a nyelvekben, amelyekben a funkciószó-ismétlés betöltheti a soron következő tartalmas szó elhalasztásának funkcióját. Fox és munkatársai kutatásai (2009 és 2010), valamint a jelen vizsgálat azonban megmutatták, hogy nemcsak az említett szerkezetű nyelvek, hanem valamennyi eddig vizsgált nyelv az ismétlés nagyarányú preferenciáját mutatja a cserével szemben. Ezt az újraindítás jelenségével magyaráztam. Mindez elvezetett egy preferenciahierarchia felállításához a vizsgált javítási műveletek között.

Irodalom

Bóna Judit 2006. A megakadásjelenségek akusztikai és percepciós sajátosságai. *Beszédkutatás 2006.* 101–114.

Clark, Herbert H. – Wasow, Thomas 1998. Repeating words in spontaneous speech. *Cognitive Psychology* 37. 201–242.

Fox, Barbara – Jasperson, Robert 1995. A syntactic exploration of repair in English conversation. In Davis, Philip W. (ed.): *Descriptive and theoretical modes. The alternative linguistics.* John Benjamins, Amsterdam, 77–134.

Fox, Barbara – Hayashi, Makoto – Jasperson, Robert 1996. Resources and repair: a cross-linguistic study of syntax and repair. In Ochs, Elinor – Schegloff, Emanuel A. – Thompson, Sandra A. (eds.): *Interaction and grammar.* Cambridge University Press, Cambridge, 185–237.

Fox, Barbara – Wouk, Fay – Hayashi, Makoto – Fincke, Steven – Tao, Liang – Sorjonen, Marja-Leena – Laakso, Minna – Hernandez, Wilfrido Flores 2009. A cross linguistic investigation of the site of initiation in same-turn self-repair. In Sidnell, Jack (ed.): *Conversation analysis: Comparative perspectives.* Cambridge University Press, Cambridge, 60–103.

Fox, Barbara – Maschler, Yael – Uhmann, Susanne 2010. A cross-linguistic study of self-repair: Evidence from English, German, and Hebrew. *Journal of Pragmatics* 42. 2487–2505.

Fromkin, Victoria A. (ed.) 1973. *Speech errors as linguistic evidence.* Mouton, The Hague–Paris.

Gósy Mária 2003. A spontán beszédben előforduló megakadásjelenségek gyakorisága és összefüggései. *Magyar Nyelvőr* 127. 257–277.

Gósy Mária 2004. A spontán magyar beszéd megakadásainak hallás alapú gyűjteménye. *Beszédkutatás 2004.* 6–18.

Gósy Mária 2005. *Pszicholingvisztika.* Osiris Kiadó, Budapest.

Gósy Mária 2008. Magyar spontánbeszéd-adatbázis – BEA. *Beszédkutatás 2008.* 194–208.

Gyarmathy Dorottya 2006. A beszédpercepciós és beszédprodukciós folyamat összefüggései a megakadásjelenségek tükrében. In Heltai Pál (szerk.): *MANYE XVI. Tanulmánykötet.* Vol. 3/2. 449–455.

Gyarmathy Dorottya 2007. Az alkohol hatása a spontán beszédprodukcióra. *Beszédkutatás 2007.* 108–121.

Gyarmathy Dorottya 2009. A beszélő bizonytalanságának jelzései: ismétlések és újraindítások. *Beszédkutatás 2009.* 196–216.

Horváth Viktória 2007. Vannak-e „női" és „férfi" megakadásjelenségek a spontán beszédben? *Magyar Nyelvőr* 131. 315–323.

Iványi Zsuzsanna 2001. A nyelvészeti konverzációelemzés. *Magyar Nyelvőr* 125. 74–93.

Jurafsky, Daniel – Bell, Alan – Fosler-Lussier, Eric – Girand, Cynthia – Raymond, William 1998. Reduction of English function words in Switchboard. In *Proceedings of the International Conference on Spoken Language Processing* (ICSLP-98). Vol. 7. 3111–3114.

Lerch Ágnes 2007. Az ismétlés mint az önjavítás eszköze a magyarban. In Gecső Tamás – Sárdi Csilla (szerk.): *Nyelvelmélet – nyelvhasználat.* Tinta Könyvkiadó, Budapest, 123–130.

Levelt, Willem J. M. 1983. Monitoring and self-repair in speech. *Cognition* 14. 41–104.

Levelt, Willem J. M. 1989. *Speaking. From intention to articulation.* MIT Press, Cambridge, MA.

Nooteboom, Sieb G. 1980. Speaking and unspeaking: detection and correction of phonological and lexical errors in spontaneous speech. In Fromkin, Victoria A. (ed.): *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand.* Academic Press, New York, 87–96.

Nooteboom, Sieb G. 2005. Lexical bias revisited: detecting, rejecting and repairing speech errors in inner speech. *Speech Communication* 47. 43–58.

Postma, Albert 2000. Detection of errors during speech production: review of speech monitoring models. *Cognition* 77. 97–131.

Postma, Albert – Kolk, Herman H. J. – Povel, Dirk-Jan. 1990. On the relation between speech errors, disfluencies and self-repairs. *Language and Speech* 33. 19–29.

Rieger, Caroline L. 2003. Repetitions as self-repair strategies in English and German conversations. *Journal of Pragmatics* 35. 47–69.

Schegloff, Emanuel A. 1997. Third turn repair. In Guy, Gregory R. – Feagin, Crawford – Schiffrin, Deborah – Baugh, John (eds.): *Towards a social science of language: Papers in honor of William Labov.* Vol. 2: Social interaction and discourse structures. John Benjamins, Amsterdam, 31–40.

Schegloff, Emanuel A. 2000. When "others" initiate repair. *Applied Linguistics* 21. 205–43.
Schegloff, Emanuel A. – Jefferson, Gail – Sacks, Harvey 1977. The preference for self-correction in the organization of repair in conversation. *Language* 53. 361–382.

Köszönetet mondok Gósy Máriának, a lektoroknak és Németh T. Enikőnek értékes megjegyzéseikért, az SZTE Pszichológiai Intézetének, hogy a felvételek elkészítését és elemzését lehetővé tették, Orosz Gábornak a kódolásban nyújtott segítségéért, valamint Gósy Máriának, hogy rendelkezésemre bocsátotta a BEA spontánbeszéd-adatbázis nyolc beszélgetését.

IMPLIKATÚRA ÉS ARCULATMUNKA TELEVÍZIÓS BESZÉLGETŐMŰSOROKBAN

Nemesi Attila László

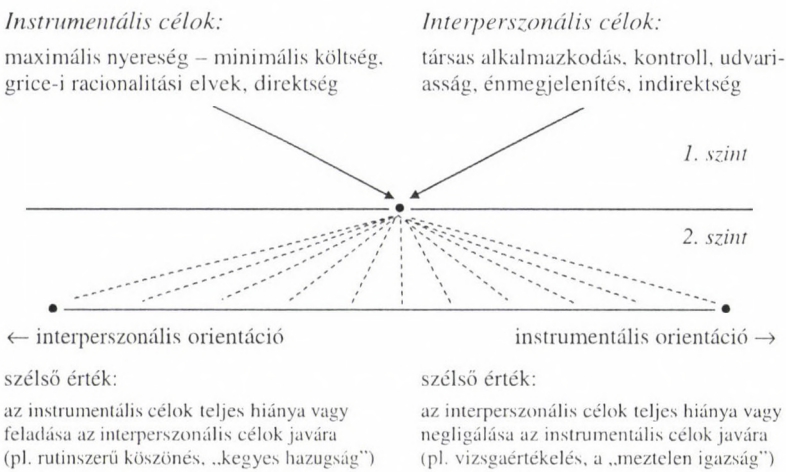
Bevezetés

Társas kontextusban a beszéd tervezésének és kivitelezésének meghatározó szociálpszichológiai pragmatikai befolyásoló tényezője az először Goffman (1955) által elemzett arculatmunka (l. Penman 1990; Holtgraves 1992; 2002; Ting-Toomey 1994; Wood–Kroger 1994; Lerner 1996; Muntigl–Turnbull 1998; Ting-Toomey–Oetzel 2003). Ebben a dolgozatban a verbális arculatmunka részeként értelmezhető implikatúrák (Grice 1975) azonosításával foglalkozom, ami azért fontos, mert az implikatúra szakirodalmában a kognitív megközelítések (l. pl. Sperber–Wilson 1986/1995; Davis 1998; 2005; Levinson 2000; Horn 2006; Borg 2009) eléggé háttérbe szorítják a diskurzusfunkciók feltárását.

Néhány korábbi munkámban a túlzás (hiperbola) gondolatalakzatának társalgásban betöltött szerepét vizsgálva kimutattam, hogy az arculatmunka – másképpen énmegjelenítés vagy benyomáskeltés – motívuma tetten érhető a sugallt jelentés egy rétegében, amelyet énimplikatúrának neveztem el (Nemesi 2003; 2004; 2009). Ezzel tulajdonképpen Goffman (1955) és Leary (1995) szociálpszichológiai modelljének „összeházasítását" kíséreltem meg Grice (1975; 1989/2011: 27–43), illetőleg a Sperber–Wilson (1986/1995) szerzőpáros implikatúraelméletével. Klasszikus magyar játékfilmeket, tehát nem spontán diskurzusokat elemeztem abból a megfontolásból, hogy egyrészt ne kelljen spekulációkba bocsátkoznom a beszélői szándékokat illetően (amelyek világosan kiderülnek a történetből), másrészt elkerüljek bizonyos etikai problémákat (az arculatmunka gyakran társas szempontból kényes, és többékevésbé leplezett törekvéseket is magában foglal), s nem mellesleg mindenki által ellenőrizhetővé tegyem megállapításaimat, hiszen e máig népszerű filmek bárki számára hozzáférhetők.

A társalgási célokat két csoportra osztottam: instrumentális és interperszonális célokra (Nemesi 2009: 132–149; 2011: 81–83). Az előbbiek a saját boldogulásunkat legjobban segítő valóságállapot-változások elérésére irányulnak, tekintet nélkül a társas kontextusra (akarok valamit, és ezt a lehető legkisebb erőbefektetéssel szeretném elérni), míg az utóbbiak a személyközi szituáció megfelelő kezelését szolgálják (udvariasság, énmegjelenítés, alkalmazkodás, kontroll stb.). Az instrumentális és az interperszonális célok vagy 1. versengenek egymással (szívességkérés, reklamáció, véleményütköztetés,

ill. kitérés a vita elől, érdekérvényesítés másokkal szemben stb.); vagy 2. egybeesnek (dicséret, udvarlás, köszönetnyilvánítás, meghívás, üdvözlet stb.); vagy 3. közömbös a viszonyuk (bejelentés, instrukcióadás, eredményhirdetés, útbaigazítás stb.); vagy 4. konfliktusban vannak (rendreutasítás, gyanúsítás, fenyegetés, szitkozódás stb.). Összehangolásuk többnyire a nyelvi közlés kisebb-nagyobb mértékű indirektségét eredményezi:



1. ábra
A társalgási célok összehangolásának modellje

Az interperszonális célok jellemzően a beszélő költségét növelő tényezőként befolyásolják a nyelvi választásokat (vö. Verschueren 1999: 55–61; Tátrai 2011: 45–50), a „legtakarékosabb" direkt stratégia helyett az indirektség változatai felé tolva el a megnyilatkozások jelentésszerkezetét (vö. *Add kölcsön a fúrógépedet! – Kölcsön tudnád adni a fúrógépedet?*). A mérlegelési folyamat, az ellenkező előjelű célok relatív erősorrendjének megállapítása racionális ágenseknél racionális kalkuláción alapul, bár számolnunk kell olyan kognitív és motivációs torzító tényezőkkel is, mint például a hangulati állapot (Forgács 2000). A Leech (1983: 107–130) által tárgyalt pragmatikai skálák (indirektség, opcionalitás, hatalom–szolidaritás stb.) a nyelvi „kimenetre" illeszthetők rá. Az egyik véglet az instrumentális cél(ok) teljes feladása, felülbírálása (például az ún. „kegyes hazugság"), a másik az interperszonális szempontokat figyelmen kívül hagyó, „orvoslás nélküli" (Brown–Levinson 1987; Nemesi 2009: 71–76; 2011: 74–77) direktség. A valóságban rendszerint az e két szélső pólus között húzódó kontinuumon próbáljuk megtalálni a közvetettség ideális fokát, és eszerint alakítjuk megnyilatkozásainkat.

Anyag és módszer

Egy nyilvános kerekasztal-beszélgetés résztvevői a szociálpszichológiai pragmatika jóslata szerint általában úgy fejezik ki véleményüket, hogy közben odafigyelnek a többiek arculatára. Négy 45 perc körüli időtartamú televíziós beszélgetőműsor – két *Záróra*-adás 2006-ból (MTV) és kettő a *Mindentudás Egyeteme*nek 2011. évi kerekasztal-beszélgetései közül – alkotja a vizsgálat korpuszát, melyek az információs, illetőleg a tudományos ismeretterjesztő műfajcsaládba tartoznak. A talk show-k (Ilie 1999; 2001; Thornborrow 2007), a hírinterjúk (Clayman–Heritage 2002; Lauerbach 2007) és a politikai vitaműsorok (Reményi 2003; Hámori 2007; Schirm 2009a; 2009b) eddig nagyobb figyelmet kaptak úgy a pragmatikában, mint az érveléselméletben (van Rees 2007). Ezért is esett a választásom az említett értelmiségi beszélgetésekre, amelyekről azt feltételezhetjük, hogy 1. bennük a kommunikáció direkt, tárgyi síkja hangsúlyos, az ének illedelmesen háttérbe vonulnak; 2. a résztvevők kivárják a fordulók végén az átmeneti pontokat, ritkán vágnak egymás szavába; 3. a műsorvezető igyekszik emelni a meghívott vendégek arculatát; 4. arculatfenyegető momentum pedig az lehet, amikor valaki pontatlanul fogalmaz vagy tárgyi tévedést követ el, és ki kell(ene) javítani, illetőleg amikor ki kell fejezni az egyet nem értést a jelenlévők legalább egyikének álláspontjával szemben (vö. a társalgási műfajok jellemzőivel: Hámori 2006).

Más háttértudásra a továbbiakban, mint az implikatúraelmélet (Grice 1975; 1989/2011: 27–42; Davis 1998; 2005; Levinson 2000; Nemesi 2009: 55–65), az arculatteória (Goffman 1955; Brown–Levinson 1987; Tracy 1990; Nemesi 2009: 71–76; 2011: 74–77), a társalgáselemzés (Atkinson–Heritage 1984; Iványi 2001; Boronkai 2009) és a diskurzusjelölők (Schiffrin 1987; Fraser 1999; Dér 2009; 2010) mibenlétének ismerete, nem lesz az olvasónak szüksége. Amire ezeken túlmenően támaszkodom, az a bevezetésben már idézett szerzőpáros, Peter Muntigl és William Turnbull kategóriarendszere.

Muntigl és Turnbull (1998) három fordulóból álló, egyetemistáktól és családtagoktól származó társalgásrészleteket elemez, melyekben valamilyen fokú egyet nem értés fejeződik ki. A szerzők négy típusba sorolják a vitatkozás beszédaktusait: 1. irrelevancia kifejezése (pl. *És akkor mi van?; Az nem érdekes; Elkanyarodsz a témától*); 2. kihívás (pl. *Várj csak, mikor is volt ez?; Honnan veszed ezt?; Fejtsd ki, légy szíves, mire gondolsz*); 3. kontradikció (pl. *De igenis számít; Ez nem így van; Épp ellenkezőleg:...*); valamint 4. ellenvélemény (pl. *Szerintem a dolgot más szempontból érdemes vizsgálni; Én viszont inkább azt mondanám, hogy...; Igen, de arról nem kell beszámolnom nektek, hogy mikor mit csinálok*). Kiemelik, hogy a kontradikció és az ellenvélemény gyakran kombinálódik egymással. Az a különbség közöttük, hogy az ellenvélemény nem pontosan az ellenkezőjét állítja a partner által mondottaknak, hanem egy új szempontot szegez azzal szembe.

E négyes felosztást az itteni céljainknak megfelelően tovább finomíthatjuk. Egyrészt alkalmasint mind a négy beszédaktus megjelenhet explicit és impli-

cit formában, másrészt bizonyára lehet őket késleltetni és enyhíteni. Az implicit forma implikatúrába foglalt, azaz sugallt beszédaktust jelent, a késleltetés legalább egy köztes fordulót vagy minimum 30 s várakozást feltételez a szóátvétel, majd a vitatkozó beszédaktus elhangzása előtt, az enyhítés pedig olyan diskurzusjelölők beiktatását, amelyeknek az a funkciója, hogy tompítsák az arculatfenyegetést. Íme tehát a taxonómiánk:

1. Irrelevancia állítása (IR)
   - explicit IR
   - késleltetett és/vagy enyhített explicit IR
   - implicit IR
   - késleltetett és/vagy enyhített implicit IR
2. Kihívás (KI)
   - explicit KI
   - késleltetett és/vagy enyhített explicit KI
   - implicit KI
   - késleltetett és/vagy enyhített implicit KI
3. Kontradikció (KO)
   - explicit KO
   - késleltetett és/vagy enyhített explicit KO
   - implicit KO
   - késleltetett és/vagy enyhített implicit KO
4. Ellenvélemény (EL)
   - explicit EL
   - késleltetett és/vagy enyhített explicit EL
   - implicit EL
   - késleltetett és/vagy enyhített implicit EL

Az explicit vitaaktusok a legerősebben arculatrombolók minden kategórián belül; a késleltetés, az enyhítés és az implicit formába csomagolás – különösen kombinálva – azt célozza, hogy a közlő csökkentse megnyilatkozásának támadó élét. Úgy vélem, a vita hőfokát jól mutatja, hogy hány és milyen típusú beszédaktus fordul elő a föntiek közül egy beszélgetésben. Elfogadva a Muntigl és Turnbull (1998: 227) által felállított sorrendet, *ceteris paribus* az IR a legsúlyosabb arculatsértés, ugyanis megkérdőjelezi a partner társalgási kompetenciájának egyik alapelemét (l. Grice: „Légy releváns!"), s ily módon korlátozza a racionális diskurzus folytatásának lehetőségét. A második legagresszívabb a KI, amely azt implikálja, hogy a másik fél nem támasztotta alá kellően állítását. Ezután következik a KO, a KO + EL és végül az EL. Lássuk, mi jelenik meg ezekből a kiválasztott négy kerekasztal-beszélgetésben.

Eredmények

Érték és minőség a könyvpiacon

Milyen lehetőségei vannak a minőségi könyvkiadásnak ma Magyarországon? Mit támogasson a kultúrpolitika? Díjazza-e a piac az igényes küllemet és a tartalmas mondanivalót? Ezekről a kérdésekről fejtette ki véleményét négy vezető hazai kiadó igazgatója – Bucsi Szabó Zsolt (Akadémiai Kiadó), Kőszeghy Péter (Balassi Kiadó), Matyi Dezső (Alexandra Kiadó) és Morcsányi Géza (Magvető Kiadó) – 2006. május 8-án (ism. 2006. július 19.) Mészáros Sándor vendégeként a *Záróra* című késő esti műsorban. A résztvevők utaltak a beszélgetés előzményeire, jól ismerték egymás álláspontját, és világosan kitűnhetett a tévénéző számára, hogy ki kinek a gondolatmenetével rokonszenvez:

(1) (a) K. P.: [...] nem lehet piacról beszélni akkor – *és itt teljesen a Zsolttal értek egyet* – akkor, amikor tényleg nyilvánvaló, hogy ezek a közvetlen piacon nem tudnak megméretődni.

(b) M. G.: *Ennyiben igaza van a Dezsőnek* – hogy volt az elején, amikor azt mondta, hogy ez nem könyvkiadási kérdés. Ez egyszerűen politikai kérdés vagy kultúrpolitikai kérdés, és [...]

Az argumentációs törésvonal az érték viszonylagosságát és a piaci megmérettetés fontosságát hangoztató M. D. és M. G., illetőleg a kulturális emlékezet eltűnése és a magas kultúra jövője miatt aggódó B. Sz. Zs. és K. P. nézete között húzódik. Az IR kivételével mindegyik vitatkozó beszédaktus többször tetten érhető a reagálásokban. Széles értelemben véve az egyetlen IR-gyanús megjegyzés – egy KO után – az alábbi:

(2) M. S.: [...] hogyha egy példát mondhatok, a *Sorstalanság*ot háromszázan vették meg, mielőtt...

M. G.: Dehogy vették meg háromszázan.

M. S.: Még annyian sem, tehát...

M. G.: Volt, amikor többen, volt, amikor kevesebben. *De a* Sorstalanság *nem jó példa, mert azzal egy csoda történt...*

M. S.: Én éppen azt akarom mondani, hogy:: hogy:: valóban ez egy egyedi példa, de:: ö::: mondhatok sok olyan példát, ö:: ö:: például a Nobel-díjas Beckettnek tizenöt évig nem fogyott el háromszáz darab egy jóval nagyobb könyvpiacon, mégis ezek a könyvek valamiképpen túl kellett éljék ezt a tizenöt évet. [...]

M. D. és M. G. a KI explicit (3a) és implicit (3b–c) módozatát veti be. M. G. hozzászólása a (3b)-ben EL + enyhített KI (l. *nem nagyon látom, nem látom egészen világosan*), a (3c) dőlttel kiemelt része pedig késleltetett KI:

(3) (a) M. D.: Az, hogy minőségi könyvkiadás, magát. *Egyet hadd kérdezzek: ez:: ez a kérdés, hogy minőségi könyvkiadás, ezt*

légy szíves, definiáld, és akkor onnantól mindenkinek könnyebb lesz a dolga. [...]

(b) K. P.: [...] természetesen van színvonalas kommersz könyv, tehát színvonalas kereskedelmi könyv, és van nagyon színvonaltalan szakkönyv is. Tehát önmagában a színvonalat nem ez dönti el. De nagyon pontosan kell látni – ez a dolog rákfenéje –, hogy a könyv nem azonos a könyvvel. Tehát, hogy mondjam... Zalatnay Saci és Maria Callas, mind a kettő énekesnő. Ennyiben a könyv az könyv. De nem egészen azonos minőséget képvisel; sőt lehet, hogy valami egészen más a funkciója – ugye, az egyiké talán inkább a szórakoztatás, a másiké, akár mondhatunk ilyen nagy szavakat, hogy művészet. Nos, ez a helyzet szerintem a könyvnél is. Legalább így válik szét.

M. G.: És akkor máris nagyon ingoványos talajon vagyunk, mer' *nem nagyon látom az evidens különbséget – miközben látom természetesen Zalatnay Sarolta és Maria Callas hanganyaga között a különbséget, de hogy:: hogy:: mondjuk, a CD-ik forgalmazásában mi a különbség, azt nem látom egészen világosan.*

(c) B. Sz. Zs.: [...] Ha a társadalom úgy dönt, hogy márpedig egy Balassi-összkiadásra szükség van, akkor ezt a társadalomnak meg kell finanszírozni, és az a könyvkiadó, az nem abban kell gondolkozzon, hogy a piacról ö:: ö:: megkapja-e a munkájának ellenértékét, hanem [...]

M. G.: [...] Én inkább visszamennék, mert változatlanul ingoványon érzem magunkat. *Ki a társadalom, és hogy dönti el, hogy valamire szükség van?*

Tanulságos megfigyelni B. Sz. Zs. késleltetett EL-ét a (3c)-beli KI-ra:

(4) B. Sz. Zs.: Szóval, visszatérve megin' ez egy:: el::: azóta azon gondolkozom, illetve próbálom a válasz[t megadni], hogy mondja meg a társadalom azt, hogy neki mi kell, és mennyit szán rá. Ugye, a társadalom bonyolult, és *amikor azt mondjuk, hogy a népakaratot valahogy ki kell fejezni, akkor azér' a társadalom megalkotta hozzá a maga intézményrendszerét – mondjuk, a parlament és ami aztán arra fölépült. Tehát igazából az ilyen apró kérdésekben is abban kéne bízni, hogy a társadalom ki tudja dolgozni maga mögé azt az intézményrendszert, amely ezekben a kérdésekben dönteni tud.* [...]

M. G. és M. D. egy-egy metapragmatikai utalása (l. a *kekeckedik* igét) jelzi, hogy tudatában vannak a KI mint argumentációs taktika vitaélező hatásának:

(5) (a) M. G.: [...] azért *kekeckedek* én ezen a társadalmi döntésen, miközben természetesen – jóllehet nem vagyok se filológus, se bölcsész, ö::: se nem foglalkozom ilyesmivel, csak távolról:: csodálom ezeket a munkákat, amit:: aminek Péter a nagy szakembere [...]

(b) M. D.: [...] de itt:: itt valahogy nem ez van, itt én azt látom, hogy itt még mindig – és ez háborít föl, tehát én ezért vagyok, most nem veled szemben, hanem maga a rendszerrel szemben egy kicsit *kekeckedő*, mert még mindig gyakorlatilag ebben nem történt meg a rendszerváltás, a támogatásokban, és engem ez borzalmasan fölháborít.

A beszélgetésben legaktívabb K. P. számos explicit KO-val nyomatékosítja egyet nem értését M. D. okfejtéseire reflektálva (6a–b), aki egyszer B. Sz. Zs.-tól is kap – igaz, inkább implicitnek minősíthető, ironikus ízű – KO-t (6c):

(6) (a) M. D.: De ez nem a kö:: ez nem könyvkiadási kérdés, hanem egy:: egy:: akár az államnak lehet egy financiális, és nem a könyvkiadóknak kell ezt föltenni. Tehát ez:: ez nem a könyvkiadás problémája.

K. P.: *Ez egy alapvető könyvkiadási kérdés, ez egy alapvető könyvkiadási kérdés*, és szó nincs arról, hogy [...]

(b) M. D.: Attól, hogy ötezer vagy tízezer forint egy olyan könyv, ami valakinek a munkájához nélkülözhetetlen, azt nem hiszem el, hogy:: hogy árkérdés lenne, hogy megveszik-e vagy sem.

K. P.: Hát *lehet, hogy te nem hiszed el, de ez attól még így van.*

(c) M. D.: [...] nálunk minden szótár kapható, ami van a piacon. Hatalmas az árkülönbség, tehát:: öö:: úgy gondolom, hogy öthatszoros a ti áraitok az ő szótáraihoz képest.

B. Sz. Zs.: *Azt csak gondolod...* [Bocsánat... ((mosolyog))

M. D.: [Rendben van. Akkor négy, hát ez mindegy, tehát igazából ennek most nincs jelentősége.

A felek néhány visszakozással (7a–b) és részleges egyetértéssel (7c–d) oldják az alapvető nézeteltérésből fakadó feszültséget:

(7) (a) K. P.: [...] Vigyázat, am:: én nem ezt mondtam, amit te:: ööö::: ahová alakítottad a szövegemet, mar... [maradjunk ennyiben.

M. D.: [*Elnézést, ha más-hogy mondtam...*

K. P.: Semmi gond, de én azt mondtam, hogy [...]

(b) K. P.: [...] Bocsáss meg, hadd mondjam azt, Géza, szóval... Természetesen nyilván van nekem is egy csőlátásom. De azér' annyira nem cső a cső, sz'al én is olvasok verset. Még mai verset is olvasok.

M. G.:*Jó, abszurd volt a...*

K. P.: Ö:: nem:: nem erről szól a történet. Termész:: Az érték és a nem érték között azér' van különbség, tényleg ne::: legyünk teljesen relativisták. [...]

(c) M. D.:*Most:: maximálisan egyetértek azzal, amit mondtál, és szerintem most nagyon jól összefoglaltad a rákfenéjét az egésznek, s nekem sem azzal van a bajom, hogy ki kell adni kritikai kiadásokat* [...]

(d) K. P.: Hát igen, csak tényleg nagyon világosan – és érdekes, *ez úgy látszik, a:: a::: békés egyetértés estéje, a mai, nagyon helyes*, de azt tényleg világosan kell látni, hogy:: amíg nincs kvázi piac teremtve [...]

Talán nem tévedünk, ha úgy értékeljük, hogy azért viszonylag nagy a KI-ok, KO-k és EL-ek száma, mert az egyes kiadók üzleti érdekei ütköznek, illetőleg a beszélgetés során feltáruló világnézeti-társadalomfilozófiai szembenállás is fokozza a vita intenzitását.

Történészek a nándorfehérvári diadalról

2006. május 15-én sugározta először, majd július 20-án ismételte meg az MTV 2-es csatornája azt a legkevésbé sem konfrontatív eszmecserét, amelyben Feledy Péter műsorvezető négy történésszel (Fodor Pál, Pálffy Géza, Pálosfalvi Tamás, Tringli István) idézte föl a nándorfehérvári csata előzményeit és magát az ütközetet. A beszélgetés mindössze három explicit KO-t tartalmaz, amelyből kettő késleltetett, egy pedig a műsorvezető kérdésként megfogalmazott bizonytalan állítását igazítja ki. Nézzük meg a legérdekesebb részletet:

(8) Fe. P.: Na most, ennek a bizonyos ö:: ö:: rigómezei csatának azér' – Magyarországtól ez még viszonylag messze van, de – *mi volt a közvetlen*, ö::: hogy mondjam, ö:: *negatívuma* a Magyar Királyság számára?

P. T.: *Közvetlen negatívuma ö:: nem volt*; felhívta a figyelmet arra, hogy itt megjelent a a::: déli, délkeleti szomszédságban egy olyan hatalom, amely alapvetően azért újra:: hatalmi viszonyokat, öö: és amely ellen az a korábbi offenzív külpolitika, amelyet az Anjou királyok folytattak, az nem ígérkezik gyümölcsözőnek [...]

Fe. P.: *Bocsásson meg, csak azért szakítom félbe, mert akkor:: hát az nem úgy volt, hogy azért* a rigómezei csatának következménye volt, hogy Szerbia meghódolt, behódolt [Törökországnak, vagy az oszmánoknak...]

Fo. P.: [Ezt:: ezt mindenképp hozzá kell tenni, igen.] *Kézzelfogható következményei lettek Magyarországra nézve*: először jelentek meg az ország területén török dúló csapatok, és...

Mint látjuk, a műsorvezető először megfogalmaz egy kérdést, majd az egyik történész explicit módon ellentmond a kérdésben rejlő előfeltevésnek. Ezután a műsorvezető késleltetve és enyhítve (l. *bocsásson meg, hát*) implicit kontradikcióval él (újabb kérdés), megerősítve az előző kérdésében megfogalmazott előfeltevést, amellyel a másik történész egyetért, és végül késleltetve cáfolja kollégájának kijelentését (*kézzelfogható következményei lettek*). Talán nem véletlen, hogy Fo. P. később két arculaterősítő megjegyzést is tesz, amellyel kompenzálni tudja a (8)-ban P. T.-t ért sérelmet (*a Tamás asszem, hogy nagyon jó helyen keresgélte a magyarázatot*; *ezt a Tamás nagyon szépen megírta*). A beszélgetés végén P. T. még hozzáfűz egy KO-t P. G. egyik gondolatához, ám ez alig hallható a résztvevők együttbeszélése és a műsoridő lejárta miatt.

Az arculatfenyegető beszédaktusok csekély száma ez esetben azzal magyarázható, hogy a négy felkészült történész a tényekre koncentrál, tárgyi tévedés nemigen fordulhat elő, amellett a műsorvezető nem a vitatott tudományos kérdésekre helyezi a hangsúlyt, hanem az ismeretterjesztésre.

Nyilvántartanak – éljünk adattudatosan!

A *Mindentudás Egyeteme*nek kerekasztal-beszélgetései, amelyek elérhetők a világhálón (http://mindentudas.hu/kerekasztalok.html), sokban hasonlítanak a *Záróra* korábbi adásaira. A moderátor, Lévai Balázs személyében is képviseli a folytonosságot, hiszen 2000 és 2005 között a *Záróra* egyik műsorvezetője volt. 2011. január 16-án három nemzetközi hírű magyar hálózatkutatót (Barabási Albert-László, Csermely Péter, Kertész János), egy „pletykaszakértő" szociálpszichológust (Szvetelszky Zsuzsanna) és egy üzleti tapasztalatokkal rendelkező fiatal cégvezetőt (Ruppert Péter) hívott meg a stúdióba, hogy az emberi dinamikáról, kivált a valamennyiünkről nap mint nap keletkező digitális adattömeg jó célokra való felhasználásának lehetőségeiről kérdezze őket. Mindjárt az elején tisztázza, hogy kölcsönösen tegeződni fognak egymással, és ez a közvetlen hangnem végig jellemzi a beszélgetést.

Kezdjük a tévedések kezelésével: a rutinos televíziós L. B. két pontatlansága Cs. P.-t enyhített explicit (9a), B. A-L.-t implicit KO-ra (9b) kényszeríti:

(9) (a) L. B.: De hogyha rögtön személyesíteni akarom a dolgot, akkor rólad tudom:: – mer' köztudomású, vagy elárultad nekünk [...] –, hogy neked például nincsen mobiltelefonod. Tehát

magyarán te egy veszélyforrást ilyen értelemben likvidálsz az életedből.

Cs. P.: *Ó, nekem van mobiltelefonom*, csak SMS-re használom kizárólag, tehát [a veszélyforrás csökken...

L. B.: [Az is:: az is rögzít adatot...

(b) L. B.: Az az érdekes, hogy ugyanakkor szer'tem többen is beleütköztetek abba a problémába, hogy amikor elkezdtetek kutatni – mondjuk, jelen esetben a::: nektek is volt egy ilyen kutatásotok, amely több millió öö:: mobiltelefon-hívás adatait vizsgálta, meg a te esetedben is, *még az első könyvednél, ugye*, hogy:: hogy:: hogy:: hogy megrémült a közvélemény attól, hogy:: hogy ezeket az adatokat más célra is lehet felhasználni, vagy egyáltalán kutatási cél kiindulás:: öö:: alapja lehet.

B. A-L.: Igen, mi:: valóban volt egy néhány hetünk, ami nagyon kényes volt; *mikor megjelent az első cikkünk a* Nature-*ben az emberi mozgásról*, akkor valahogy az emberek azon a cikken keresztül és különösen az arról a cikkről megjelenő sajtón keresztül jöttek rá, hogy [...]

A (9a)-ban Cs. P. emotívája (*Ó*) és L. B. közbevetése érdemel külön figyelmet az arculatvédés szempontjából, a (9b)-ben pedig B. A-L. úgy javítja ki L. B. pontatlanságát – az *igen* szócskával vezetve be a fordulót –, hogy azt alig lehet észrevenni (nem az első könyvének megjelenése kapcsán, hanem pár évvel később történt, amire L. B. utalt). Egyetlen további KO fordul elő a beszélgetésben, melynek „szenvedő alanya" Sz. Zs.:

(10) Sz. Zs.: Ez a szóban elhangzó információra is vonatkozik ugyanúgy.

K. J.: *Nem egészen ugyanúgy, nem egészen ugyanúgy.*

((L. B. egy rövid történetet mesél a fiáról, aki a Facebookon egy órára személyiséget cserélt egy barátjával, majd mindenki nevet Sz. Zs. egy megjegyzésén.))

K. J.: *Igen, szóval nem egészen ugyanarról van szó*, mert teljesen virtuális személyiségeket lehet létrehozni. Hogyha valakivel találkozom, akkor azér' az nem egy virtuális személy, és megvannak azok a:: ö::: – nyilván te ezt sokkal jobban tudod nálam – azok a:: technikák, amelyekkel legalábbis nagy valószínűséggel szűrni lehet ezeket a [...]

K. J. a megismételt KO enyhítésével (*Igen, szóval nem egészen*) és Sz. Zs. szakmai kompetenciájának megerősítésével (*nyilván te ezt sokkal jobban tudod nálam*) mérsékli az arculatfenyegetést.

Ami a KI-okat illeti, Cs. P. és B. A-L. magát a műsorvezetőt hozza egy helyütt zavarba, pedig L. B. szerepéből adódóan nem feltétlenül a saját véleményének szószólója:

(11) L. B.: Most azért egy picit visszautalnék arra, amit János mondott, tehát hogy:: hogy az az előnye, hogy akkor pontosabban meg tudják marketing szempontból határozni azt, hogy kik is vagyunk, de hát ez t'onképpen valahol a hátránya is, nem? Tehát, hogy:: hogy most már léteznek interaktív reklámfelületek, mondjuk, hogy odalépek elé, rögtön lát egy kamera, és belövi, hogy én hány éves lehetek, hogy fehér vagyok, férfi, és a többi, és egyből olyan hirdetéssel támad meg engem, amelynek célcsoportja lehetek.

Cs. P.: *De hát ez most baj?* Hát:: hát:: az em[bernek az életében...

L. B.: [Csak felvetés...

Cs. P.: *Okvetlenül baj?* Tehát az embernek az életében így megszűnik a szemét – vagy nem megszűnik, de legalábbis mérséklődik. Azt hiszem, nem vagyok egyedül azzal a borzasztó érzéssemmel, hogy információs szempontból szeméttengerben élek. [...]

B. A-L.: *Tehát a kérdés az, hogy inkább női alsóneműreklámot szeretnél látni?*

((Nevetés.))

L. B.: Ha engem kérdezel, akkor egyébként igen. ((Nevetés.)) De, mondjuk, ee:: nyilván azért nagyon sokan szerintem a saját személyiségi jogaik védelmében esetleg azt mondják, hogy ne azonosítsanak rögtön, tehát hogy:: hogy maradjon meg ez a semleges tér, én most az utcán sétálok, engem tessék békén hagyni.

B. A-L.: *Te ezt mondanád? Egyébként őszintén így gondolod?*

L. B.: Hát, szerintem ez egy kompe:: Hát, nyilván itt, hogy:: a fogyasztói társadalom különböző kérdéseivel összefügg...

B. A-L.: *Ne társadalmi kérdésként vesd fel. Te, te magad.*

L. B.: Enge':: engem valószínűleg nem zavarna.

B. A-L.: És az érdekes az, hogy ha nagyjából körülmegyünk, és megkérdezünk minden embert, akkor kiderül, hogy nagyjából senkit se zavar, de mindig fölemelkedünk egy ilyen magasabb síkra, és azt mondjuk, ez egy társadalmi probléma.

L. B. arculatmentő közbevágásával (*Csak felvetés...*) nem tudja elérni, hogy Cs. P. kérdésbe foglalt implicit KI-át ne kövessék újabbak, viszont B. A-L. humorral fűszerezi a beszélgetést, L. B. pedig végül bólogatással jelzi, hogy meggyőzőnek tartja B. A-L. érvelését. Utóbbi egyébként az egyedüli, aki – a tegeződés mellett – megkapja egy ízben L. B.-től a „professzor úr" címzést,

és ugyancsak ő az, aki egy másik partnerrel szemben is alkalmazza a kérdésként megformált KI-t egy enyhített EL (*Én emiatt kevésbé aggódok*) után:

(12) Cs. P.:És, ugye, sokszor előfordult az, hogy:: hogy gyönyörű önéletrajz, fantasztikus interjú, s utána a személyzetis megnézte, vagy a hr-es megnézte a Facebookot, és azt mondta, hogy köszönöm szépen, ebből az emberből mi nem kérünk. Tehát:: ilyen öö átcsatolások is vannak.
B. A-L.: *Én emiatt kevésbé aggódok.* Valóban voltak erre példák, de viszont azt is tudjuk, hogy a médiának van egy hhö képessége arra, hogy ha egy ilyet talál, akkor abból ekkora lesz, igaz.
Cs. P.:Ez is igaz, ez is igaz.
B. A-L.: Én:: én, én kihívásnak tartom azt, hogy valaki 20 év múlva, egy hr-es találjon olyan fiatalt, akinek ne legyen valami kényes dolog a Facebookján. *És a következő kérdés, amit feltennék: felvennél-e te egy olyan fiatalt, akinek abszolúte soha semmi kényes dolog nem volt a Facebookján?* [...]

B. A-L. közel 20 éve él az Egyesült Államokban, mely individualizmusáról közismert, így a KI-ok kedvelése esetleg az amerikai tudományos vitakultúra hatása is lehet nála (vö. Hofstede 1991; Ting-Toomey–Oetzel 2003).

Néhány további példa az EL-ek bevezetésére (l. a részleges egyetértés és az *azért valahol* diskurzusjelölők közreműködését az EL tompításában):

(13) (a) B. A-L.: *Na most ilyen szempontból igazad van, de* van egy alapvető különbség [...]
(b) Cs. P.: *De a kettő azér' valahol átmenet egymásba*, tehát...
(c) R. P.: *De, csak hogy még vissza::lépve egy:: kicsit az előző kérdésre, hogy sz'ntem a Facebookon azért nagyon nehéz virtuális ö:: ö:: személyiséget létrehozni*, hiszen ott a barátoknak vissza kell igazolni, tehát [...]
(d) B. A-L.: *Ez jogos, viszont* emellett a megdöbbentő az, hogy [...]

Melyek a magyar versenyképesség befolyásolható tényezői?

Lévai Balázs 2011. március 6-án Bogsch Erik (vegyészmérnök, a Richter Gedeon Rt. vezérigazgatója), Bojár Gábor (fizikus, a Graphisoft szoftverfejlesztő cég elnöke), Chikán Attila (közgazdász, egyetemi tanár, volt gazdasági miniszter), Czakó Erzsébet (szociológus, tanszékvezető egyetemi docens) és Szepesi Balázs (közgazdász, a Nemzetgazdasági Minisztérium egyik államtitkárának kabinetfőnöke) részvételével igyekezett körüljárni hazánk nemzetgazdaságának fejlődési feltételeit. Ebből a beszélgetésből – szemben az előzővel – nemcsak az IR, hanem a KI példái is teljesen hiányoznak, akárcsak a nándorfehérvári diadalról szóló *Záróra*-adásból. A szakemberek inkább több oldalról megvilágítják a kérdést, mintsem vitába bocsátkoznának egymással. Három momentumot mindazonáltal kiragadhatunk. Az első az, amikor L. B.

kissé bizonytalanul (*ha jól értem*) összefoglalja Ch. A. előző gondolatmenetét, aki közbeszól, és kijavítja egy KO-val:

(14) L. B.: [...] akkor mely területekre kéne koncentrálni, mer' amit professzor úr mondott, abból nekem az szűrődött le, hogy:: hát hogy mondjam, középtávon, mondjuk, öt–nyolc éven belül, ezeken a k:: ezeken a::: intézményi feltételeken, ha jól értem, azér' nagyon nehéz lesz változtatni. Melyek...
Ch. A.: *Az:: azér' nagyon sokat lehet rajta változtatni.*
L. B.: Jó:: Akkor ez::
Ch. A.: Az értékekről volt szó, hogy azokat [nagyon nehéz, az intézményi rendszeren, azon lehet.
L. B. [Világos, de mondjuk, a:: a:: Öö, igen. Hogy akkor pont e:: pont erre akartam fókuszálni, hogy akkor melyik, melyik, melyik területre érdemes ö::: leginkább koncentrálni, vagy melyik területen lehet a leginkább elvárni a versenyképességet.

Újfent az *azért* diskurzusjelölőt látjuk felbukkanni enyhítő szerepben [vö. (4), (7b), (8), (10), (11), (13b–c), (15), (16)], illetőleg L. B.-nak a KO-val szembeni védekezése [vö. M. S.-éval a (2)-ből] árul el sokat a verbális arculatmunka mibenlétéről.

A második illusztráció egy több fordulón át késleltetett EL-nyel indul. Azt is jól illusztrálja, hogy a KO és az EL között nem mindig könnyű éles határvonalat húzni:

(15) B. G.: *Hát én hadd vitatkozzak Erikkel, a:: Ez már régi vitánk*, hogy a kormányzati támogatás vagy a kitörési pontok vagy egy-egy iparág ö:: kiemelt támogatása, az mennyire lehet helyénvaló. Én az' hiszem, hogy amikor a Nokia nagy lett, akkor még nem támogatta őt a finn állam. [...] Magyarországon, amikor igazi erős gyógyszeripar alakult ki – mondjuk, ezt biztos jobban tudod, mint én, de én úgy tudom, hogy amikor Richter Gedeon megcsinálta ezt a gyárat, akkor nem állt mögé az állam, tehát nem az állam találta ki, hogy Magyarország kitörési pontja a gyógyszeriparban lesz, hanem fölismerte azt egy jó vállalkozó, nevezetesen Richter Gedeon. De...
B. E.: *De, de azér' itt közbe kell, hogy vágjak: viszont hátrányosan semmiképpen nem kapott...*
B. G.: Hátrányosan nem, hátrányosan biztosan nem...
B. E.: *Egy. Kettő: téves, mer' igenis az állam támogatta, tehát:: ö:: indirekt eszközökkel akkor is*, tehát:: ö::: én úgy gondolom, hogy egyébként a mostani helyzetet semmiképp nem lehet összevetni egy:: évszázad előtti helyzettel. [...]

B. E. első reagálása EL-nek, a második KO-nak tekinthető (l. *téves*, *igenis*), a folytatás pedig (*én úgy gondolom...*) ismét EL-nek.

A harmadik idézetben Ch. A. azt implikálja, hogy B. G. általánosságban fogalmaz (s talán téves következtetéseket von le), miközben a valóság árnyaltabb megközelítést kíván:

(16) Ch. A.: *Azér' ebbe':: ebbe' nagyon nagyok a különbségek asszem cégek és cégek között. Te a:: a:: az' hiszem, hogy a:: élvonalát látod ennek a dolognak, mer:: az egyik ö:: dolog az az, hogy:: hogy minden jel arra mutat, hogy a multik szerepéről meg a kis-közepes vállalatok helyzetéről is olyan sok sztereotípia hangzik el, ami általánosságban, hogyha átlagolnánk a dol::ni lehetne a dolgokat, lehet, hogy igaz lenne, de ha rendesen megnézzük, akkor minden szektorban, minden ágazatban teljesen más a multik szerepe Magyarországon, mások a kis-közepes vállalatok, tehát ha:: ha:: ha:: ha:: csak azt nézzük, hogy úgy általában mi van, akkor nagyon sok téves következtetés::öö::re van ö:: ö:: lehetőség.* [...]

Érzékelhető tehát a beszélgetésből, hogy vannak nézetkülönbségek a helyzetértékelésben, de ezeknek a meghívottak csak pár esetben adnak hangot, és rendre megjegyzik, ha valamiben – legalábbis részlegesen – egyetértenek a másik által mondottakkal.

Következtetések

E dolgozat fő újdonsága, hogy árnyalja Muntigl és Turnbull (1998) négyes osztályozását a vitatkozás beszédaktusairól, lehetővé téve a társalgás szerkezeti vizsgálata mellett a finomabb diskurzuselemzést. A célom az volt, hogy kitűnjön az implikatúrák szerepe az arculatmunkában, és tudományos igénnyel ellenőrizhessük a televíziós beszélgetőműsorok konfrontatív jellegére vonatkozó benyomásainkat. A tárgyalt implikatúrák implicit KI-ok, KO-k és EL-ek alakját öltik, de az enyhített aktusokat is úgy értelmezem, mint amelyek a nem enyhített megfelelőjüket és azt a gyenge implikatúrát (vö. Nemesi 2009: 137–149) hordozzák, hogy a beszélő tiszteletben tartja a partner pozitív arculatát (közbevágáskor talán a negatívat is). Kiderült, hogy a KI-okat a résztvevők szinte mindig kérdésbe burkolják, és tipikus a részleges egyetértés, illetőleg az enyhítő diskurzusjelölők használata, melyek közül gyakorisága miatt külön is kiemelendő az *azért* szócska.

Hogy össze tudjuk hasonlítani az arculatmunka vizsgált eszközeinek számszerű viszonyait az egyes beszélgetésekben, érdemes egy táblázatba rendezni az adatokat. Jól látszik, hogy az *Érték és minőség a könyvpiacon* című vita volt a legkonfrontatívabb, a nándorfehérvári diadalról szóló a „legbékésebb", a két *Mindentudás Egyeteme* kerekasztal pedig köztes helyet foglal el képzeletbeli skálánkon:

1. táblázat: A vitatkozó beszédaktusok előfordulása a négy beszélgetésben

Beszédaktusok	A nándorfehérvári diadal	Érték és minőség	Éljünk adattudatosan!	A magyar versenyképesség
explicit IR	–	–	–	–
– enyhített	–	–	–	–
– késleltetett	–	1	–	–
– enyh./késl.	–	–	–	–
implicit IR	–	–	–	–
– enyhített	–	–	–	–
– késleltetett	–	–	–	–
– enyh./késl.	–	–	–	–
explicit KI	–	1	1	–
– enyhített	–	–	–	–
– késleltetett	–	–	–	–
– enyh./késl.	–	–	–	–
implicit KI	–	2	4	–
– enyhített	–	1	1	–
– késleltetett	–	1	–	–
– enyh./késl.	–	–	–	–
explicit KO	1	8	–	4
– enyhített	–	–	2	1
– késleltetett	2	–	–	–
– enyh./késl.	–	–	1	–
implicit KO	–	1	2	–
– enyhített	–	–	–	–
– késleltetett	1	–	–	–
– enyh./késl.	–	–	–	–
explicit EL	–	6	3	1
– enyhített	–	5	2	1
– késleltetett	–	–	–	2
– enyh./késl.	–	2	1	–
implicit EL	–	1	1	1
– enyhített	–	–	–	1
– késleltetett	–	–	–	–
– enyh./késl.	–	–	–	–
Összesen:	4	29	18	11

Az egyik megszólaltatott tudós, Barabási Albert-László a műsorvezető fölvetésére – hogy tudniillik arra bizonyára ő is odafigyel, mit ír le, milyen szavakat használ a Facebookon – azt fejtegeti, hogy minden helyzetben próbálunk megfelelni az adott műfaj normáinak: „mint ahogy itt is figyelek arra, hogy bizonyos módon kell beszélni egy kerekasztalnál, és másképp beszélek a fiammal [...] játék közben, és megint másként beszélek, tegyük fel, egy bárban valakivel". A kerekasztalviták alkalmat teremtenek arra, hogy a felek kifejtsék álláspontjukat, a grice-i maximáknak megfelelve válaszoljanak a moderátor kérdéseire, kapcsolódjanak az előttük szólóhoz egyetértést vagy ellenvéleményt kifejezve, s mindezt általában úgy teszik, hogy saját arculatuk ne sérüljön, a többiekét pedig ne fenyegessék oly módon és mértékben, hogy abból sértődés vagy nyílt konfliktus származzon. Ezzel függ össze, hogy az IR aktusa alig vagy egyáltalán nem dokumentálható a vizsgált beszélgetésekben (jóllehet a hétköznapi viták visszatérő eleme), KI-okra is csak két műsorban akadt példa, KO-k és EL-ek azonban mindenütt megjelentek.

Irodalom

Atkinson, J. Maxwell – Heritage, John (eds.) 1984. *Structures of social action: Studies in conversation analysis.* Cambridge University Press, Cambridge.

Borg, Emma 2009. On three theories of implicature: Default theory, relevance theory and minimalism. *International Review of Pragmatics* 1. 63–83.

Boronkai Dóra 2009. *Bevezetés a társalgáselemzésbe.* Ad Librum, Budapest.

Brown, Penelope – Levinson, Stephen C. 1987. *Politeness: Some universals in language usage.* Cambridge University Press, Cambridge.

Clayman, Steven – Heritage, John 2002. *The news interview: Journalists and public figures on the air.* Cambridge University Press, Cambridge.

Davis, Wayne 1998. *Implicature: Intention, convention, and principle in the failure of Gricean theory.* Cambridge, Cambridge University Press.

Davis, Wayne 2005. Implicature. In Zalta, Edward N. (ed.): *The Stanford encyclopedia of philosophy.* (Summer 2005 Edition.) http://plato.stanford.edu/entries/implicature. (A letöltés ideje: 2011. november 22.)

Dér Csilla Ilona 2009. Mik is a diskurzusjelölők? In Keszler Borbála – Tolcsvai Nagy Gábor (szerk.): *Diskurzus a grammatikában – grammatika a diskurzusban.* Tinta Könyvkiadó, Budapest, 293–303.

Dér Csilla Ilona 2010. On the status of discourse markers. *Acta Linguistica Hungarica* 57. 3–28.

Forgács József 2000. Az érzelem és a stratégiai kommunikáció: az érzelem hatása a szóbeli kérések megfogalmazására és értelmezésére. *Magyar Pszichológiai Szemle* 55. 145–178.

Fraser, Bruce 1999. What are discourse markers? *Journal of Pragmatics* 31. 931–952.

Goffman, Erving 1955. On face-work: An analysis of ritual elements in social interaction. *Psychiatry* 18. 213–231.

Grice, H. Paul 1975. Logic and conversation. In Cole, Peter – Morgan, Jerry L. (eds.): *Syntax and semantics. Vol. 3: Speech acts.* Academic Press, New York. 41–58.

Grice, H. Paul 1989/2011. *Tanulmányok a szavak életéről.* Gondolat Kiadó, Budapest.

Hámori Ágnes 2006. A társalgási műfajokról. In Tolcsvai Nagy Gábor (szerk.): *Szöveg és típus. Szövegtipológiai tanulmányok.* Tinta Könyvkiadó, Budapest, 157–181.
Hámori Ágnes 2007. Párbajokról nyelvészetileg. Miniszterelnöki viták jellegzetességei társalgáselemzési szempontból. *Századvég* 43. 3–35.
Hofstede, Geert 1991. *Cultures and organizations: Software of the mind.* McGraw-Hill, London.
Holtgraves, Thomas M. 1992. The linguistic realization of face management: Implications for language production and comprehension, person perception, and cross-cultural communication. *Social Psychology Quarterly* 552. 141–159.
Holtgraves, Thomas M. 2002. *Language as social action: Social psychology and language use.* Lawrence Erlbaum, Mahwah.
Horn, Laurence R. 2006. Implicature. In Horn, Laurence R. – Ward, Gregory (eds.): *The handbook of pragmatics.* Blackwell, Oxford, 3–28.
Ilie, Cornelia 1999. Question–response argumentation in talk shows. *Journal of Pragmatics* 31. 975–999.
Ilie, Cornelia 2001. Semi-institutional discourse: The case of talk shows. *Journal of Pragmatics* 33. 209–254.
Iványi Zsuzsanna 2001. A nyelvészeti konverzációelemzés. *Magyar Nyelvőr* 125. 74–93.
Lauerbach, Gerda 2007. Argumentation in political talk show interviews. *Journal of Pragmatics* 39. 1388–1419.
Leary, Mark R. 1995. *Self-presentation: Impression management and interpersonal behavior.* Westview Press, Boulder.
Leech, Geoffrey N. 1983. *Principles of pragmatics.* Longman, London.
Lerner, Gene H. 1996. Finding "face" in the preference structures of talk-in-interaction. *Social Psychology Quarterly* 59. 303–321.
Levinson, Stephen 2000. *Presumptive meanings: The theory of generalized conversational implicature.* MIT Press, Cambridge (MA).
Muntigl, Peter – Turnbull, Willam 1998. Conversational structure and facework in arguing. *Journal of Pragmatics* 29. 225–256.
Nemesi Attila László 2003. A túlzás szerepe a személyközi retorikában. *Általános Nyelvészeti Tanulmányok* XX. 195–219.
Nemesi Attila László 2004. What discourse goals can be accomplished by the use of hyperbole? *Acta Linguistica Hungarica* 51. 351–378.
Nemesi Attila László 2009. *Az alakzatok kérdése a pragmatikában.* Loisir Kiadó, Budapest.
Nemesi Attila László 2011. *Nyelv, nyelvhasználat, kommunikáció. Hét tanulmány.* Loisir Kiadó, Budapest.
Penman, Robyn 1990. Facework & politeness: Multiple goals in courtroom discourse. *Journal of Language and Social Psychology* 9. 15–38.
Reményi Andrea Ágnes 2003. Kampányviták 2002-ben. Interakciós elemzés. In Sárközy Erika – Schleicher Nóra (szerk.): *Kampánykommunikáció.* Akadémiai Kiadó, Budapest, 65–99.
Schiffrin, Deborah 1987. *Discourse markers.* Cambridge University Press, Cambridge.
Schirm Anita 2009a. Televíziós beszélgetőműsorok nyelvi szabályszerűségei. In Gecső Tamás – Sárdi Csilla (szerk.): *A kommunikáció nyelvészeti aspektusai.* Tinta Könyvkiadó, Budapest, 222–227.

Schirm Anita 2009b. The role of questions in talk shows. In Dynel, Marta (ed.): *Advances in discourse approaches*. Cambridge Scholars Publishing, Cambridge, 147–173.

Sperber, Dan – Wilson, Deirdre 1986/1995. *Relevance: Communication and cognition*. Blackwell, Oxford.

Tátrai Szilárd 2011. *Bevezetés a pragmatikába. Funkcionális kognitív megközelítés*. Tinta Könyvkiadó, Budapest.

Thornborrow, Joanna 2007. Narrative, opinion and situated argument in talk show discourse. *Journal of Pragmatics* 39. 1436–1453.

Ting-Toomey, Stella (ed.) 1994. *The challenge of facework: Cross-cultural and interpersonal issues*. State University of New York, Albany.

Ting-Toomey, Stella – Oetzel, John G. 2003. Cross-cultural face concerns and conflict styles: Current status and future directions. In Gudykunst, William B. (ed.): *Cross-cultural and intercultural communication*. Sage, Thousand Oaks, 127–147.

Tracy, Karen 1990. The many faces of facework. In Giles, Howard – Robinson, W. Peter (eds.): *Handbook of language and social psychology*. Wiley, Chichester, 209–226.

van Rees, M. A. 2007. Discourse analysis and argumentation theory: The case of television talk. *Journal of Pragmatics* 39. 1454–1463.

Verschueren, Jef 1999. *Understanding pragmatics*. Arnold, London.

Wood, Linda A. – Kroger, Rolf O. 1994. The analysis of facework in discourse: Review and proposal. *Journal of Language and Social Psychology* 13. 248–277.

A tanulmány a Bolyai János kutatási ösztöndíj támogatásával készült.

AZ ISMÉTLÉSEK AUTOMATIKUS OSZTÁLYOZÁSA SPONTÁN BESZÉDBEN

Beke András – Gyarmathy Dorottya

Bevezetés

A beszédprodukció egyes részfolyamatai nagyrészt időben párhuzamosan mennek végbe. A beszélő az éppen aktuális közlésegység artikulációs megvalósítása közben már a következő gondolatsor fogalmi, grammatikai, fonológiai és artikulációs tervezését és megvalósítását végzi. A szimultán működés miatt a folyamat bármely szintjén keletkezhet zavar, amely a felszíni szerkezetben különféle diszharmonikus jelenségeket, összefoglaló néven megakadásjelenségeket eredményezhet. Ezek a spontán beszéd mintegy egyharmadát teszik ki (Gósy 2005). Funkcionális szempontból elkülöníthetők egymástól a bizonytalanságok (szünetek, ismétlések, újraindítások, nyújtások), amelyek a fogalmi és a nyelvi tervezés között fennálló nehézségekre utalnak, illetve a felszíni szerkezetben hibaként realizálódó téves kivitelezések (sorrendiségi hibák, téves kezdés, téves szótalálás stb.), amelyek a nyelvi tervezéstől az artikulációs kivitelezésig a beszédprodukciós folyamat bármely szintjén jelentkezhetnek.

Felolvasott beszéden az automatikus felismerés közel 90%-os eredménnyel működik (Furui 2003), a spontán beszédben ez az eredmény azonban szignifikánsan csökken. Ez részben a megakadásjelenségeknek köszönhető (Nakatani–Hirschberg 1994; Stouten–Martens 2003; Stouten et al. 2006). Adda-Decker és munkatársai (2003) igazolták, hogy a felismerési hiba (WER, word error rate) 12%-át a megakadásjelenségek okozzák. Stouten és Martens (2003) kimutatták, hogy ha a kitöltött szüneteket automatikusan osztályozzák a spontán beszédben, akkor a beszédfelismerő eredményei javulást mutatnak (a WER értéke 51,4%-ról 47,6%-ra csökkenthető). Mindezen eredmények igazolják, hogy amennyiben létezne egy olyan modell, amelyik képes felismerni a megakadásjelenségeket, a beszédfelismerő rendszerek hibázási aránya csökkenthető lenne (Reddy–Hasegawa-Johnson 2006).

A kitöltött szünetek, azaz a hezitálások osztályozására több eredményes nemzetközi és hazai kísérlet is született. A különféle metódusokkal működő beszédfelismerő rendszerek például a japán nyelvben 85-90%-os biztonsággal képesek osztályozni az egyes hezitációtípusokat (vö. Masataka et al. 1999; Wu–Yan 2002), míg magyar nyelven Beke és Horváth (2012) HMM-alapú osztályozó algoritmussal 98,33%-os felismerési arányt ért el. A jelen

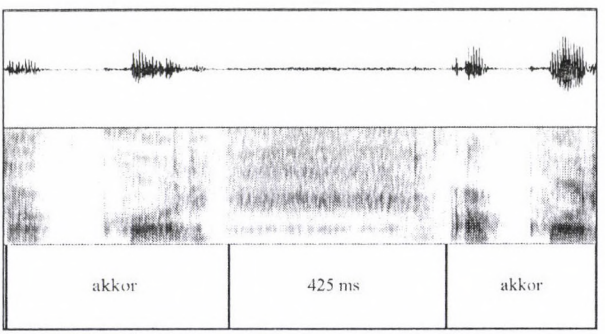
kutatás egy másik, a beszélő bizonytalanságára visszavezethető megakadásjelenség, az ismétlés automatikus osztályozására irányul.

Az ismétlés a hezitáláshoz hasonlóan nyelvtől függetlenül az egyik leggyakoribb, a spontán beszédben előforduló megakadásjelenség (vö. Clark–Wasow 1998; Gósy 2003; Bóna 2006; Gyarmathy 2006, 2007; Horváth 2007). A közlésben megjelenő ismétlések mint nem akaratlagos szóismétlések a beszélő bizonytalanságát jelzik (például: *és akkor ott ott ott ott elmondtam egy sztorit arról, hogy mi hogy indultunk; de de de nem tudtunk úgy szakmailag eszmét cserélni*). Az ismétlés sok esetben nem akasztja meg a beszéd artikulációs folyamatosságát, a beszélő ugyanis gyakran szünettartás nélkül ismétli meg az adott szót (Clark–Wasow 1998; Horga 2008; Gyarmathy et al. 2009; Gyarmathy 2009). Ez a megakadás a többi bizonytalansági megakadásjelenséghez hasonlóan a beszélő számára a mentális lexikonban való kereséshez, a grammatikai átalakításhoz, illetve a közlés tartalmi és formai felülvizsgálatához biztosít megfelelő időt. Clark és Wasow (1998) azzal magyarázza a jelenség létrejöttét, hogy a megismételt szó magasabb aktivációs szinttel rendelkezik, így a beszélő csak ezt képes újra kiejteni. Egy másik hipotézis szerint az artikuláció gyorsabban lezajlik, mint ahogy azt a kognitív folyamatok követni tudnák, ezért újraindul a már meglévő fonetikai terv megvalósítása, maga után vonva az adott szó ismételt kiejtését (Blackmer–Mitton 1991). Az ismétlések általában a funkciószavakat érintik (Fox–Jasperson 1995; Gyarmathy et al. 2009), ami magyarázható a spontán beszédbeli gyakoriságukkal, a mentális lexikonban való klisészerű tárolásukkal, de azzal is, hogy az ismétlés elsődleges funkciója az időnyerés, ezért a beszélő olyan szót választ, amely tartalmilag nem változtatja meg a közlést.

A jelenség automatikus felismertetése több problémát is felvet. Az egyik probléma az, hogy noha az ismétlés során a percepció számára azonos fonémasorok hangzanak el, a beszélő ugyanazt a lexikai egységet ismétli meg nagyon rövid időn belül, az első és a második kimondás akusztikai vetülete mégis különböző lehet (Benkenstein–Simpson 2003). A felolvasott beszéddel, illetve az izolált szavas ejtéssel ellentétben a spontán beszéd nagy variabilitást mutat, ami az automatikus felismerést különösen bonyolulttá teszi (vö. Beke–Szaszák 2010; Xie et al. 2004; Halpern 2006). A beszéd akusztikai vetülete még akkor is nagy különbségeket mutathat, ha ugyanaz a hangsor hangzik el kétszer egymás után rövid időn belül (Gósy 2009). Jól szemlélteti ezt az 1. ábra rezgés- és hangszínképe, ahol a beszélő az *akkor* szót ismételte meg 425 ms-os időkülönbséggel. A két kimondás akusztikai szerkezete láthatóan különbözik egymástól.

A másik probléma abban áll, hogy ugyan az ismétlések felismertetése elsőre egyszerűnek tűnhet – hiszen a felismerőnek csak ki kell választania a korpusz azon szavait, amelyek egymás után kétszer vagy többször fordulnak elő, és ezeket megakadásként kell azonosítania – bizonyos nyelvekben (pl. angol, német, holland) azonban – így a magyarban is – a helyzet ennél sokkal bo-

nyolultabb. Sok esetben egy adott szó megismétlése nem tekinthető megakadásnak, grammatikailag teljesen indokolt és korrekt pl. *az az ember vette meg az esernyőt; nagyon-nagyon örülök ennek; fel-feldobott kő.* Ezek az esetek azonban csak szemantikai elemzés alapján különíthetők el a megakadásoktól.



1. ábra
Az *akkor* (425 ms) *akkor* rezgés- és hangszínképe

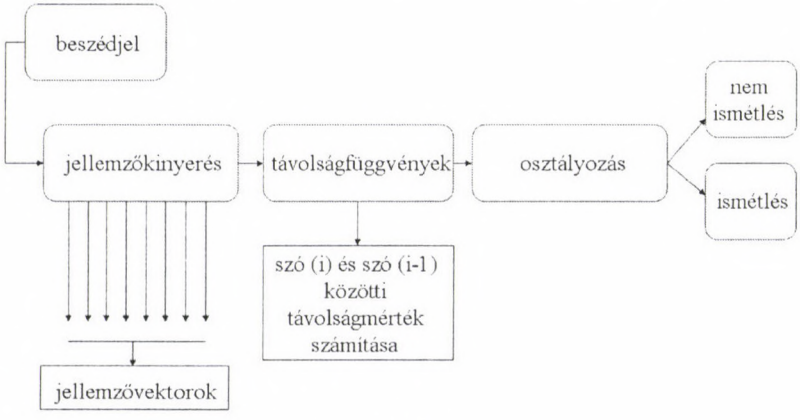
Az ismétlések automatikus osztályozására a külföldi szakirodalomban számos példát találunk. Léteznek akusztikai alapú (Ward 1991; Schriberg 1995; Plauche–Shriberg 1999), prozódiai alapú (Stolcke et al. 1999; Shriberg et al. 1997) és nyelvi alapú (Stolcke et al. 1999; Stolcke–Shriberg 1996; Liu et al. 2003) modellek. Az akusztikai alapú felismerők elsősorban az ismétlések két tagjának időtartamát, illetve a köztük lévő szünet hosszát veszik alapul. A prozódiai alapú osztályozók szóhatárokkal dolgoznak oly módon, hogy az egyes szavakból kivonják az akusztikai-prozódiai jellemzőket (pl. időtartam, szünet, normalizált F_0-érték), amelyek az osztályozás alapjául szolgáló döntési fák jellemzővektorait adják. A nyelvi modellek pedig a természetes beszédmegértést kívánják modellezni, amelynek érdekében a tanító adatbázisban előforduló megakadásokat címkézik, majd a felismerés során ezen szósorozatok valószínűségét kívánják megbecsülni (Rangarajan–Narayanan 2006). A felismerés eredményességét több tényező befolyásolja, úgy mint az adott nyelv, a beszéd minősége, a hangminőség, illetve az ismételt szavak komplexitása. Ravikumar és munkatársai (2008) olvasott beszéden alapuló felismerő rendszere 83%-os eredményességgel azonosítja a szótagismétléseket, lehetővé téve ezáltal a dadogó beszélők objektív osztályozását.

A spontán beszédben azonban romlik az ismétlések felismerése, így például Shriberg és munkatársai (1997) prozódiaalapú felismerő rendszere mindössze 77,5%-os pontossággal bír. A magyar nyelvre ez idáig még nem készült a spontán beszéd ismétléseinek felismerését lehetővé tevő osztályozó algoritmus. A jelen kutatás fő kérdése az volt, hogy az ismétlések (kivéve a

nem megakadásjelenség funkciójú ismétlés) osztályozhatók-e automatikusan, és ha igen, milyen eredménnyel. Feltételeztük, hogy ez a jelenség jó eredménnyel osztályozható spektrális akusztikai paraméterekkel, és szupport vektor géppel.

Kísérleti személyek, anyag, módszer

A BEA adatbázisból 7 beszélő spontán narratíváját választottuk ki (2 férfi és 5 nő, átlagéletkoruk 25 év). A hangfájlokat szószinten kézzel annotáltuk a Praat programban (Boersma–Weenink 2009). Az annotálás során jelöltük azon szavakat, amelyek felszíni megakadásként ismétlődtek, tehát a megakadásjelenségnek tekinthető ismétléseket vettük figyelembe, a grammatikailag, illetve pragmatikailag indokoltakat nem. Összesen mintegy 4 órányi spontán beszédben 280 db ismétlést címkéztünk (Praat 5.1.34). Az ismétlések osztályozásakor az akusztikai jelből jellemzőket vontunk ki (l. a *Jellemzőkinyerés* című részben), majd egy szó és az előtte álló szó jellemzői között kiszámítottuk a köztük levő távolságmértéket (l. *A hasonlóság mérése* című részben), majd ezt az értéket használtuk fel az osztályozáskor (2. ábra).



2. ábra
Az ismétlések osztályozásának sematikus folyamatábrája

Jellemzőkinyerés

A beszédtechnológiai kutatások többsége a rövid idejű spektrális burkológörbe érzeti transzformációján alapuló eljárásokat alkalmazza (MFCC: Mel Frequency Cepstral Coefficients; PLP: Perceptual Linear Prediction) (Mermelstein 1976; Hermansky 1990). A Mel-frekvenciás kepsztrális együtthatókra történő átalakításon alapuló eljárást (MFC) igen széles körben használják; többek között ennek segítségével modellezik az emberi hallást is. Lé-

nek azonban más, például az emberi percepció transzformációján alapuló (PLP) lényegkinyerő algoritmusok is. A jelen kutatásban mindkét akusztikai jellemzőt alkalmaztuk. Mind az MFC, mind a PLP kinyerésekor különböző beállításokat használtunk. Elsősorban különböző számú együtthatókat számoltunk; először 12, majd 24 koefficienst tartalmazó MFC-t és PLP-t számítottunk. Mindkét esetben megvizsgáltuk, hogy az energia és a delták (delta: a jellemző változás mértéke, delta-delta: a változás sebessége) hozzáadása javít-e az osztályozás eredményén. Mindezeknek megfelelően összesen 8 különféle jellemzőkinyerést hajtottunk végre.

Lényegkiemelés

Az egyes szavakból kinyert akusztikai tulajdonságvektorokon lényegkiemelést végeztük. Ez azt jelenti, hogy a jellemzőkre Gauss-keverék modell (GMM: Gaussian Mixture Model) segítségével megbecsüljük azok eloszlásparamétereit. A GMM univerzális eloszlásbecslő (függvényapproximátor), jól kezelhető, nem igényel komplex számítást, mégis pontosan megbecsülhetők vele a függvény paraméterei, ami alatt a haranggörbe (Gauss-görbe) paramétereinek becslését értjük. A jelen esetben a GMM-et mint lényegkiemelő eljárást alkalmaztuk. Az adatok k-komponensű Gauss mintaátlagainak a kezdeti értékeit a k-means (k-közép) klaszterezési algoritmus alapján választottuk meg.

A Gauss-keverék modell paramétereinek, mint például a mintaátlag, tapasztalati szórás, kovariancia mátrix becslésére számos algoritmus létezik. Ezek közül a legtöbbször alkalmazott eljárás a maximum likelihood criterion (a legnagyobb valószínűség elve), amelyet az iteratív Expectation-Maximization (EM) (Dempster et al. 1977; McLachlan–Krishnan 1997) algoritmussal szokás megvalósítani. Általában kevesebb mint 10 iteráció elégséges az EM algoritmusnak ahhoz, hogy elérje a paraméterek elégséges konvergenciáját. A GMM optimalizálása után kapott mintaátlagokat nyertük ki, amelyek egy-egy szó akusztikai képét reprezentálták.

A hasonlóság mérése

Az i-edik szó és az i–1-edik szóra kapott lényegkiemelt jellemzővektorok között távolságokat mértünk különböző távolságfüggvényekkel. A beszélőfelismerésben általánosan használt módszer az egyes beszélői minták hasonlóságának mérésére a mértani távolság alkalmazása (Atal 1974; Markel–Davis 1979; Schwartz et al. 1982). Ennek számolására számtalan módszer létezik. A gyakorlatban legtöbbször használt távolságfüggvény az euklidészi távolság és a négyzetes euklidészi távolság. Egy másik lehetséges módszer a Kullback–Leibler-távolság , amely a két eloszlás között fellépő távolságot írja le (Cambell 1997; Bishop 2006). A harmadik eljárás a dinamikus idővetemítés (DTW) (Sakoe–Chiba 1978), ami különböző hosszúságú vektorok közötti távolság mérésére alkalmazható. A negyedik távolságfüggvény, amelyet a jelen tanulmányban használtunk, a Bhattacharyya-távolság (Kailath 1967), amely a két eloszlás között fellépő távolságot írja le. Ezen távolságértékek adták a négyelemű tulajdonságvektorát az ismételt és a nem ismételt szavak-

nak. Ezek a jellemzők adták a bemenetét az ismételt és nem ismételt szó osztályozónak.

Osztályozás

Az ismételt és a nem ismételt szavak osztályozásához szupport vektor gépet használtunk. Az SVM (Support Vector Machine) a felügyelt tanulási módszerek családjába tartozik, célja egy olyan szeparáló hipersík keresése, amely jól választja el egymástól a két osztály elemeit (lehet többosztályos is). Az SVM-ek működésének lényege, hogy az eredeti megfogalmazásában még komplex nemlineáris megoldást igénylő feladatot, azaz a feladatból származó mintákat nemlineáris transzformációk segítségével egy, a bemeneti mintatér dimenziójánál több dimenziós térbe transzformálja, ahol az már lineárisan megoldható. A módszer egyik legnagyobb előnye, hogy egy garantált felső korlátot ad az approximáció általánosítási hibájára. Egy másik fontos jellemzője, hogy a tanulási algoritmus törekszik a modell méretének minimalizálására (ritka modellt alkot), ami a hiba rovására történik, de mértéke egy paraméterrel szabályozható (Schölkopf et al. 1999). A hagyományos SVM alkalmazásának legnagyobb akadálya a módszer nagy algoritmikus komplexitása és a nagy memóriaigény, ami tipikusan a nagy adatmennyiség kezelését teszi lehetetlenné. A problémára számos megoldás született. Ezek az algoritmusok többnyire iteratív megoldások, melyek a nagy optimalizálási feladatot kisebb feladatok sorozatára bontják. A nemlineáris osztályozáshoz a legelterjedtebb függvény a radiális bázis függvény alkalmazása. Ekkor az SVM-nek két szabadon állítható paramétere van. Ezen szabad paraméterek értékeinek optimalizálásra Hsu és munkatársai (2003) tettek javaslatot. Az eljárásban tehát a C a hibázási paramétert (penalty parameter), és γ-t az RBF kernel függvény (gaussi függvény) szórásparaméterének értékét keressük. Az SVM optimalizálását kizárólag a tanító halmazon végezzük keresztvalidációs eljárással. A keresztvalidációt használva az adathalmazt *i* részre osztjuk. Ezután *i* lépésben alkalmazzuk a teszt- és tanítóhalmazok módszerét. Minden lépésben egy résszel mint teszthalmazzal, a többi k–1 résszel, mint tanító halmazzal értékeljük az adott paraméterekkel rendelkező SVM működését. Keresztvalidációval a teljesítmény mértéke:

$$P_M = \frac{1}{i}\sum_{j=1}^{i} p_j ,$$

ahol p_j az *i*. körben a teszthalmazon a hiba. A módszer előnye, hogy kevésbé számít, hogyan választjuk szét az adatokat. Az SVM C és gamma paramétereit Hsu és munkatársai által javasolt térben érdemes keresni, ami a következő: C = $\{2^{-5}; 2^{-3}; \ldots; 2^{13}; 2^{15}\}$; a γ: $\{2^{-15}; 2^{-13}; \ldots; 2^{1}; 2^{3}\}$. Így a háromszoros keresztvalidáció mellett 3 × 11 × 11 = 363 tanítás-tesztelést végeztünk. Az optimális C és gamma paraméter páros a legjobb osztályozás értékénél volt.

Kiértékelés

Az osztályozás eredményességét sokféleképpen mérhetjük (Ian et al. 2011). Bináris osztályozás esetén a tévesztési mátrix a következő (1. táblázat):

1. táblázat: Tévesztési mátrix bináris osztályozás esetén

		Aktuális feltétel	
		Pozitív	**Negatív**
A teszt eredménye	**Pozitív**	A feltétel teljesül + pozitív teszt = TP (True Positives)	A feltétel nem teljesül + pozitív teszt = FP (False Positives)
	Negatív	A feltétel teljesül + negatív teszt = FN (False Negatives)	A feltétel nem teljesül + negativ teszt = TN (True Negatives)

Az osztályozó teljesítményének mérésére számos mérőszám létezik: Accuracy = (TP + TN)/(TP + TN + FP + FN), vagyis a helyesen felismert adatok aránya; sensitivity = recall = TP/(TP + FN), vagyis a pozitívak helyesen felismert aránya; specificity = TN/(FP + TN), vagyis a negatívak helyesen felismert aránya; precision = TP/(TP + FP), vagyis a pozitívként felismertek hányad része volt valóban pozitív. Ezek mellett jó mérőszám az F-measure: 2*Precision*Recall/(Precision + Recall) a harmonikus közepe a precison és recall értékeknek.

Az SVM optimalizálásakor az accuracy (felismerési arány: itt osztályozási arány) és az F-measure (F-mérték) értékeket vettük figyelembe.

Eredmények

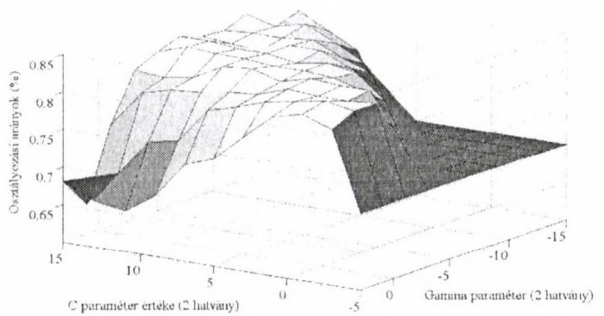
Az ismétlések osztályozására MFCC-t és PLP-t használtunk akusztikai jellemzőként. Mindkét esetben külön tanítottunk, illetve teszteltünk 12 és 24 koefficienst tartalmazó és energiát, deltát, delta-deltát tartalmazó, valamint nem tartalmazó akusztikai jellemzőkkel.

A 12 koefficiensből álló MFCC esetén akkor értük el a legjobb SVM beállítást, ha a gamma és a C értéke is 2^{-1} volt (3. ábra). A teszteléskor az osztályozás eredménye 81,60%-volt, ekkor az F-érték 79,06% volt.

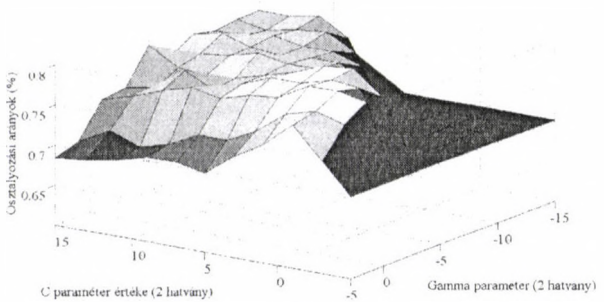
Ha a 12 koefficiens mellé kiszámoltuk az energiát és az első két deriváltat, akkor az osztályozó legjobb értéke akkor adódott, ha a gamma értéke 2^{-1}, a C értéke 2^{1} volt (4. ábra). A teszteléskor 71,93%-os volt az osztályozás értéke, míg az F-érték 73,83% volt.

Az MFC koefficienseinek számát 24-re növelve javulást értünk el. Az optimalizálás során a gamma = 2^{1}, C = 2^{-1} volt (5. ábra). A tesztelés során kapott eredmény 82,87%-volt, az F-érték 79,07% volt.

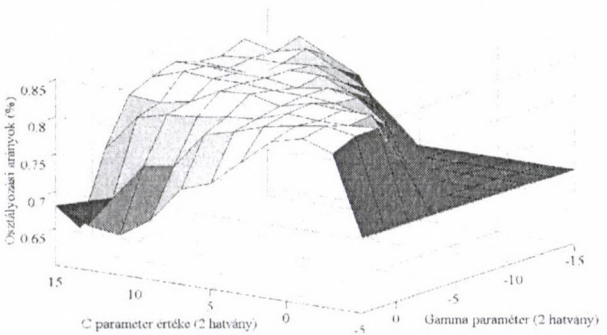
Az 24 MFC mellé számított energiával és deltákkal a rendszer eredménye romlott (gamma = 2^{-3}; C = 2^{-1}). Az osztályozás eredménye jóllehet magas, 80,33%, de az F-érték igen alacsony, 73,44%.



3. ábra
Az SVM optimalizálása 12 MFC esetén

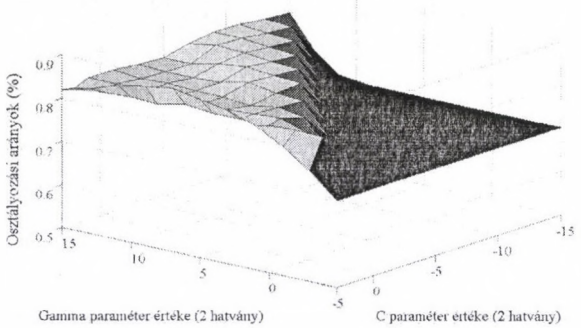


4. ábra
Az SVM optimalizálása 12 MFC + Δ + ΔΔ esetén



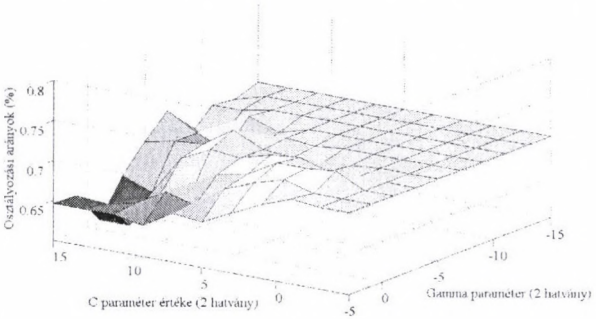
5. ábra
Az SVM optimalizálása 24 MFC esetén

A kutatás során akusztikai jellemzőként kimértük a PLP-t is. A 12 PLP-vel optimalizált rendszer a legjobb eredményt akkor adta, ha a gamma = 2^{-3}, C = 2^{15} (6. ábra). A teszteléskor az osztályozás eredménye 87,77%-volt, az F-érték pedig 82,57%. Ez javulást jelent a 12 MFC-hez és a 24 MFCC-hez képest is.



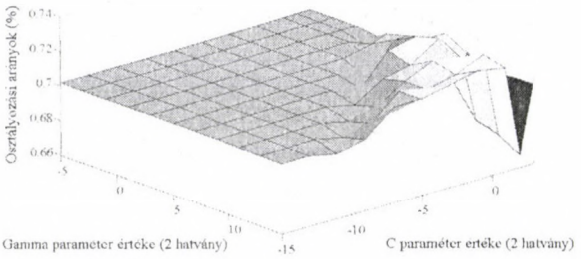
6. ábra
Az SVM optimalizálása 12PLP esetén

Ha a 12 PLP mellé kiszámoljuk az energiát és a deriváltakat, akkor az eredmények – hasonlóan az MFC-hez – romlanak (gamma = 2^{-1}, C = 2^{-1}). A teszteléskor elért osztályozási arány 71,93%, míg az F-érték 68,50% volt (7. ábra).



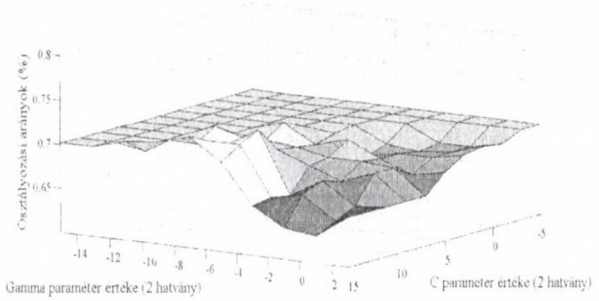
7. ábra
Az SVM optimalizálása 12 PLP + Δ + ΔΔ esetén

A 24 PLP-vel tanított rendszerünk jóval elmaradt a többi paraméterrel működő rendszerétől. Az SVM paraméterei a legjobb eredménynél a következők voltak: gamma = 2^{3}, C = 2^{7}. A helyes osztályozás aránya 72,29%, míg az F-érték 69,34% volt (8. ábra).



8. ábra
Az SVM optimalizálása 24 PLP esetén

A 24 PLP-vel, az energiával és a deriváltakkal működő rendszer közel azonos eredményt adott, mint a 24 PLP. A tanítás során az SVM akkor adta a legjobb eredményt, ha a gamma értéke 2^3 volt, míg a C értéke 2^3. A tesztelés során a helyes osztályozási arány 72,62%, míg az F-érték 69,74%-volt (9. ábra).



9. ábra
Az SVM optimalizálása 24 PLP + Δ + ΔΔ esetén

Az osztályozók összehasonlítása

Az összes kialakított osztályozó közül a 12 koefficienst tartalmazó PLP-vel értük el a legjobb eredményt. Mind az osztályozás aránya, mind az F-érték ennél az osztályozónál volt a legmagasabb. Az MFCC esetében akkor értük el a legjobb eredményt, ha 24 koefficienst alkalmaztunk bemeneti paraméterként. Mind a PLP, mind az MFCC esetében rontott az eredményeken, ha hozzáadtuk az energiát és a deltákat (2. táblázat).

2. táblázat: Az osztályozók eredményei

A koefficiens száma		Kiértékelés	MFC	PLP
12	-	Rec	81,60%	**87,77%**
		F-érték	79,06%	**82,57%**
	e, Δ, ΔΔ	Rec	71,93%	71,93%
		F-érték	73,83%	68,50%
24	-	Rec	**82,87%**	72,29%
		F-érték	**79,07%**	69,34%
	e, Δ, ΔΔ	Rec	80,33%	72,62%
		F-érték	73,44%	69,74%

Következtetések

A jelen kutatásban azt vizsgáltuk, hogy milyen eredménnyel lehet automatikusan osztályozni az ismétléseket a spontán beszédben. Az ismétlések osztályozásához akusztikai jellemzőként 12, illetve 24 koefficienst tartalmazó MFCC-t és PLP-t használtunk, az energiával, deltákkal együtt és azok nélkül.

Lényegkiemelésként Gauss-keverék modellt alkalmaztunk, hogy az egyes jellemzőkre megbecsüljük azok eloszlásparamétereit. Az egyes szavakat ezek a tulajdonságvektorok jellemzik. Az ismétlések osztályozásakor egy szó és az előtte álló szó közötti hasonlóságot kell mérnünk. A jelen vizsgálatban ennek mérésére euklidészi távolságot, négyzetes euklidészi távolságot, Kullback–Leibler-távolságot, DTW-t, Bhattacharyya-távolságot alkalmaztunk. Így megkaptuk a két szó akusztikai jellemzői közötti távolságmértékeket. Ezeket a távolságmértékeket használtuk az SVM osztályozó bemeneti paramétereként. Az SVM-et a tanító halmazon optimalizáltuk keresztvalidációval és kimerítő kereséssel.

A legjobb eredményt a 12 koefficienst tartalmazó PLP-vel értük el, amelynek helyes osztályozási értéke 87,77% volt, míg az F-értéke 82,57%. Az eredmények azt mutatták, hogy sem az energia, sem a delták nem javítanak az osztályozási arányon, ami azzal magyarázható, hogy számos esetben igen rövidek a beszédminták (pl. névelők, két-három hangból álló kötőszók ismétlésekor).

Az eredmények tehát igazolták hipotézisünket, miszerint az ismétlések jó osztályozási eredménnyel detektálhatók a spontán beszédben. Az eredményeink javítására újabb kísérletet tervezünk, amely nagyobb korpuszt tartalmaz.

Irodalom

Adda-Decker, Martine – Habert, Benoît – Barras, Claude – Adda, Gilles – Boula de Mareuil, Philippe – Paroubek, Patrick 2003. A disfluency study for cleaning spontaneous speech automatic transcripts and improving speech language models. In: *Proceedings of ITRW on Disfluency in Spontaneous Speech.* Göteborg, 67–70.

Atal, Bishnu S. 1974. Effectiveness of linear prediction characteristics of the speech wave for automatic speaker identification and verification. *Journal of the Acoustical Society of America* 55/6. 1304–1312.

Beke András – Horváth Viktória 2012. A hezitációs jelenségek gépi osztályozása a spontán beszédben. (Megjelenés alatt.)

Beke, András – Szaszák, György 2010. Automatic recognition of schwa variants and stress syllables in spontaneous Hungarian speech. *Acta Linguistica Hungarica* 57. 229–253.

Benkenstein, Ramona – Simpson, Adrian P. 2003. Phonetic correlates of self-repair involving word repetition in German spontaneous speech. In: *Proceedings of DiSS '03. Disfluency in Spontaneous Speech Workshop, 5-8 September 2003. Göteborg University, Sweeden.* Robert Eklióund (ed.): *Gothenburg Papers in Theoretical Linguistics 90.* 81–84.

Bishop, Christopher M. 2006. *Pattern recognition and machine learning.* Springer Science & Business Media, New York.

Blackmer, Elisabeth R. – Mitton, Janet L. 1991. Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition* 39. 173–194.

Boersma, Paul – Weenink, David 2009. Praat: doing phonetics by computer (Version 5.1.34) [Computer program]. http://www.praat.org

Bóna Judit 2006. A megakadásjelenségek akusztikai és percepciós sajátosságai. *Beszédkutatás 2006.* 101–113.

Campbell, William M. – Campbell, Joseph P. – Reynolds, Douglas – Singer, Elliot – Torres-Carrasquillo, Pedro A. 2006. Support Vector Machines for speaker and language recognition. *Computer Speech and Language* 20/2–3. 10–29.

Hsu, Chih-Wei – Chang, Chih-Chung – Lin, Chih-Jen 2003. *A practical guide to support vector classification.* Technical report. Department of Computer Science, National Taiwan University. http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf

Clark, Herbert H. – Wasow, Thomas 1998. Repeating words in spontaneous speech. *Cognitive Psychology* 37. 201–242.

Fox, Barbara - Jasperson, Robert 1995. A syntactic exploration of repair in English conversation. In Davis, Philip W. (ed.): *Alternative linguistics: Descriptive and theoretical modes.* Current Issues in Linguistic Theory 102. Benjamins, Amsterdam, 77–134.

Furui, Sadaoki 2003. Recent advances in spontaneous speech recognition and understanding. In: *Proceeding of IEEE Workshop on Spontaneous Speech Processing and Recognition.* Tokyo, 1–6.

Gósy Mária 2003. A spontán beszédben előforduló megakadásjelenségek gyakorisága és összefüggései. *Magyar Nyelvőr* 127. 257–277.

Gósy Mária 2005. *Pszicholingvisztika.* Osiris Kiadó, Budapest.

Gósy Mária 2009. Szóejtés és szóészlelés: változatosság és adaptálódás. *Beszédkutatás 2009.* 46–75.

Gyarmathy Dorottya 2006. A beszédpercepció és beszédprodukciós folyamat összefüggései a megakadásjelenségek tükrében. In Heltai Pál (szerk.): *MANYE XVI. Tanulmánykötet.* Vol. 3/2. 449–455.

Gyarmathy Dorottya 2007. Az alkohol hatása a spontán beszédprodukcióra. *Beszédkutatás 2007.* 108–121.

Gyarmathy Dorottya 2009. A beszélő bizonytalanságának jelzései: ismétlések és újraindítások. *Beszédkutatás 2009.* 196–217.

Gyarmathy Dorottya – Gósy Mária – Horváth Viktória 2009. A rejtett és a felszíni önmonitorozás temporális jellemzői. In Keszler Borbála – Tátrai Szilárd (szerk.): *Diskurzus a grammatikában – grammatika a diskurzusban.* Tinta Kiadó, Budapest, 46–55.

Halpern, Jack 2006. The contribution of lexical resources to natural language. In *Processing of CJK Languages. ISCSLP 2006.* 768–780.

Horga, Damir 2008. Repetitions in interrupted speech production. *Beszédkutatás 2008.* 157–171.

Horváth Viktória 2007. Vannak-e „női" és „férfi" megakadásjelenségek a spontán beszédben? *Magyar Nyelvőr* 131. 315–323.

Kailath, Thomas 1967. The divergence and Bhattacharyya distance measures in signal selection. In: *Proceedings of the Institute of Electrical and Electronic Engineers, Transactions on Communication Technology* 15/1. 52–60.

Liu, Yang – Shriberg, Elizabeth – Stolcke, Andreas 2003. Automatic disfluency identification in conversational speech using multiple knowledge sources. In: *Proceeding of EUROSPEECH.* Geneva, 957–960.

Markel, John D. – Davis, Steven B. 1979. Text-independent speaker recognition from a large linguistically unconstrained time-spaced data base. In: *Proceedings of the Institute of Electrical and Electronic Engineers, International Conference on Acoustics* 27/1. 74–82.

Masataka, Goto – Katsunobu, Itou – Satoru, Hayamizu 1999. A real-time system detecting filled pauses for spontaneous speech. In: *Proceedings of. European Conference Speech Communication and Technology.* 227–230.

Nakatani, Christine H. – Hirschberg, Julia 1994. A corpusbased study of repair cues in spontaneous speech. *Journal of the Acoustical Society of America* 95/3. 1603–1616.

Plauche, Madelaine C. – Shriberg, Elisabeth 1999. Data-driven subclassification of disfluent repetitions based on prosodic features. In: *Proceedings of International Congress of Phonetic Sciences.* Vol. 2. San Francisco, 1513–1516.

Rangarajan, Vivek – Narayanan, Shrikanth 2006. Analysis of disfluent repetitions in spontaneous speech recognition. In: *Proceedings of the European Signal Processing Conference (EUSIPCO).* http://www.eurasip.org/Proceedings/Eusipco/Eusipco2006/papers/1568981801.pdf

Ravikumar, K. – Reddy, Balakrishna – Rajagopal, R. – Nagaraj, H. 2008. Automatic detection of syllable repetition in read speech for objective assessment of stuttered disfluencies. In: *Proceedings of World Academy Science, Engineering and Technology.* 270–273.

Reddy, Rajiv M. – Hasegawa-Johnson, Mark A. 2006. Analysis of pitch contours in repetition-disfluency using stem-ML. http://cogcomp.cs.illinois.edu/sites/mclc/finalPapers/rreddy_AT_uiuc.edu__reddy_mclc2006.pdf

Sakoe, Hiroaki – Chiba, Seibi 1978. Dynamic programming algorithm optimization for spoken word recognition. In: *Proceedings of the Institute of Electrical and Electronic Engineers, Transactions on Acoustics, Speech and Signal Processing,* 26/1. 43–49.

Schölkopf, Bernhard – Burges, Christopher J. C. – Smola, Alexander J. (eds.) 1999. *Advances in Kernel Methods: Support Vector Learning.* MIT Press, Cambridge.

Schwartz, Richard – Roucos, Salim – Berouti, M. 1982. The application of probability density estimation to text independent speaker identification. In: *Proceedings of International Conference Acoustics, Speech, and Signal Processing*. Paris, 1649–1652.

Shriberg, Elisabeth 1995. Acoustic properties of disfluent repetitions. In: *Proceedings of International Congress of Phonetic Sciences*. Stockholm, 384–387.

Shriberg, Elisabeth – Bates, Rebecca – Stolcke, Andreas 1997. A prosody-only decision-tree model for disfluency detection. In *Proceedings of EUROSPEECH*. Vol. 5. 2383–2386.

Stolcke, Andreas – Shriberg, Elizabeth 1996. Statistical language modeling for speech disfluencies. In: *Proceedings of ICASSP*. Vol. 1. Atlanta, GA, 405–409.

Stolcke, Andreas – Shriberg, Elizabeth – Hakkani-Tür, Dilek – Tür, Gökhan 1999. Modeling the prosody of hidden events for improved word recognition. In: *Proceedings of EUROSPEECH*. Vol. 1. Budapest, 307–310.

Stouten, Frederik – Martens, Jean-Pierre 2003. A feature-based filled pause detection technique for dutch. In: *IEEE Intl. Workshop on ASRU*. 309–314.

Stouten, Frederik – Duchateau, Jacques – Martens, Jean-Pierre – Wambacq, Patrick 2006. Coping with disfluencies in spontaneous speech recognition: Acoustic detection and linguistic context manipulation. *Speech Communication* 48. 1590–1606.

Xie, Huayang – Andreae, Peter – Zhang, Mengjie – Warren, Paul 2004. Detecting stress in spoken english using decision trees and Support Vector Machines. *ACSW Frontiers 2004*. 145–150.

Witten, Ian H. – Frank, Eibe – Hall, Mark A. 2011. *Data mining: Practical machine learning tools and techniques*. 3rd Edition. Morgan Kaufmann, San Francisco.

Wu, Chung-Hsien – Yan, Gwo-Lang 2004. Acoustic feature analysis and discriminative modeling of filled pauses for spontaneous speech recognition. *Journal of VLSI Signal Processing Systems* 36. 91–104.

Ward, Wayne 1991. Understanding spontaneous speech: The phoenix system. In *Proceedings ICASSP*. 365–367.

A kutatás a 78315. sz. OTKA keretében folyt.

THE PRODUCTION OF THE AMERICAN ENGLISH LOW FRONT VOWEL BY BILINGUAL SPANISH- AND ENGLISH-SPEAKING CHILDREN AND THEIR MONOLINGUAL PEERS

Ferenc Bunta – Michelle Norton

Introduction

Vowel acquisition in bilingual and monolingual children has received less attention than the acquisition of consonants, because the former are generally claimed to be established earlier than the latter (cf. Goldstein–Cintrón 2001). By 36 months of age, monolingual English-speaking children become highly accurate in their vowel productions (Donegan 2002; Vorperian–Kent 2007). Bilingual phonological research has also found that vowels are produced relatively early and generally more accurately than consonants (Gildersleeve-Neumann et al. 2008; Goldstein–Cintrón 2001); however, certain American English vowels (most notably lax and rhotacized ones) are acquired later than the rest of the English vowel system. One particular vowel that bilingual speakers may have difficulty acquiring is the English low front vowel, /æ/, because it is a linguistically marked vowel (see Archangeli–Langendoen 1997) that is known to pose challenges to some non-native speakers of English (Bunta–Major 2004; Jia et al. 2006).

The purpose of the present study is to investigate how bilingual Spanish- and English-speaking four- to five-year-old children and their monolingual age-matched English-speaking peers produce a linguistically marked American English vowel, /æ/, that is also missing from the Spanish vowel inventory. Acoustic measurements of duration and the first two formants were used to evaluate the production of the vowel in stressed position and compare the productions of bilingual children to those of their monolingual peers to gain insights into the structural organization of the phoneme.

(i) Spanish and English vowel inventory differences

The size and structure of vowel inventories have significant and wide-ranging effects on how bilingual individuals acquire vowels in their target languages (see Flege 1989; Gildersleeve-Neumann et al. 2008; Iverson–Evans 2007; Levey–Cruz 2004). Iverson and Evans (2007) found that speakers of languages with relatively large vowel inventories collapsed fewer English vowels than their peers whose languages have relatively smaller and less complex vowel systems. Native speakers of Norwegian and German – both

having relatively large vowel inventories – were more accurate in recognizing English vowels, while native speakers of Spanish – whose vowel inventory was much smaller – displayed collapsed English vowel categories and overall lower accuracy rates (Iverson–Evans 2007). Furthermore, Flege (1989) noted that the size of an individual's native language vowel inventory does affect the placement of the vowels in the other language, but vowel inventory size does not have an effect on the precision of tongue positioning.

A comparison of the Spanish and English vowel inventories reveals important differences between the vowel systems of the two languages. The Spanish inventory consists of 5 vowels: /i e a o u/ as opposed to the American English vowel inventory that contains up to 15 vowels (/i ɪ ɛ æ ʌ ɚ ɑ ɔ ʊ u ɔɪ aɪ aʊ eɪ oʊ/, according to Gildersleeve-Neumann et al. 2008). While there can be dialectal variation in the vowel systems of both languages (especially in English), the English vowel inventory is much larger than its Spanish counterpart by any measure. In addition, unlike Spanish vowels, English ones can also undergo significant vowel reduction in unstressed syllables (Dauer 1983), an issue that is not explored in the present study, but one that further complicates speech acquisition in Spanish- and English-speaking bilinguals. Moreover, some English vowels (for example /æ/) are markedly different from Spanish ones, having no obvious targets for substitution. Consequently, Spanish- and English-speaking bilingual children have to acquire two very different vowel systems and this is likely to affect how they produce vowel sounds compared to their monolingual English-speaking peers. One of the most challenging vowels for bilingual Spanish- and English-speaking bilinguals to acquire is the low front lax American English vowel (Levey–Cruz 2004), which makes this vowel a good candidate for investigation.

(ii) Vowel acquisition by monolinguals and bilinguals with a focus on /æ/

In his description of English vowel development during the first four years of life, Kent (1992) noted that [æ] occurred among the early "vocants", albeit these early vocalizations were non-phonemic. Notably, [æ] disappeared from children's productions during the second year even though it had been prominent in babbling, and the low front vowel resurfaced as a stable phoneme only in the fourth year of life in monolingual English-speaking children. Davis and MacNeilage (1990) also found that [æ] occurred in early disyllables and was strongly favored in babbling, but its accuracy in early words produced was not very high, indicating that while /æ/ may occur in early productions, it may not have a stable phonemic representation in early segmental acquisition.

In linguistic terms, /æ/ is also often characterized as a marked segment, as argued by Archangeli and Langendoen (1997), who noted that the preferred location of low vowels was in the back. The authors named this the LOWBACK markedness constraint, favoring a [+low] vowel to be associated

with the [+back] feature. Taken together, it appears that while the American English low front vowel occurs as early as during babbling, it is a vowel that takes a relatively long time to acquire in a stable manner, compared to other English vowels.

Studies also indicate that Spanish- and English-speaking bilingual children have difficulty with acquiring lax and rhotic vowels (Gildersleeve-Neumann et al. 2008; Goldstein–Cintrón 2001). This is not surprising, considering that research on bilingual child phonological acquisition found evidence of cross-language interaction (cf. Fabiano-Smith–Goldstein 2010b; Goldstein–Bunta in press). It is a well-known fact that bilingual individuals are not two monolinguals in one (Grosjean 1989), and while proficient and early bilinguals may have separate phonological systems for their target languages, interaction between the sound systems of those languages is likely to occur (cf. Paradis–Genesee 1996). Interaction (or interdependence) has been more recently been defined in terms of transfer, acceleration (faster rate of acquisition by bilinguals than their monolingual peers), and deceleration (slower rate of acquisition of some phonological phenomena in bilinguals than monolinguals) (Fabiano-Smith–Goldstein 2010b). All of these general phenomena can be observed in the vowel acquisition of bilinguals as compared to their monolingual peers.

In a study on vowel perception by Spanish- and English-speaking bilinguals and their monolingual English-speaking peers, Levey and Cruz (2004) found that both monolinguals and bilinguals had difficulty with differentiating /ɑ/ versus /ɔ/ and /æ/ versus /eɪ/, but there was also an interaction between bilingualism and certain vowel contrasts, suggesting that bilinguals had more difficulty than their monolingual English-speaking peers when it came to the /ɑ/ - /ɔ/ , /æ/ - /eɪ/, /ɪ/ - /i/, and /u/ - /ʊ/ contrasts. Furthermore, Spanish- and English-speaking bilinguals are not the only bilinguals who have difficulty acquiring the English low front vowel; Hungarian-English and Chinese-English bilinguals also have problems producing /æ/ (Bunta–Major 2004; Jia et al. 2006).

Flege (1991) also noted that /æ/ and /ɪ/ proved to be the most challenging vowels to identify correctly for bilingual Spanish-English and monolingual Spanish listeners. Even some of the more experienced English as a second language learners identified /æ/ as the Spanish "a" most of the time. This indicates that some of the "difficult" and novel English phonemes may be mapped onto native categories and assimilated in such a way that they are not perceived (or produced) as novel sounds but as variants of existing native Spanish vowels (Højen–Flege 2006). Flege (1995) called the phenomenon of identifying second language segments as native phonemes *equivalence classification* whereby the forming of new non-native phoneme categories are blocked by native language phonemes if the new categories are sufficiently similar to existing ones. Evidence of identifying novel sounds as native

language phonemes has been found by a number of studies on bilingual phonology (Højen–Flege 2006; Mora–Keidel–Flege 2011).

Alternatively, bilinguals may overcompensate and produce a novel sound in an exaggerated fashion in an effort to make the new sound category markedly different from existing ones (e.g., Italian learners of English overshooting /eɪ/, Flege–Schirru–MacKay 2003). Guion (2003) found that Quichua- and Spanish-speaking bilinguals raised some of their native language vowels to maintain separation between the two vowel inventories. Furthermore, Oh et al. (2011) also noted that early learners of English, whose native language was Japanese, gradually deflected the Japanese /a/ away from the English /ʌ/ over the course of one year to maintain differentiation between the two phonemes. However, the authors also found that the Japanese and English /i/ categories moved closer to each other, based on acoustic measurements of the first two formants.

Cross-language interaction of specific phonological phenomena may occur even if overall separation of the phonological systems is evident in bilinguals (cf. Fabiano-Smith–Barlow 2010). Interestingly, while interaction appears to be unidirectional in late bilinguals (i.e., the first language has various effects on the later acquired language), interaction in simultaneous and early bilinguals seems to be bi-directional; each language affecting the other (Baker–Trofimovich 2005; Oh et al. 2011). The implications of possible interaction in Spanish-English bilinguals regarding the English vowel, /æ/, are that both the duration and the formants may be affected. One possible manifestation of cross-language interaction could be that bilingual Spanish- and English-speaking children would produce the English /æ/ as a shorter vowel than their monolingual English-speaking peers would, because as Delattre (1966) noted, English vowels are generally longer than their Spanish counterparts, and this applies to /æ/ that tends to be a relatively long vowel in English (Peterson–Lehiste 1960). Considering the location of /æ/ and the fact that it is a novel vowel unlike any Spanish vowel, differences may also be expected in the space mapped on the basis of the first two formants (F_1 and F_2), because F_1 and F_2 are highly relevant in identifying vowels (Jongman–Fourakis–Sereno 1989; Peterson–Barney 1952; Vorperian–Kent 2007). Thus, the productions of English /æ/ may differ in both duration and the first two formants when uttered by bilingual Spanish- and English-speaking children as compared to their monolingual English-speaking peers, especially at an age when children's vowel systems are still undergoing development.

(iii) The present study

Most of the studies that use acoustic analyses to investigate the vowel production of bilingual individuals report on adults only (Flege 1989; Piske et al. 2011), and even the ones that involve children have school-age participants but no preschoolers (Baker–Trofimovich 2005; Jia et al. 2006; Oh et al. 2011; Trofimovich et al. 2003). The present study focuses on bilingual Spanish-

and English-speaking preschool-age children, because in children of this age significant phonological development takes place in simultaneous bilinguals. Investigating the production of a novel and linguistically marked English vowel, /æ/, can provide valuable insights into how phonological development proceeds in bilingual children and how cross-language effects interact with linguistic markedness and general speech development. Considering that the participants of the present study are bilingual as first language or child second language learners with at least three years of exposure to English, we expect our bilingual participants to differentiate the two languages, but we also expect to find differences between our bilingual and monolingual participants due to the expected cross-language interaction and linguistic markedness involving the target vowel, /æ/. To that end, we investigate the durational and spectral profile of /æ/, and we posit two hypotheses that are grounded in existing research and the characteristics of our target population.

Hypotheses

(i) Hypothesis 1 (H1). Duration is a relatively salient feature of English vowels (see Peterson–Lehiste 1960) even though vowel duration per se is not a phonemic feature that differentiates vowels in English (or Spanish). Spanish lacks long vowels, so bilingual Spanish-English children may not use duration as a primary cue to vowel identification. Based on existing research and the characteristics of the participants, we expect to find differences in the durational properties of /æ/ as produced by bilingual children speaking English versus their monolingual English-speaking peers such that the /æ/ vowel tokens produced by bilingual children will be shorter than those produced by their monolingual English-speaking peers.

(i) Hypothesis 2 (H2). The American English /æ/ is not only missing from the Spanish inventory (cf. Levey–Cruz 2004), but it is also a linguistically marked phoneme. We expect to find differences in the first formant and/or the second formant of the /æ/ vowels produced by bilingual children and those produced by their monolingual English-speaking peers.

Method

(i) Participants

The data reported in the present paper consist of a sub-set of a larger project whose purpose was to investigate the acquisition of speech rhythm by bilingual Spanish- and English-speaking children and their monolingual peers in each language (see Bunta–Ingram 2007). The present study included 10 typically developing, monolingual American English-speaking children and 10 typically developing, Mexican Spanish- and American English-speaking bilingual peers, all of whom were living in the Southwestern United States at the time of the study. The monolingual and bilingual children were matched on age, ranging from 4;0 to 5;2 (mean age of the monolingual English children:

4;6, SD: 4 months; mean age of the bilingual participants: 4;7, SD: 5 months). The participants' parents completed detailed questionnaires about their children's language background (e.g., socio-economic status, proportion of use of each language in the household, family history of speech and language impairment, etc.) and gave consent for their child to participate. The children themselves also gave assent to participating in the study. In order for a child to be considered bilingual in this study, s/he had to be able to maintain a conversation in both languages, be able to produce spontaneous full sentences in both languages, and the parent-reported language use had to exceed 20% in each language, a benchmark commonly used in bilingual child language research (cf. Gutierrez-Clellen–Kreiter 2003; Pearson et al. 1997). Both the monolingual English-speaking and the bilingual Spanish- and English-speaking children who participated in the present study were born in the United States, with the exception of one bilingual child who arrived in Phoenix, Arizona at nine months of age. All of the bilingual participants in the present study were bilingual first language learners or child second language learners (see Meisel 2004), because they were exposed to both Spanish and English from birth (except for one child who arrived in the US at the age of 9 months). For the most part, the bilingual participants were exposed to Spanish via parental interactions at home, television and (to a more limited extent) other means. English exposure of the bilingual child participants occurred mostly via peer interactions, television, and preschool language use. There were no parent or teacher concerns regarding the participants' speech or language.

(ii) Materials and procedure

The participants first passed a bilateral hearing screening at 25 dB hearing level (HL), using pure tones of 500, 1000, 2000, and 4000 Hz. To ensure that the participants had typically developing and age-appropriate language, the Expressive Communication subscale of the Preschool Language Scales 4 (Zimmerman–Steiner–Pond 2002a; 2002b) was administered to each child in English, and the bilingual children also completed the Spanish version. All of the children who participated in the present study passed the age-appropriate level, demonstrating typical, age-appropriate language skills.

The larger study also included a picture elicitation task with twenty-six images eliciting quasi-naturalistic connected speech samples to investigate speech rhythm acquisition (Bunta–Ingram 2007). The children were shown a depiction of an action that contained the target words, and they were prompted to describe what was happening. In the present study, we extracted "apple" and "cat" from short sentences in order to analyze the stressed vowel, /æ/. The target words occurred in the same position and carried main stress within the sentence. The words were clearly identifiable in the speech stream, and they were extracted for analysis. One token of each target word was measured for each child.

The procedure for collecting the data was the same for each participant, beginning with a hearing screening, followed by the PLS-4, and finally administering the picture elicitation task. The experimenters were fluent in the language that was tested (for the present study, English), and they were familiar with working with children with clinically and linguistically diverse backgrounds.

(iii) Equipment and measurements

Recordings were made using the Shure Presenter system with a wireless WL93 lapel microphone and a Creative SoundBlaster Audigy 2 NX 24-bit sound card with direct input into a Dell Inspiron 300m laptop computer. The monaural recordings were saved directly onto the hard drive of the laptop as uncompressed, standard sound wave (.wav) files with a sampling rate of 44.1 kHz and a quantization rate of 16 bits. The digital sound files were used for the acoustic analyses described below.

For the present study, the American English low front vowel, /æ/, was analyzed in the words *apple* and *cat*, measuring the duration of the vowel and the first two formants using Wavesurfer (Sjölander–Beskow 2005). The words always occurred in stressed position and in the same carrier phrase. Vowel duration, including formant transitions, was measured based on established criteria using a combination of the time waveform and the spectrogram (e.g., Ladefoged 2003; Peterson–Lehiste 1960; Shah 2002). Vowel duration measurements were considered in agreement if they were within 10 milliseconds of each other, an accepted practice in measuring vowel length (see Peterson–Lehiste 1960). Formant values were obtained using the Formant Plot window in Wavesurfer. The spectrogram that served as the basis of the formant plot had the following properties: 350 Hz bandwidth, 0.8 pre-emphasis filter, 5 kHz range view, Hamming window, and 1024 points for the FFT window length. The duration measurements and the formant readings were verified both via inter-rater and intra-rater reliability. Inter-rater and intra-rater reliability were 98.3% and 96.6%, respectively.

Results

(i) Results for the first hypothesis

According to the first hypothesis, bilingual Spanish- and English-speaking participants were expected to have shorter low front, /æ/, vowels relative to their monolingual English-speaking peers. An independent samples t-test was conducted to compare group performance involving the duration of /æ/. No significant differences were found between monolingual English-speaking children and their bilingual peers speaking English for either the word *apple* [$t(18) = 0.664$ at $p = 0.52$, $d = 0.31$] or the word *cat* [$t(18) = -0.283$ at $p = 0.78$, $d = -0.14$]. Thus, the first hypothesis (H1) was not supported by our data, and the relatively small effect sizes also point to the lack of a substantial difference between the groups on the duration measures.

The mean duration for /æ/ in *cat* was 0.194 sec (SD: 45.8) for the monolingual English-speaking group and 0.201 (SD: 64.4) sec for the bilingual children. In the word *apple*, vowel durations were measured at 0.193 sec (SD: 72.9) when produced by monolingual English-speaking children, and bilingual children speaking English produced an /æ/ vowel (in *apple*) that was 0.171 sec (SD: 76.5), on average. These findings will be revisited in the discussion of the paper.

(ii) Results for the second hypothesis

On the basis of our second hypothesis, we predicted that there would be differences in the location of the vowel as measured by the first two formants when comparing the American English /æ/ produced by monolingual versus bilingual speakers. Independent samples *t*-tests were conducted to compare group performance involving the first two formants of /æ/. No differences were found for the second formant between monolingual and bilingual children when producing /æ/ in *apple* [$t(18) = 1.289$ at $p = 0.21$, $d = 0.61$] or *cat* [$t(18) = -0.874$ at $p = 0.39$, $d = -0.42$]. However, when comparing the monolingual and bilingual groups' realizations of /æ/, differences were found for the first formant of the target vowel in both words: *apple* [$t(18) = 2.428$ at $p = 0.026$, $d = 1.14$] and *cat* [$t(18) = 2.602$ at $p = 0.019$, $d = 1.26$]. Generally, the effect sizes matched the significant results by showing large effects (Cohen's *d* values) for first formant differences and more moderate effect sizes for the second formant differences. Overall, the results were supportive of the second hypothesis, indicating that the bilingual participants differed from their monolingual English-speaking peers in the placement of the vowel /æ/, especially in the vowel height dimension (further explored in the discussion section).

In comparing the averages for the second formant of the vowel /æ/ for the two groups, monolingual English-speaking children had mean values of $F_2 = 1636$ Hz (SD: 513.6 in the word *cat*) and $F_2 = 1741$ Hz (SD: 457.4 in the word *apple*) while bilingual children had means of $F_2 = 1893$ Hz (SD: 758.3 in the word *cat*) and $F_2 = 1508$ Hz (SD: 339.7 in the word *apple*). Considering the first formant values, monolingual English-speaking children had averages of $F_1 = 1057$ Hz (SD: 177.9 in the word *cat*) and $F_1 = 1054$ Hz (SD: 290.1 in the word *apple*), and bilingual children's means were $F_1 = 770$ Hz (SD: 293.7 in the word *cat*) and $F_1 = 742$ Hz (SD: 284.7 in the word *apple*). The results will be further evaluated in the discussion of the paper.

Discussion

The findings of the present study provide interesting insights into how simultaneous Spanish- and English-speaking bilingual children produce the American English /æ/ as compared to their monolingual English-speaking peers. Overall, the results indicate that monolingual and bilingual children do not differ significantly in how long they produce the vowel /æ/ in stressed position. However, bilingual and monolingual children do differ in where

they produce the low front vowel. Notably, the /æ/ vowels in *apple* and *cat* appear to be produced in a higher location in the oral cavity by bilingual Spanish- and English-speaking children than their monolingual English-speaking peers. Nonetheless, the findings deserve further exploration, because some of the results may indicate patterns that are not immediately obvious from our results.

(i) Duration of /æ/

Previous studies found that bilinguals make use of all available acoustic cues when acquiring their languages, even if some of those features are not used to differentiate phonemes in their native language (Iverson–Evans 2007; Morrison 2009). For example, despite the fact that vowel duration is not a salient or phonologically distinctive feature in Spanish, Escudero and Boersma (2004) found that some Spanish-English bilinguals did use duration as a cue to differentiate English /i/ from /ɪ/, relying on a non-native cue. In contrast, in the present study, we did not find significant differences in how monolingual English-speaking children and their bilingual Spanish- and English-speaking peers produce the English /æ/ in terms of its duration. This finding may be due to bilingual children matching their monolingual peers in vowel duration even though Spanish vowels are relatively shorter than the English /æ/. Alternatively, it is possible that a durational difference exists, but the variances in the two groups (bilingual versus monolingual) obscured overall group differences. Notably, Figure 1 illustrates that monolingual English-speaking children display relatively more consistent durational properties than their bilingual peers speaking English when it comes to producing /æ/ in *apple* versus *cat*.

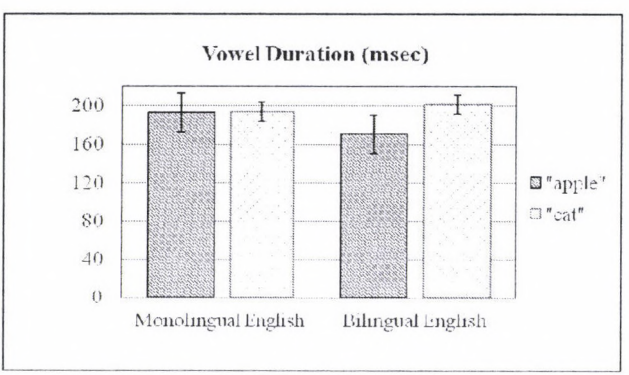


Figure 1.
Duration of American English /æ/ in *apple* and *cat* produced by monolingual English-speaking children and their bilingual Spanish-English peers

The consistency evident in the duration of /æ/ across different words produced by monolingual English-speaking children may indicate relatively stable vowel representations. Conversely, the lack of stable vowel durations in *apple* and *cat* produced by bilingual Spanish- and English-speaking children may reflect the lack of well-established representations, at least in terms of appropriate use of vowel duration (cf. Bosch–Ramon-Casas 2011). Nonetheless, based on our data, we cannot reach a definitive conclusion as to whether or not the durational differences (or lack thereof) are real, and future studies should involve vowel pairs where durational properties can be better investigated (such as comparing the duration of the English /æ/ and /ɛ/ and including more tokens with each vowel).

(ii) First and second formants

Generally, the first formant is inversely related to tongue height (i.e., high vowels tend to have lower F_1 than low vowels). It appears that the bilingual participants in the present study raised their vowels compared to their monolingual peers due to a combination of native language effects (Spanish /e/) and the linguistically marked nature of the low front vowel (Archangeli & Langendoen 1997). However, it is important to note that the bilingual participants in our study did not simply substitute the Spanish /e/ or any other Spanish vowel for the English /æ/; rather, they did differentiate the novel English vowel from their Spanish vowels, but did not match the appropriate English target when compared to their monolingual peers.

Another important observation is that F_1 values displayed considerable consistency within each group; namely, the first formant values of /æ/ – albeit different **across** the monolingual and bilingual groups – were fairly stable across the words **within** each group, unlike the variability attested with respect to duration in the bilingual group (see Figure 2 below).

The second formant is related to tongue advancement in that, generally, as the tongue moves forward, F_2 increases. The bilingual participants in the present study did not differ from their monolingual peers when it came to the second formant, suggesting that the front-back dimension of the vowel /æ/ was not significantly different between the monolingual English-speaking group and their bilingual peers. As with the results regarding the durational properties of the vowel, caution needs to be exercised when interpreting the F_2 results, because there was considerable variation within the bilingual group that needs further exploration in future studies (see Figure 3). Considering the group averages with respect to F_2 values, it must be noted that monolingual English speakers had about a 100 Hz difference in the F_2 of the vowel in *cat* (1636 Hz) as opposed to the first vowel in *apple* (1741 Hz), but the same difference for bilingual participants was much larger (almost 400 Hz: 1893 Hz in *cat* versus 1508 Hz in *apple*). As with duration, it appears that monolingual children have more stable productions of /æ/ compared to their bilingual peers, as measured by F_2.

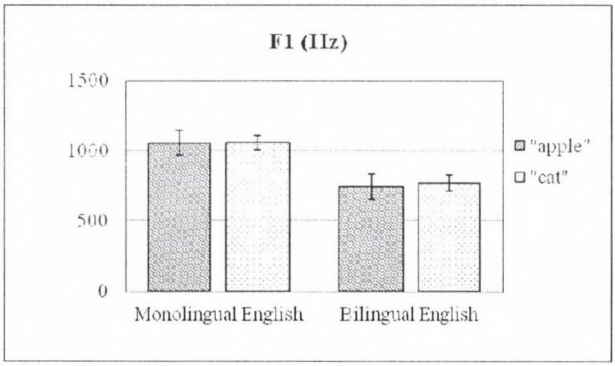


Figure 2.
First formant (in Hertz) of American English /æ/ in *apple* and *cat* produced by monolingual English-speaking children and their bilingual Spanish-English peers

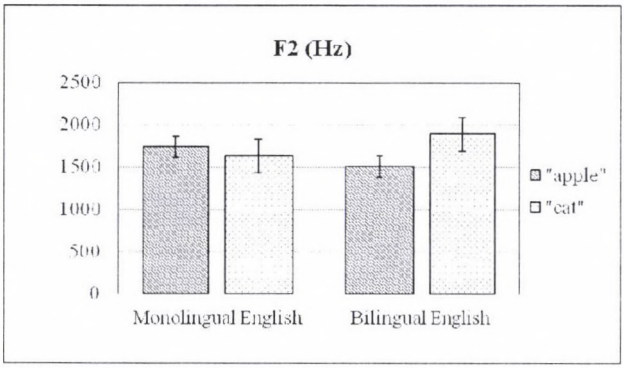


Figure 3.
Second formant (in Hertz) of American English /æ/ in *apple* and *cat* produced by monolingual English-speaking children and their bilingual Spanish-English peers

Taken together, our results pertaining to the first two formants of the English vowel /æ/ produced by bilingual Spanish- and English-speaking children and their monolingual English-speaking peers suggest that bilingual children produce the target vowel with an elevated tongue, relative to monolingual English speakers of the same age. In addition, monolingual children have generally more stable productions of the English low front vowel than their bilingual Spanish- and English-speaking peers, as measured by duration and

the second formant. The F_1 values for our monolingual four-to-five-year-old participants are highly consistent with the seminal work conducted by Peterson and Barney, whose F_1 value for /æ/ for the child participants was 1010 Hz (similar to our 1055 Hz). The F_1 values for /æ/ produced by our bilingual participants (about 750 Hz) were much closer to the F_1 value for /ɛ/ found by Peterson and Barney (690 Hz), which supports the idea that our bilingual participants raised their English low front vowels. The second formant group mean values found in our study ranged from 1508 Hz to 1890 Hz for bilingual children and 1636 Hz to 1741 Hz for our monolingual participants as compared to 2320 Hz noted by Peterson and Barney. It seems that the children who participated in our study formed the low front vowel further back in the oral cavity than Peterson and Barney's participants, and this difference may be due to dialectal or other variation, an issue that may be explored in future studies.

(iii) Other issues

Bilingual Spanish- and English-speaking children encounter two powerful factors in phonological development that make the acquisition of the English low front vowel challenging: (1) the Spanish vowel system is much smaller and less complex than the English one (e.g., 5 Spanish vowels instead of 15, no vowel reduction, no duration issues), and (2) the /æ/ is a linguistically marked vowel (Archangeli–Langendoen 1997). It is not surprising that the bilingual children who participated in our study displayed differences from their monolingual peers in producing /æ/. Most notably, monolingual English-speaking children produced the /æ/ with higher F_1 values than their bilingual peers. In addition, the monolinguals also displayed more consistency in the two target words (*cat* and *apple*) than their peers who were bilingual both in the duration and second formant domains. If variability is to be taken as indication of ongoing acquisition and lack of experience (cf. Trofimovich et al. 2003), our findings suggest that bilingual Spanish- and English-speaking children are lagging behind their monolingual peers with respect to the acquisition of /æ/, consistent with the findings of Fabiano-Smith and Goldstein (2010a) and Kehoe (2002). Kehoe (2002) found evidence for bilingual Spanish- and German-speaking children acquiring German vowel duration slower than their monolingual German-speaking peers. Evidence of delayed acquisition was also found in stop consonant production by Fabiano-Smith and Goldstein (2010a).

A closer analysis of the differences and similarities in producing the English /æ/ by monolingual and bilingual children in our study allows us to comment on categorical assimilation and dissimilation issues. Had equivalence classification blocked the formation of the /æ/ category, our data would have shown that bilingual Spanish- and English-speaking children substituted /e/ or /a/ for the English low front vowel. Instead of substitutions, we found evidence for both separation from Spanish vowels, and also of non-native-

like values for English. Previous studies have also found intermediate values in the productions of bilinguals even in cases when separation of the two languages was evident (Flege–Eefting 1987; Fowler et al. 2008; MacLeod–Stoel-Gammon 2009). More interestingly, our bilingual participants displayed differences from their monolingual peers despite acquiring both languages before their first birthday, and the variability of the bilingual children's productions relative to their monolingual peers may also indicate that acquisition in the former group is still in progress, relative to their monolingual peers' regarding the placement and the duration of the vowel.

In the present study, we did not focus on three factors that have dramatic effects on bilingual language development: age of acquisition, length of exposure, and intensity of exposure to the target languages, because all of our participants started learning their languages during the first year of life, used both languages and had over three years of exposure to both English and Spanish (the latter if they were bilingual). Nonetheless, there is a considerable amount of literature devoted to the effects of age of acquisition, length of residence, and intensity of first and second language use on success with the target languages of bilinguals (Flege–Schirru–MacKay 2003; Levey–Cruz 2004; Piske et al. 2011; Trofimovich et al. 2003, and many others).

(iv) Limitations and directions

While this study provided interesting insights into how bilingual Spanish-English children produce a vowel that is difficult to acquire, the work needs to continue to disentangle some of the effects that cannot be observed based on our limited data. A more comprehensive approach is needed, expanding acoustic and segmental analyses to more sounds while at the same time considering other measures of phonological skills, such as phonological whole-word measures (e.g., Bunta et al. 2009). In addition to further phonological and acoustic analyses, more participants are needed to draw definitive conclusions, because the present study involved only a limited number of children. Increasing the number of participants and expanding their age range hold the promise of revealing more insights into bilingual phonological development.

Future studies should also investigate the effects of using different elicitation techniques on vowel productions. For example, studies should consider more carefully the differences or similarities between samples elicited via picture prompts, reading, and unscripted conversations with the experimenter (i.e., spontaneous samples). It appears that Flege (1991) found that most non-native English speakers and monolingual Spanish speakers identified /æ/ as *a* influenced by orthography, but in the present study, the substitute was a more raised vowel rather than a low back one. These differences warrant systematic future research.

The present study is also limited to investigating Spanish- and English-speaking bilingual children, so adding more and hitherto understudied lan-

guage combinations could reveal new information about phonological development. For example, investigating the vowel systems of language pairs such as Hungarian and English would allow for comparing two highly complex but also very different vowel and phonological systems. Iverson and Evans (2007) claimed that speakers whose languages had complex vowel systems (such as Norwegian or German) were better able to identify and categorize English vowels. However, Norwegian and German are both related to English, so it is unknown how Hungarian learners of English would fare with two complex and yet very different systems. In a pilot study involving two trilingual Farsi-, English- and Hungarian-speaking children, Navracsics (1999) found that phonemes unique to Hungarian occurred early in acquisition and were produced accurately (including /œ/ and /y/, p. 109). Moreover, most of the substitutions attested in the productions of these trilingual children were consistent with the patterns that occurred in monolingual Hungarian-speaking children (p. 110). However, in the early stages of vowel development, the trilingual children in Navracsics' study did substitute [i:] for /e:/ and [ɒ] for /o/ in Hungarian and [i] for /ɛ/ in English. According to Navracsics (1999), the [i] substitutions were due to Farsi influences. In any event, more research on the development of Hungarian- and English-speaking bilinguals may yield results that could be very different from patterns attested in other language combinations, and therefore, such a line of research has the potential of revealing valuable novel insights that have potentially significant theoretical implications for bilingual phonological development.

The main findings of our study were that bilingual Spanish- and English-speaking children produced the English /æ/ with a higher place feature than their monolingual peers, but the duration and F_2 values were commensurate for the two groups. Bilingual participants also displayed more overall variability than their monolingual peers when producing the English low front vowel, which could be a sign of ongoing and possibly incomplete acquisition. The available avenues for continuing to explore bilingual phonological development are numerous, and they range from acoustic analyses (such as the ones presented in this paper) to linguistic measures (such as phonological features or whole-word analyses). Future studies should adopt a comprehensive approach to investigating phonological acquisition, incorporating elements of different analyses to gain a more complete picture of how the sound systems of bilingual children evolve. The nature of phonological development in bilingual children also necessitates conducting more longitudinal studies, because the best way to observe phonological development in children is via long-term research that follows the progress of the participants so that various factors affecting phonological development may be teased apart and analyzed more carefully.

References

Archangeli, Diana – Langendoen, D. Terence (eds.) 1997. *Optimality Theory: An overview*. Blackwell Publishers, Cornwall, UK.

Baker, Wendy – Trofimovich, Pavel 2005. Interaction of native- and second-language vowel system(s) in early and late bilinguals. *Language and Speech* 48. 1–27.

Bosch, Laura – Ramon-Casas, Marta 2011. Variability in vowel production by bilingual speakers: Can input properties hinder the early stabilization of contrastive categories? *Journal of Phonetics* 39/4. 514–526.

Bunta, Ferenc – Fabiano-Smith, Leah – Ingram, David – Goldstein, Brian A. 2009. Phonological whole-word measures in three-year-old bilingual children and their monolingual peers. *Clinical Linguistics and Phonetics* 23/2. 156–175.

Bunta, Ferenc – Ingram, David 2007. The acquisition of speech rhythm by bilingual Spanish- and English-speaking four- and five-year-old children. *Journal of Speech, Language, and Hearing Research* 50/4. 999–1014.

Bunta, Ferenc – Major, Roy C. 2004. An Optimality Theoretic account of Hungarian ESL learners' acquisition of /ɛ/ and /æ/. *International Review of Applied Linguistics in Language Teaching* 42. 277–298.

Davis, Barbara L. – MacNeilage, Peter F. 1990. Acquisition of correct vowel production: A quantitative case study. *Journal of Speech and Hearing Research* 33. 16–27.

Dauer, Rebecca M. 1983. Stress-timing and syllable-timing re-analysed. *Journal of Phonetics* 11. 51–62.

Delattre, Pierre 1966. A comparison of syllable length conditioning among languages. *International Review of Applied Linguistics in Language Teaching* 4. 183–198.

Donegan, Patricia 2002. Normal vowel development. In Ball, Martin J. – Gibbon, Fiona (eds.): *Vowel disorders*. Butterworth-Heinemann, Boston. 1–35.

Escudero, Paola – Boersma, Paul 2004. Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition* 26. 551–585.

Fabiano-Smith, Leah – Barlow, Jessica A. 2010. Interaction in bilingual phonological acquisition: Evidence from phonetic inventories. *International Journal of Bilingual Education and Bilingualism* 13/1. 81–97.

Fabiano-Smith, Leah – Goldstein, Brian A. 2010a. Early-, middle-, and late-developing sounds in monolingual and bilingual children: An exploratory investigation. *American Journal of Speech-Language Pathology* 19. 66–77.

Fabiano-Smith, Leah – Goldstein, Brian A. 2010b. Phonological acquisition in bilingual Spanish-English speaking children. *Journal of Speech, Language and Hearing Research* 53. 160–178.

Flege, James E. 1989. Differences in inventory size affect the location but not the precision of tongue positioning in vowel production. *Language and Speech* 32. 123–147.

Flege, James E. 1991. The interlingual identification of Spanish and English vowels: Orthographic evidence. *The Quarterly Journal of Experimental Psychology* 43A/3. 701–731.

Flege, James E. 1995. Second language speech learning: Theory, findings, and problems. In Strange, Winifred (ed.): *Speech perception and linguistic experience: Issues in cross-language research*. York Press, Timonium, MD. 233–277.

Flege, James E. – Eefting, Wieke 1987. The production and perception English stops by Spanish speakers in English. *Journal of Phonetics* 15. 67–83.

Flege, James E. – Schirru, Carlo – MacKay, Ian R. A. 2003. Interaction between the native and second language phonetic subsystems. *Speech Communication* 40. 467–491.

Fowler, Carol A. – Sramko, Valery – Ostry, David J. – Rowland, Sarah A. – Hallé, Pierre 2008. Cross language phonetic influences on the speech of French-English bilinguals. *Journal of Phonetics* 36. 649–663.

Gildersleeve-Neumann, Christina E. – Kester, Ellen S. – Davis, Barbara L. – Peña, Elizabeth D. 2008. English speech sound development in preschool-age children from bilingual English-Spanish environments. *Language, Speech and Hearing Services in Schools* 39. 314–328.

Goldstein, Brian A. – Bunta, Ferenc in press. Positive and negative transfer in the phonological systems of bilingual speakers. *The International Journal of Bilingualism.*

Goldstein, Brian A. – Cintrón, Patricia 2001. An investigation of phonological skills in Puerto Rican Spanish-speaking 2-year-olds. *Clinical Linguistics and Phonetics* 15. 343–361.

Grosjean, François 1989. Neurolinguists, Beware! The bilingual is not two monolinguals in one person. *Brain and Language* 36. 3–15.

Guion, Susan 2003. The vowel systems of Quichua-Spanish bilinguals - Age of acquisition effects on the mutual influence of the first and second languages. *Phonetica* 60/2. 98–128.

Gutierrez-Clellen, Vera – Kreiter, Jacqueline 2003. Understanding child bilingual acquisition using parent and teacher reports. *Applied Psycholinguistics* 24. 267–288.

Højen, Anders – Flege, James E. 2006. Early learners' discrimination of second-language vowels. *Journal of Acoustical Society of America* 119. 3072–3084.

Iverson, Paul – Evans, Bronwen G. 2007. Learning English vowels with different first-language vowel systems: Perception of formant targets, formant movement, and duration. *Journal of Acoustical Society of America* 122. 2842–2854.

Jia, Gisela – Strange, Winifred – Wu, Yanhong – Collado, Julissa – Guan, Qi 2006. Perception and production of English vowels by Mandarin speakers: Age-related differences vary with amount of L2 exposure. *Journal of Acoustical Society of America* 119. 1118–1130.

Jongman, Allard – Fourakis, Marios – Sereno, Joan A. 1989. The acoustic vowel space of modern Greek and German. *Language and Speech* 32. 221–248.

Kehoe, Margaret 2002. Developing vowel systems as a window to bilingual phonology. *The International Journal of Bilingualism* 6/3. 315–334.

Kent, Raymond D. 1992. The biology of phonological development. In Ferguson, Charles A. – Menn, Lise – Stoel-Gammon, Carol (eds.): *Phonological development: Models, research, implications.* York Press, Timonium, MD, 65–90.

Ladefoged, Peter 2003. *Phonetic data analysis: An introduction to fieldwork and instrumental techniques.* Blackwell, Oxford, UK.

Levey, Sandra – Cruz, Denise 2004. The discrimination of English vowels by bilingual Spanish/English and monolingual English speakers. *Contemporary Issues in Communication Science and Disorders* 31. 162–172.

MacLeod, Andrea A. N. – Stoel-Gammon, Carol 2009. The use of voice onset time by early bilinguals to distinguish homorganic stops in Canadian English and Canadian French. *Applied Psycholinguistics* 30/1. 53–77.

Meisel, Jürgen 2004. The bilingual child. In Bhatia, Tej K. – Ritchie, William C. (eds.): *The handbook of bilingualism.* Blackwell, Malden, MA, 91–113.

Mora, Joan C. – Keidel, James L. – Flege, James E. 2011. Why are the Catalan contrasts between /e/-/ɛ/ and /o/-/ɔ/ so difficult for even early Spanish-Catalan bilinguals to perceive? In Wrembel, Magdalena – Kul, Małgorzata – Dziubalska-Kołaczyk, Katarzyna (eds.): *Achievements and perspectives in the acquisition of second language speech.* New Sounds 2010, Vol. 2. Peter Lang, Bern, Switzerland, 183–193.

Morrison, Geoffrey S. 2009. L1-Spanish speakers' acquisition of the English /i/ - /ɪ/ contrasts II: Perception of vowel inherent spectral change. *Language and Speech* 52. 437–462.

Navracsics, Judit 1999. *A kétnyelvű gyermek* [The bilingual child]. Corvina, Budapest.

Oh, Grace E. – Guion-Anderson, Susan – Aoyama, Katsure – Flege, James E. – Akahane-Yamada, Reiko – Yamada, Tsuneo 2011. A one-year longitudinal study of English and Japanese vowel production by Japanese adults and children in an English-speaking country. *Journal of Phonetics* 39/2. 156–167.

Pearson, Barbara – Fernandez, Sylvia – Lewedeg, Vanessa – Oller, D. Kimborough 1997. The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics* 18. 41–58.

Peterson, Gordon E. – Lehiste, Ilse 1960. Duration of syllable nuclei in English. *Journal of the Acoustical Society of America* 32. 693–703.

Piske, Thorsten – Flege, James E. – MacKay, Ian R. A. – Meador, Diane 2011. Investigating native and non-native vowels produced in conversational speech. In Wrembel, Magdalena – Kul, Małgorzata – Dziubalska-Kołaczyk, Katarzyna (eds.): *Achievements and perspectives in the acquisition of second language speech.* New Sounds 2010, Vol. 2. Peter Lang, Bern, Switzerland, 195–205.

Shah, Amee 2002. *Temporal characteristics of Spanish-accented English: Acoustic measures and their correlation with accentedness ratings.* (Doctoral dissertation, The City University of New York, 2002). Dissertation Abstracts International, 63-09B, 4311.

Sjölander, Kare – Beskow, Jonas 2005. *WaveSurfer* (Version 1.8.5) [Computer software]. Retrieved December, 2005, from http://www.speech.kth.se/wavesurfer/

Trofimovich, Pavel – Baker, Wendy – Flege, James E. – Mack, Molly 2003. Second language sound learning in children and adults: Learning sounds, words of both? In Beachley, Barbara – Brown, Amanda – Conlin, Frances (eds.): *Proceedings of the 27th Boston University Conference on Language Development.* Cascadilla Press, Sommerville, MA, 775–786.

Vorperian, Houri K. – Kent, Raymond D. 2007. Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research* 50. 1510–1545.

Zimmerman, Irla L. – Steiner, Violette G. – Pond, Roberta E. 2002a. *Preschool language scale fourth edition (English Edition).* Pearson, San Antonio, TX.

Zimmerman, Irla L. – Steiner, Violette G. – Pond, Roberta E. 2002b. *Preschool language scale fourth edition (Spanish Edition).* Pearson, San Antonio, TX.

Acknowledgments: Funding for this study was provided in part by the New Faculty Research Program granted by the University of Houston. The authors thank the children and their parents for choosing to participate in the study.

VIRTUÁLIS MONDATOK GYERMEKEK SPONTÁN BESZÉDÉBEN

Neuberger Tilda

Bevezetés

A hosszabb szövegek megértésének egyik alapfeltétele a megfelelő tagolás. Ez mind az írott, mind a beszélt nyelvben így van. Az írott nyelvben az egységekre bontás a befogadónak (olvasónak) rendszerint nem okoz problémát a tipográfiának, illetve a központozásnak köszönhetően. A szöveg formai megjelenése alapján ugyanis a mondatok és a bekezdések világosan elkülönülnek egymástól. *A magyar helyesírás szabályai* (1984) szerint az írásjelek tükrözik a mondatok szerkezetét, tagolódását, részeik egymáshoz kapcsolódását, s emellett nagy kezdőbetűk állnak minden szöveg, valamint – pont, kérdőjel, felkiáltójel és bizonyos esetekben kettőspont után – minden új mondat élén. A beszélt nyelvben nem támaszkodhatunk efféle „mankókra" az elhangzott közlések szegmentálásakor. Az írás és a beszéd a nyelv két különböző, de egyenrangú megjelenési formája, normarendszerük között a legnagyobb különbség talán a mondattanban található (Lanstyák 2009). Az írott nyelv tudatosabban szerkesztett, és kevesebb benne a tévesztés, hiszen lehetőség van a szövegjavításra (Érsok 2006). Emellett statikus, a befogadónak nem csak egyszer van alkalma elolvasni a szöveget. Ezzel szemben a beszélt nyelv, pontosabban a spontán, kötetlen beszéd előkészítés nélküli, időhöz kötött, egyszeri. Normája lazább, mint az írott nyelvé. Jellemzőek rá a grammatikai közbeékelések, a váratlan bővítmények, az egyeztetési tévedések, a tartalmi és formai ismétlések (Keszler 1983). A közlések gyakran félbeszakadnak, hiányosak, megjelenhetnek mondatátszövődések, az értelmezéshez a beszédhelyzet, a kontextus nagy szerepet kap. Mivel a gondolatok kiválogatása és nyelvi formába öntése, a tervezés és a kivitelezés szinte egyidejűleg mennek végbe, ez magában hordozza a diszharmónia lehetőségét, amelynek a felszínen megjelenő következményei a különféle megakadásjelenségek is (Gósy 2004). Mindebből következik, hogy a beszélt nyelvben a tagolás nehézkesebb és kevésbé egyértelmű, mint írásban.

Spontán beszéd esetén a beszélő már a beszédprodukció mikrotervezési szakaszában (Levelt 1989) igyekszik megtervezni, hogy hogyan tagolja a mondanivalóját, az artikulációs megvalósítás azonban túlnyomórészt ösztönösen zajlik. A szegmentálást befolyásolják egyfelől bizonyos fiziológiai tényezők (például a légzés), másfelől maga a közölni kívánt tartalom, a szöveg típusa, a beszédhelyzet (Váradi 2008), valamint nagyok az egyéni különbsé-

gek (Kohler 1983). A hallgató a beszédmegértési folyamat során szegmentálja az elhangzottakat, vagyis bizonyos egységekre bontja a közlést, amelyek lehetnek mondatnyi vagy annál hosszabb szövegrészletek, például bekezdések („paragraph", vö. Lehiste 1979).

A *Magyar értelmező kéziszótár* (Juhász et al szerk. 1987: 965) a mondatot úgy határozza meg, hogy az „a beszédnek az a legkisebb, rendszerint több szavas egysége, amellyel valamit kijelentünk, kívánunk vagy kérdezünk". A hagyományos grammatikák szerint a mondat „egy vagy több szóból áll, zárt szerkezet jellemzi. A mondat a legnagyobb nyelvi egység, amely a nyelv szabályai, mintái szerint nyelvi elemekből megszerkeszthető. Egyben a legkisebb nyelvi egység, amely egy adott beszédhelyzetben közlésegységgé (a nyelvi kommunikáció egységévé) válhat" (Kugler 2000: 371). A beszélt nyelvben a szövegeket úgynevezett virtuális mondatokra oszthatjuk, amelyek „többé-kevésbé azonos paraméterekkel jellemezhetők, s ezeket a paramétereket a hallgatók képesek tagoló funkcióban használni a spontán közlések hallgatásakor" (Gósy 2003a: 22). A spontán beszéd szegmentálását a tartalmi sajátosságok, a szintaktikai szerkezet és bizonyos szupraszegmentális tényezők mentén végezhetjük el, utóbbiak közé sorolható a szünet, az alaphangmagasság változása, az intenzitás csökkenése vagy a szünet előtti utolsó szó időtartalmának növekedése (Gósy 2004). Ezek a tényezők együttesen segítik a tagolást, de eltérő fontosságúak a sikeres szegmentálásban, hierarchikusan szerveződnek a percepcióban: a szemantikai koherencia alapvető fontosságú, a prozódiai eszközök közül a szünet meghatározóbb az észlelésben, mint az alapfrekvencia változása (Gósy–Kovács 2007).

A magyar kutatásokból már sok információnk van a felnőttek spontán beszédének tagolásáról (pl. Wacha 1988, Gósy 2003a, Váradi 2008, 2009), de virtuális mondatokat valószínűleg a gyermekek beszédében is találhatunk, mivel a spontán beszéd tervezése hasonlóan történik a nyelvelsajátítás egyes fokain, mint a felnőtteknél, igaz, eltérő gyakorlottsággal. Ismeretes, hogy az anyanyelv-elsajátítás során az első jelentéssel bíró hangsorok egyéves korban jelennek meg (Vértes O. 1953; S. Meggyes 1971; Gósy 1981; Piaget 1970). A szómondatok vagy holofrázisok elnevezés arra utal, hogy formailag a szavakhoz, funkciójukat tekintve a felnőttnyelvi mondatokhoz hasonlatosak. Megértésüket nagyban befolyásolja a szituáció, s a jelentés az intonáció függvényében változik (Lengyel 1981). Ezeket követik a kétszavas közlések, illetve a távirati stílusú beszéd. A korai „mondatok" telegrafikus jellegét a funkciószavak (névelő, létige stb.) és a szóvégződések (pl. ragok) elhagyása kelti (Crystal 2003). A kétéves kor tájékán produkált megnyilatkozásoknak a hosszúsága elérheti a három-négy szót, amelyeket a gyermekek többféleképpen kombinálnak, valamint egyre több toldalékkal látnak el. Hároméves kor után megjelennek az egynél több tagmondatból álló mondatok, gyakorivá válnak a párbeszédek. Hatéves korra pedig a gyermekek már képesek bonyolultabb tar-

talmak kifejezésére, összefüggő szóbeli szövegek létrehozására, beszédüket különféle megakadásjelenségek tarkítják (Gósy 2005; Horváth 2009).

A hatéves gyermekek eltérő szintaktikai összetettségű közléseket produkálhatnak (S. Meggyes 1981). A grammatikai komplexitás mértékét a magyar gyermekek spontán beszédében az úgynevezett KFM-módszerrel vizsgálhatjuk (Gerebenné et al. 1992). Ez az amerikai Lee és Canter (1971) *Developmental Sentence Scoring* elnevezésű kritériumrendszerének magyarra való adaptálása. A közlésegységek fejlődési mutatója azt méri, hogy a gyermek spontán beszédét milyen szófajú szavak, mennyire bonyolult szerkezetek építik fel. Az adaptálás során a magyar kutatók a hatéves korra megadtak egy tapasztalati értéket, amely elvárható az ilyen idős gyermekek közléseiben, ez az érték 11 (vö. Horváth 2006). A Horváth Viktória (2006) kutatásában részt vett hatéves gyermekek átlagos KFM-értéke 14,84 volt, a két szélső értéket (4,51–27,88) egy-egy fiú beszédében találta a szerző. A jelen tanulmány szerzője egy korábbi vizsgálatban hasonló eredményekre jutott (Neuberger 2011): az óvodások közlésegységeinek fejlődési mutatója átlagosan 13,45, a szélső értékek pedig 7,93–20,21 (egy-egy fiúnál). A KFM-módszer alkalmazásához elengedhetetlen, hogy a gyermekek szövegeit szegmentáljuk, hiszen a számításokhoz mind a helyes virtuális mondatok száma, mind az összes virtuális mondat számára szükség van. A virtuális mondatokra bontás minden esetben szubjektív, és ez befolyásolhatja a különböző vizsgálatok KFM-re vonatkozó eredményeit. De vajon mi alapján húzhatjuk meg a virtuális mondatok határait a gyermekek szövegeiben?

A jelen kutatás célja megismerni, hogy milyen eszközöket alkalmaznak a gyermekek tagoló funkcióban a spontán beszédben, illetve hogy a hallgatók mi alapján szegmentálják a gyermekek közléseit. Hipotézisünk szerint a beszéd grammatikai komplexitása befolyásolja a hallgatók határjelzéseit: egy kevésbé összetett szövegben könnyebb és egyöntetűbb a virtuális mondatok meghatározása, mint egy szintaktikailag komplexebb beszédrészletben. A fiatalabb gyermekek általában nem összefüggő monológokat hoznak létre, hanem gyakran támaszkodnak a kísérletvezető kérdéseire, s ez eleve tagolja valamennyire a szövegeiket. Másrészt az egyszerűbb szerkezetek létrehozásához kevesebb szó szükségeltetik, valamint rövidebbek is, ami megkönnyíti a hallgatók feldolgozását, a megértést. Mindezeknek köszönhetően pedig valószínűsíthető, hogy a gyermekek kevésbé összetett közléseit egyszerűbb szegmentálni, mint a felnőttek bonyolultabb szövegeit.

Anyag, módszer, kísérleti személyek

A jelen kísérlethez két hatéves óvodás fiú beszédanyagát használtuk fel. Tíz óvodástól rögzítettünk spontán beszédet – digitális hangfelvevő segítségével, a megszokott óvodai környezetben, minden gyermekkel egyénileg. A gyermekeket arról kérdeztük, hogy mivel szeretnek játszani, mi a hobbijuk, illetőleg milyen óvoda utáni elfoglaltságaik vannak. Minden óvodás szövegét leje-

gyeztük, majd kiszámoltuk bennük a közlésegységek fejlődési mutatóját (KFM), s végül kettőt választottunk ki. A szövegek kiválasztásának fő szempontja az volt, hogy a grammatikai összetettségük eltérő legyen, azaz relatíve nagy különbség legyen a kettő közlésegységeinek fejlődési mutatójában. A morfológiailag és szintaktikailag egyszerűbb szöveg (Gy1) KFM-értéke 12,5, az összetettebbé (Gy2) pedig 21,6. Az első gyermek KFM-értéke a hat éves korban elvárt 11-es értéket nem sokkal haladja meg, míg a második gyermek közlésegységeinek fejlődési mutatója az általános iskola 3. osztályába járó gyermekek eredményeihez közelít (az ő átlageredményük egy vizsgálatban: 20,36; vö. Neuberger 2011). Mindkét gyermek ép hallású, ép intellektusú, nincs beszédhibájuk.

A két gyermek hanganyagát alapul véve elkészítettük a tesztlapokat, amelyeken a gyermekek szövegei szerepeltek a kísérletvezető kérdéseivel együtt, helyesírásban lejegyezve, de központozás nélkül. A percepciós tesztben 45 magyar szakos egyetemi hallgató vett részt, életkoruk 21–42 közötti, az átlagéletkor 24 év. Közülük 37 nő és 8 férfi. A kísérleti személyek azt a feladatot kapták, hogy a hanganyag kétszeri meghallgatása során jelöljék be a mondathatárokat a leírt szövegben függőleges vonalakkal. A percepciós teszt végén egy értékítéletet is kértünk a hallgatóktól arra vonatkozóan, hogy véleményük szerint melyik szöveget volt nehezebb szegmentálni (Gy1 vagy Gy2). A mondatvégjelöléseket összesítettük és elemeztük. A hanganyaggal összevetve megvizsgáltuk azokat a helyeket, ahol az adatközlők mondatvéget észleltek. Az elemzés az észlelés egyöntetűségére (hány adatközlő jelölt az adott helyeken), valamint a szünetezések és a szakaszvégi dallammenetek vizsgálatára terjedt ki. Megmértük funkciótól függetlenül az összes szünet időtartamát. A szünethossz alsó határának az általánosan elfogadott 100 ms-os értéket tekintettük (Gósy 2004), ugyanis ez a percepció számára jól azonosítható. A hangmeneteket előzetesen három fonetikus (saját percepcióján alapuló) minősítése alapján határoztuk meg, majd a Praat 5.3.02-es verziószámú programban (Boersma–Weenink 2011) lemértük az alaphangmagasság változására vonatkozó értékeket, az auditíven észlelt és a dallamgörbén látott frekvenciaváltozásra támaszkodva. Az előzetes percepciós ítéletek javarészt összhangban voltak egymással, ahol nem egyöntetű döntés született, ott a műszeres elemzésre támaszkodva döntöttünk a dallammenet típusáról. A szakaszvégi intonációs egység minimum és maximum frekvenciaértéke közötti 10%-os eltérésig lebegőnek minősítettük az adott szakasz hangmenetét, annál nagyobb változás esetén ereszkedőnek (ha negatív irányú), illetve emelkedőnek (ha pozitív irányú) tekintettük a dallammenetet (vö. Kassai 1998: 213; Markó 2005: 61). Szökő vagy eső zárlatot nem különítettünk el, hiszen ezek megjelenési aránya általában kismértékű a többi dallamtípushoz képest (Markó 2009), és a jelen kutatás anyagában az előzetes percepciós ítéletek szerint sem állapítottunk meg ilyen típusú hanglejtésformákat.

Megállapítottuk a jelölések és az akusztikai sajátosságok összefüggéseit. Az annotálást és a méréseket a Praat programban, a statisztikai elemzéseket (egyváltozós varianciaanalízis és Pearson-féle korrelációelemzés) az SPSS 13.0-s verziószámú szoftver segítségével végeztük.

Eredmények

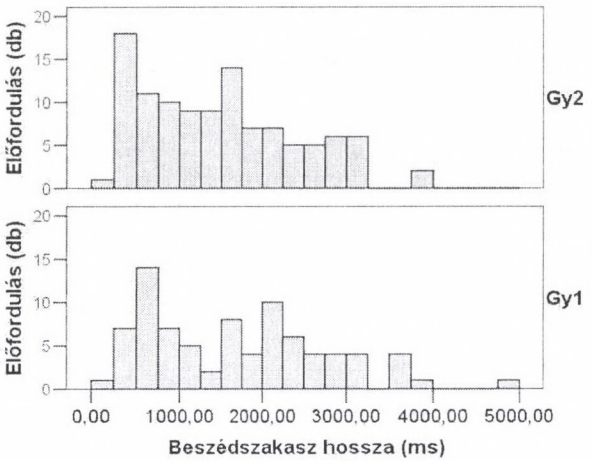
A szövegek temporális jellemzői

A temporális jellemzők körébe soroltuk a szövegeknek azokat a szupraszegmentális sajátosságait, amelyek az időtartam segítségével megragadhatók. Ilyen a szövegek teljes időtartama, a beszédszakaszok és a szünetek száma és időtartama, a szünetek időtartamának aránya a beszédidőhöz képest, valamint az artikulációs és a beszédtempó.

Az alacsonyabb KFM-értékű, vagyis szintaktikailag egyszerűbb szöveg (Gy1) teljes időtartama 275 másodperc, a magasabb KFM-értékűé, vagyis az összetettebbé (Gy2) pedig 253 másodperc. Ez az időtartam magában foglalja a kísérletvezető kérdéseinek idejét is, valamint a hallgatások (szóátvételt biztosító jelkimaradások) hosszát is, azok nélkül – vagyis tisztán a gyermek beszédét számolva – az első szöveg (Gy1) 226, a második (Gy2) 238 másodperc hosszú. Az időtartamértékek is mutatják, hogy az első gyermek kevesebbet beszélt, a kísérletvezetőnek több kérdést kellett feltennie neki, hogy beszélésre ösztönözze. Ezért a szöveg több pontján dialogikus részek jöttek létre. Ezzel szemben a másik gyermek összefüggőbb szöveget hozott létre. Ebből a különbségből is adódhat a KFM-értékek közötti eltérés.

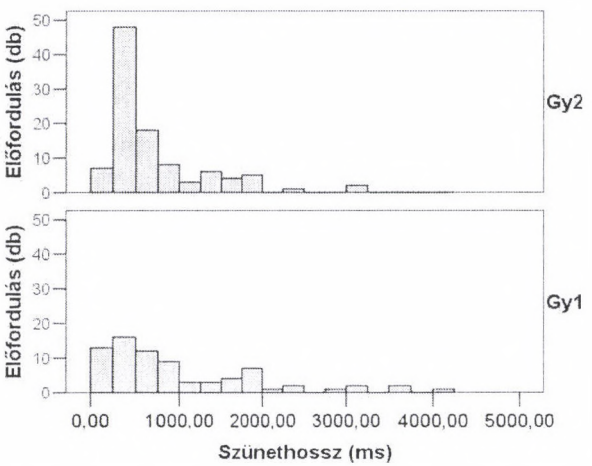
A hanganyagokat beszédszakaszokra bontottuk, amelyek két (néma vagy kitöltött) szünet (vagy egy szünet és egy hallgatás) közötti szöveges egységek (Markó 2005). Meghatároztuk ezek számát és időtartamát mind a két szövegben (1. ábra).

Az első gyermek szövege összesen 82 darab, a másodiké 110 darab beszédszakaszra osztható. Az első esetben (Gy1) a beszédszakaszok átlagos hosszúsága 1715 ms (átlagos eltérés: 868 ms), a második esetben (Gy2) pedig 1502 ms (átlagos eltérés: 752 ms), vagyis az egyszerűbb szerkezeteket használó gyermek hosszabb időtartamú szakaszokban beszélt (ez összefügghet az artikulációs tempójával is). A legrövidebb beszédszakaszt az első kisfiú produkálta, ez 63 ms hosszú volt: egy, két néma szünettel határolt *a* névelő. A leghosszabb beszédszakasz 4891 ms-os volt, vagyis majdnem 5 másodperces. Ezt is az első gyermek szövegében találtuk, és így hangzott: *a lasszóesés, hogy fönn megpörgünk és úgy esünk le, hogy a lábunk föl v-*. A második gyermek legrövidebb beszédszakasza 196 ms-os volt, a leghosszabb pedig 3873 ms-os. A statisztikai elemzés nem mutatott ki szignifikáns különbséget a két szöveg között a beszédszakaszok hosszát illetően [egytényezős ANOVA: $F(1, 191) = 2{,}270$; $p = 0{,}134$)], de az a tendencia látszik, hogy az összetettebb közléseket produkáló gyermek rövidebb szakaszokban beszélt, beszédét több szünet tagolta.



1. ábra
A különböző hosszúságú beszédszakaszok előfordulása a két szövegben

Elemeztük a néma, a kitöltött és a kombinált szünetek számát és időtartamát, a hallgatásokat kizártuk az elemzésből (2. ábra).



2. ábra
A különböző hosszúságú szünetek előfordulása a két szövegben

Míg a Gy1 szövegben összesen 77 darab szünetet találhatunk, addig a Gy2 szövegben 102 darabot, vagyis az első gyermek átlagosan 2,9 másodpercenként, a második gyermek pedig 2,3 másodpercenként tartott szünetet. Az egyszerűbb szerkezeteket használó kisfiú beszédében kevesebb, de hosszabb szünet fordult elő, mint a komplexebb szövegben, ahol viszont több rövid szünet osztotta részekre a mondanivalót. Míg az első esetben a szünetek átlagos hossza 1114 ms volt (átlagos eltérés: 846), addig a második esetben 718 ms (átlagos eltérés: 439 ms). Az egytényezős ANOVA szignifikáns különbséget mutatott a két szöveg szünetidőtartamai között: $F(1, 178) = 8{,}398$; $p = 0{,}004$. A legrövidebb szünet 102 ms-os volt, a leghosszabb pedig 8261 ms-os, vagyis több mint 8 másodperces. Mindkettőt a Gy1 szövegben adatoltuk. A Gy2 szövegben 107 ms és 3235 ms a két szélsőséges szünetidőtartam értéke. A szünetek tartam szerinti eloszlását a 2. ábrán láthatjuk (a jobb átláthatóság kedvéért a szélsőséges 8 másodperces szünet kivételével, amely a Gy1 szövegben jelent meg).

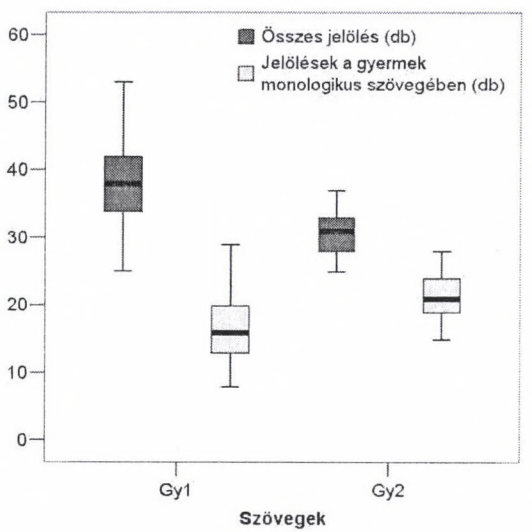
Az összes szünet időtartamának aránya a teljes beszédidőhöz képest nagymértékben beszélőfüggő, 10%-tól akár 60%-ig is szórhat. Magyar felnőttek spontán beszédére vonatkozó adatok szerint a szövegek teljes időtartamának átlagosan 20–30%-át teszik ki a szünetek (vö. Gósy 2003b; Markó 2005). A jelen kísérletben részt vevő gyermekek esetében a szünetek időtartama a teljes beszédidőhöz viszonyítva Gy1 beszédében 37,6%-os, Gy2 beszédében pedig 30,6%-os. A grammatikailag egyszerűbb felépítésű szövegben tehát magasabb a szünetek aránya, mint a komplexebb szövegben. Ez abból fakadhat, hogy az első gyermek bizonytalanabb és kevésbé motivált volt a beszélgetés során, gyakran nem tudta, mit mondjon.

A Gy1 szövegben a szünetek 88%-a néma szünetként realizálódott, 7%-a volt kombinált típusú, és mindössze 5%-a volt önálló kitöltött szünet. A Gy2 szövegben néma szüneteket 91%-ban, kombináltat 8%-ban és hezitálásokat csupán 1%-ban adatoltunk. Láthatjuk tehát, hogy már a hatévesek beszédében is jelen vannak a hezitálások, sőt olykor kombinálják is a hezitálásokat a néma szünettel bizonytalanság miatt vagy időnyerés céljából.

Megmértük mind a két hanganyag tempóértékeit. Az alacsonyabb KFM-értékű szöveg 5,1 hang/s-os beszédtempóval hangzott el, a komplexebb pedig 7,5 hang/s-ossal. A beszédtempó megállapításához a szünetek időtartamát is bele kell számolnunk, és mint láthattuk, az az első gyermek esetében nagyobb arányú. A második gyermek artikulációs tempója is gyorsabb volt, 9,9 hang/s-os, míg az elsőé 7,8 hang/s. Az egyváltozós varianciaanalízis szerint szignifikáns különbség mutatható ki a két kisfiú artikulációs tempójában: $F(1, 191) = 15{,}489$; $p < 0{,}001$. A grammatikai komplexitás nagyobb foka tehát együtt járt a gyorsabb beszéd- és artikulációs tempóval, ami arra utalhat, hogy a második gyermek anyanyelvi szintje az adott életkorban magasabb szintű, mint az elsőé.

A virtuális mondatok észlelése

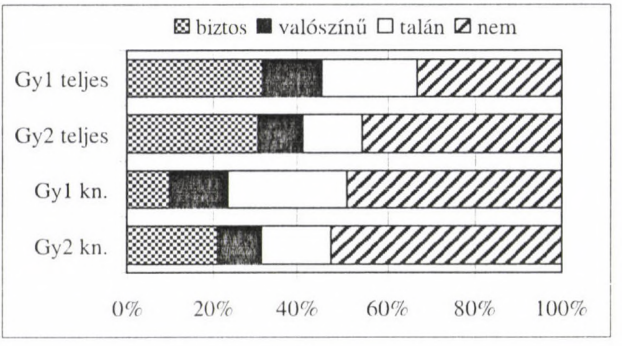
A felnőtt kísérleti személyek a szövegek különböző helyein észleltek virtuális mondatvéget, a jelöléseket összesítettük, és azt találtuk, hogy az első szövegben 75, a másodikban 62 olyan hely volt, ahol legalább egyetlen személy jelölt mondathatárt. Mivel a tesztlapok (és a hangzó anyagok is) tartalmazták az interjúkészítő kérdéseit, a beszélőváltáskor történő jelöléseket figyelmen kívül hagyva is összesítettük a jelöléseket, hiszen ezeken a helyeken egyértelmű volt a szegmentálás percepciós szempontból (3. ábra). Ha pusztán a gyermek szövegén belüli jelöléseket tekintjük, az első szövegben 51 helyen, a másodikban pedig 57 helyen észlelt legalább egy adatközlő mondathatárt. Az alacsonyabb KFM-értékű szövegben egy adatközlő átlagosan 17 mondatvéget jelölt be (átlagos eltérés: 4,1), a magasabb KFM-értékű szövegben pedig átlagosan 21 mondatot állapítottak meg (átlagos eltérés: 2,8). Az egyes kísérleti személyek észleletei között igen nagy egyéni különbségeket fedezhettünk fel. A Gy1 szövegben volt olyan személy, aki 29 helyen észlelt mondatvéget, és volt olyan, aki csupán 8 helyen. A Gy2 szövegben a legtöbb jelölés 28 volt, a legkevesebb 15 jelölés.



3. ábra
A virtuális mondatok észlelése a két gyermek szövegében

Megvizsgáltuk, hogy mennyire egyöntetű a hallgatók észlelése, illetve szegmentálása. Aszerint, hogy az adatközlők hány százaléka jelölt ugyanazon

a helyen mondatvéget, négy kategóriát állítottunk fel. „Biztos" mondatvégnek tekintettük azokat a helyeket, ahol az adatközlők nagy többsége, minimum 90%-a tett határjelölést. „Valószínű" mondatvégnek ítéltük, ha az adatközlők 60-90%-a egyöntetűen jelölt. „Talán" virtuális mondat vége található azokon a helyeken, ahol az adatközlők 20–60%-ban így gondolták. Nem vettük mondatvégnek azokat a helyeket, ahol az adatközlők kevesebb mint 20%-a húzott be jelölést. Az összes lehetséges helyet besoroltuk valamelyik kategóriába, az eredményeket a 4. ábra mutatja.



4. ábra
A mondatvégek jelölésének egyöntetűsége a gyermekek teljes és a kísérletvezető kérdései nélküli (kn.) szövegében

Megállapítottuk az észlelés egyöntetűségét kétféle módon: egyrészt a teljes szövegekben történt jelöléseket, másrészt azokat a jelöléseket, amelyek a gyermekek összefüggő szövegében jelentek meg, tehát eltekintettünk a beszélőváltásokkor megjelenő jelölésektől. Teljesen egyöntetű észlelés, vagyis 100%-os egyetértés a Gy1 teljes szövegben 8 helyen volt (ami az összes hely 10,7%-a), és mind olyan helyzetben fordult elő, ahol beszélőváltás történt. Azokban a szövegekben, ahol a gyermek a kísérletvezető kérdéseire néhány szóval, rövidebb, egyszerűbb szerkezetekkel válaszol, a hallgatók az egyes beszélők fordulói mentén szegmentálják a szöveget, így a (KFM-módszerhez szükséges) virtuális mondatokra osztás kevésbé problematikus. A Gy2 teljes szövegében 6 helyen jelölt az összes adatközlő mondatvéget (ez a lehetséges helyek 8,8%-a), de ebből 2 a gyermek szövegén belül jelent meg, vagyis nem a beszélőváltás miatt jelölték az adatközlők. Összehasonlítva felnőttekre vonatkozó eredményekkel azt láthatjuk, hogy a gyermekek beszédének szegmentálásakor valamivel kisebb volt az egyetértés a hallgatók között, mint a felnőttek spontán narratívája esetén. Gósy (2003a) vizsgálatában ugyanis az összes kísérleti személy a jelölések 33%-ánál hozott azonos ítéletet. Egy másik, ugyancsak felnőttek virtuális mondatait elemző kísérletben (Markó 2010)

a határjelölések 8%-ában értett egyet minden tesztelő, vagyis ott hasonló eredmény született, mint a jelen kísérletben.

Az első gyermek esetében az adatközlők nagy többsége, 90%-a egyöntetűen tagolta a szöveget a lehetséges helyek 32%-ánál. Lehetséges helyeknek (100%) tekintjük a szövegben azokat a pontokat, ahol legalább egy hallgató észlelt mondathatárt. A második szövegben 31%-os volt az ilyen jelölések aránya, vagyis közel azonos a két teljes szövegben. Ha eltekintünk a párbeszédes részektől, az első gyermek (Gy1 kn.) szövegében az összes lehetséges hely 10%-ánál tapasztalhatunk ilyenfajta egyöntetűséget az észlelésben, a másodiknál (Gy2 kn.) pedig 21%-ban. Ez azt mutatja, hogy az első szövegben az adatközlők nagyrészt az interjúkészítő kérdéseire támaszkodva szegmentálták a szöveget, és ez segítette elő az egyöntetű jelöléseket; a második szövegben pedig már a gyermek monologikus szövegében is egyöntetűen elkülönítettek virtuális mondatokat. Mindkét szövegben nagyfokú volt a hallgatók bizonytalansága a mondatészlelés során, ezt mutatja a sok olyan hely, ahol csupán egyetlen személy észlelt mondatvéget. Ilyen az első szövegben 10 esetben, a másodikban 15 esetben fordult elő. Ezek adódhattak a kísérleti személyek figyelmetlenségéből is. Számos olyan pontja van a szövegeknek, ahol a hallgatóknak csupán kis része (20% alatt) jelölt mondatvéget. A szövegekben az összes lehetséges hely mintegy 45%-át teszik ki az ilyen jelölések. A „valószínű" és a „talán" kategóriába tartozó jelölések megjelenése (12%, illetve 19%-ban) is a spontán szövegek virtuális mondatokra bontásának nehézségeit és a hallgatók percepciós érzékenységének különbségeit mutatja.

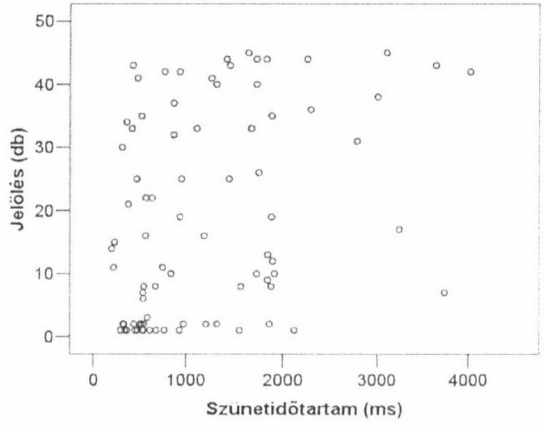
A szegmentálás a beszédszakaszok és a szünetek függvényében

Tanulmányoztuk a beszédszakaszok és az észlelés összefüggéseit, vagyis azt, hogy a jelölések milyen arányban történtek a beszédszakaszok végén. Azt állapíthattuk meg, hogy a nyelvtanilag egyszerűbb felépítésű szövegben (Gy1) a beszédszakaszok 67%-ánál, a komplexebb szövegben (Gy2) a beszédszakaszok 42%-ánál jelölt mondatvéget legalább egy személy. Korábbi, felnőttekre vonatkozó adatok szerint a hallgatók a beszédszakaszok 23%-ánál szegmentáltak egyöntetűen spontán narratívában, és 70%-ánál felolvasott szövegben (Váradi 2009). Ez abból fakad, hogy felolvasás során a beszéd létrehozása folyamatosabb, mint spontán narratívában, a beszélő (felolvasó) a leírt szöveg segítségével a szüneteket igyekszik oda időzíteni, ahol egy-egy szemantikai-szintaktikai egység lezárul. A spontán beszéd során a gondolatok megtervezése és nyelvi formába öntése kevésbé zökkenőmentes, a szünetek (és a megakadások) jelzik az egyes folyamatok és szintek működési összehangolatlanságát (Gósy 1998). Ennek következménye, hogy a spontán narratívákat általában több szünet tarkítja, mint a felolvasást, amelyek azonban nem feltétlenül esnek szerkezethatárra. Valószínűleg, mivel komplexebb volt a második gyermek beszéde (vagyis bonyolultabb gondolatokat kellett grammatikailag átalakítania), több gondot jelentett számára a tervezés és a kivitelezés, így több szünetet produkált, és több beszédszakaszra osztható a szövege. Ezek-

nek a szüneteknek nagy százaléka viszont nem egy-egy szemantikai és szintaktikai szerkezet határára esett, ezért a hallgatók ezekben az esetekben nem jelöltek virtuálismondat-véget. A szüneteknél erősebb határjelzőnek bizonyultak tehát a szemantikailag és szintaktikailag lezárt egységek.

A hallgatók tehát nem minden beszédszakasz végét (szünet megjelenését) tekintették határjelző funkciójúnak, de ahol mondatvéget jelöltek, az nagymértékben esett szünetre, azaz beszédszakasz végére. Az egyszerűbb szerkezetű szövegben (Gy1) a jelölések 91%-a, az összetettebb szövegben (Gy2) a jelölések 78%-a szakasz végére, vagyis szünetre esett. A többi jelölés beszédszakaszon belül történt, szünet-előfordulás nélkül. Ezekben az esetekben is a szemantikai és a szintaktikai szerkezetek játszottak nagy szerepet.

Feltételeztük, hogy a hosszabb szünetek gyakrabban jelennek meg a virtuális mondatok között, és jobban keltik a befejezettség érzetét a hallgatókban, akik így többször jelölnek mondathatárt a hosszabb szünetek alatt. Az első szöveg esetén nem találtunk egyértelmű függvénykapcsolatot (korrelációanalízis) a szünet hossza és a jelölések száma között, a második szöveg esetén azonban közepes, szignifikáns összefüggés volt közöttük (r = 0,462; p = 0,01). Összesítve a két szöveg adatait (5. ábra) csupán gyenge, szignifikáns összefüggés mutatható ki a gyermekek beszédében a szünet időtartama és a virtuálismondat-észlelés között (r = 0,248; p = 0,05).



5. ábra
A szünet időtartama és a jelölések száma közötti összefüggés

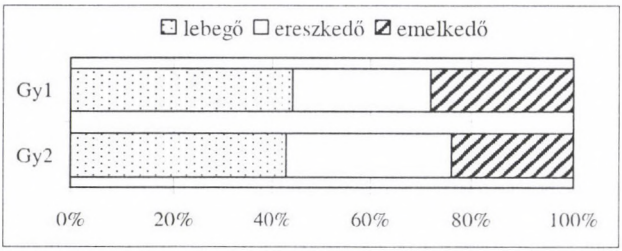
A felnőttek körében végzett kutatások is kitértek ennek a kérdésnek a vizsgálatára. Az egyik kísérletben közepesen erős összefüggést mutattak ki (Gósy 2003a), egy másikban pedig nem találtak korrelációt a szünethossz és a jelö-

lések száma között (Váradi 2009). A szünetek tehát segítik a mondanivaló szegmentálását, de az nem feltétlenül az időtartamukkal van összefüggésben.

Az alaphangmagasság változása és a mondatészlelés összefüggései

A beszélők a hanglejtéssel is képesek érzékeltetni egy-egy szövegegység lezárt vagy le nem zárult voltát. Külföldi és magyar kísérletek is igazolták, hogy az alaphangmagasság csökkenése használatos a befejezettség érzékeltetésére (Swerts et al. 1994; Moosmüller 1997; Gósy 2004). Egy összehasonlító esettanulmány megállapította, hogy a mondathatárok hanglejtéssel való érzékeltetése hatvan éve egyértelműbb volt, mint napjainkban, amikor is gyakran hiányzik a világos, tagoló intonáció, az ereszkedő hanglejtést sokszor a szintaktikai, szemantikai szerkezetlánc végén hirtelen emelkedő váltja fel (Vallent 2005). A közbeszédben és egyes nyelvészeti munkákban olykor negatívan minősítik a közlés végi dallamemelést, kiváltképp a szóközárat. Erről egy tanulmány rögzített beszédkorpusz alapján kimutatta, hogy valóban gyakoribb spontán beszédben, mint olvasásban, de a többi dallamtípushoz képest nem jelenik meg kirívó arányban (Markó 2009).

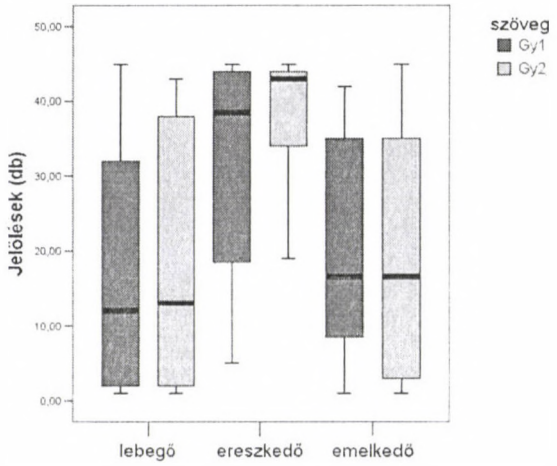
Megfigyeltük, hogy a gyermekek hogyan élnek az alaphangmagasság változtatásával a spontán beszédük során. Mindkét szövegben hasonló arányban jelentek meg a különböző dallammenet-típusok: a lebegő, az emelkedő és az ereszkedő. Az első szövegben (Gy1) a szakaszvégi hangmenetek 44%-a lebegő, 28%-a ereszkedő és 28%-a emelkedő volt. A második szövegben (Gy2) 43%-ban találhattunk lebegő, 33%-ban ereszkedő és 24%-ban emelkedő dallammeneteket az elemzett beszédszakaszok végén (6. ábra).



6. ábra
A különböző típusú dallammenetek a két szövegben

Elemeztük, hogy a három dallammenettípus esetén milyen mértékben észleltek a hallgatók virtuálismondat-határt (7. ábra). A legtöbb jelölés azokon a helyeken jelent meg, ahol a dallammenet ereszkedő volt. Ez mind a két szövegben így volt: a Gy1 szövegben az ereszkedő dallammenetek esetében az átlagos jelölések száma 32 volt, ami azt jelenti, hogy ezeken a helyeken átlagosan 32 személy jelölt mondathatárt egyöntetűen a 45-ből. A Gy2 szövegben az ereszkedő dallammeneteknél átlagosan 38 személy észlelt határokat,

vagyis ennél a dallamtípusnál jelöltek a legegyöntetűbben. Ezt követi az emelkedő (átlagosan 20), majd a lebegő dallammeneteknél (átlagosan 18) történt egyöntetű jelölések száma. Az ereszkedő hanglejtésforma esetén észleltek tehát a legnagyobb arányban mondatvéget a hallgatók, az emelkedő és a lebegő hangmenet kevésbé keltette a befejezettség érzetét.



7. ábra
A virtuális mondatok észlelése a dallammenetek szerint

A virtuális mondatok közös jellemzői

Szemügyre vettük azokat a helyeket a szövegben, amelyeket a hallgatók minimum 90%-a megjelölt, vagyis amelyeket a „biztos" mondatvég kategóriába soroltunk. Összesen 17 ilyen pozíciót találtunk a két szövegben. Ezeknek a közös jellemzőik, hogy beszédszakasz végére estek, vagyis szünet követte őket. A követő szünetek időtartama relatíve hosszú volt: az első szövegben átlagosan 2305 ms (774–4021 ms), a második szövegben átlagosan 1492 ms (430–3103 ms). Összehasonlításképpen az első szövegben az összes szünet átlagos időtartama 1114 ms, a másodikban 718 ms, tehát a „biztos" mondatvéghelyeken kétszer olyan hosszú szünetek fordultak elő, mint az átlagos szünetidőtartam. Mindegyik „biztos" virtuális mondat vége grammatikai szerkezetek határán fordult elő, amely önálló szemantikai egységet is alkotott. Például az első gyermek szövegében a következő két virtuális mondat között egyöntetűen jelöltek határt a hallgatók (a tagoló jelölést függőleges vonallal jelöltük): *a barátaimhoz közel lakunk és néha át szoktunk menni | és az egyik barátomat úgy hívják hogy J. a másikat Ö. és a harmadikat B.* A második gyermek szövegéből egy példa: *A. néni nem engedi hogy hogy annyi-*

szor hogy játsszunk vele merthogy most ragasztotta meg | és a M.-val nagyon szokat szoktunk játszani rajzolni repülőket hajtogatni. Az egyszerűbb szerkezetekből álló szövegben *és* kötőszó követte ezeket a virtuális mondatokat, a nyelvtanilag bonyolultabb szövegben pedig *és, aztán* vagy egy újabb virtuális mondat következett. A vizsgált szakaszvégek dallammenete 58%-ban ereszkedő, 24%-ban emelkedő és 18%-ban lebegő. A mondatzárás megállapításában tehát szerepet játszott a szintaktikai, szemantikai szerkezetek lezártsága, a szünettartás, valamint az intonációs megvalósítás.

A szegmentálás nehézsége a hallgatók szerint

A hallgatóktól a feladat elvégzése után azt kértük, hogy ítéljék meg, hogy melyik szövegben (Gy1 vagy Gy2) volt nehezebb bejelölni a virtuális mondatokat. Azt feltételeztük, hogy az alacsonyabb KFM-értékű szöveget lesz könnyebb tagolni, hiszen a rövidebb, egyszerűbb nyelvtani szerkezetek megkönnyítik az észlelést-megértést. A hipotézisünk beigazolódott, ugyanis a kísérleti személyek kétharmada (67%) úgy értékelte, hogy a Gy2 szöveg esetében okozott nagyobb gondot a szegmentálás. A hallgatóknak ez az ítélete ellentmond azzal az eredménnyel, hogy melyik szövegben döntöttek egyöntetűbben a virtuális mondatok határairól. Bár az egyszerűbb szöveget (Gy1) ítélték könnyebben tagolhatónak, ebben a szövegben mutatkozott több bizonytalanság a mondatészlelés során, és a grammatikailag összetettebb szövegben (Gy2) született egyöntetűbb eredmény.

Következtetések

Kísérletünkben azt vizsgáltuk, hogy a gyermekek milyen eszközöket használnak beszédük tagolására, illetőleg hogy a beszéd szerkesztettsége összefüggésben áll-e a szegmentálással. A spontán narratívák virtuális mondatokra bontása a felnőttek összetett közléseiben problematikusabbnak tűnik, mint a gyermekek kevésbé komplex beszédében. A korábbi, felnőttekre vonatkozó elemzések kimutatták, hogy a mondat szintaktikai struktúrája és a prozódiai megvalósítás között nem mindig tökéletes a kapcsolat, és hogy a szemantikai koherencia a döntő a hallgatók észlelésében (Frazier et al. 2003). Felmerül a kérdés, hogy vajon a gyermekek ugyanazokat a szupraszegmentumokat alkalmazzák-e tagoló funkcióban, mint a felnőttek. A jelen vizsgálatban a tagoló funkciójú szupraszegmentumokat a szintaktikai összetettség függvényében – egy nyelvtanilag egyszerűbb és egy komplexebb felépítésű szövegben – elemeztük. A hallgatók az egyszerűbb szöveget (Gy1) ítélték könnyebben tagolhatónak, a mondatészlelés azonban egyöntetűbb volt a második, komplexebb szövegben (Gy2). Ez abból eredhet, hogy a jobb beszédteljesítményt mutató gyermek hosszabb, bonyolultabb közléseket produkált, beszéd- és artikulációs tempója gyorsabb volt, ami megnehezíthette a hallgatók észlelését-megértését, és több odafigyelést igényelt tőlük. Ezzel egy időben viszont pontosabban alkalmazta a beszéd tagoló funkciójú eszközeit, ami segíthette a bonyolult tartalmak szegmentálását. Ezek az

eszközök megegyeznek a felnőttek beszédében tapasztaltakkal: a szintaktikai, a szemantikai és a prozódiai tényezők együttesen játszottak szerepet, bár eltérő jelentőséggel. A mondatvégek észlelése az esetek nagy részében beszédszakaszok végén történik, tehát ahol szünetet tart a beszélő. A hosszabb szünetek nem minden esetben a virtuális mondatok végén fordultak elő, de ahol az adatközlők nagy része mondathatárt jelölt, ott relatíve hosszú szünet volt. A szünethossz és a mondatvégjelölések között csupán gyenge összefüggést találtunk. A szakaszzáró dallammenetek közül főként az ereszkedő kelti a befejezettség érzetét. A gyermekek valószínűleg még bizonytalanabbul alkalmazták ezeket az eszközöket, mint a felnőttek, ezért lehet az, hogy a hallgatók a gyermekek beszédét kevésbé egyöntetűen szegmentálták, mint azt a korábbi kísérletekben a felnőtteknél megfigyelték. Mint említettük, a közlésegységek fejlődési mutatójának meghatározásához a gyermekek spontán beszédét virtuális mondatokra kell osztanunk. A gyermeki beszéd szegmentálásához tehát ugyanazokat a tagoló funkciójú eszközöket használhatjuk fel, mint amelyeket a felnőttek beszédének tagolásakor is. A mondanivaló tagolásának (ösztönös) megtervezése valószínűleg összefügg azzal, hogy milyen a gyermek anyanyelvi szintje. A jobb beszédkészség tehát megmutatkozik abban, hogy a gyermek bonyolult szerkezeteket tartalmazó, összefüggő spontán beszéd létrehozására képes, és hogy a mondanivalóját hatékonyabban, a hallgató befogadását megkönnyítve tudja tagolni.

Irodalom

A magyar nyelv helyesírási szabályai 1984. Akadémiai Kiadó, Budapest.

Boersma, Paul – Weenink, David 2011. *Praat: doing phonetics by computer.* (Version 5.3.02) [Computer program]. http://www.praat.org (A letöltés ideje: 2011. október 10.)

Crystal, David 2003. *A nyelv enciklopédiája.* Osiris Kiadó, Budapest.

Érsok Nikoletta Ágnes 2006. Szóbeliség és/vagy írásbeliség. *Magyar Nyelvőr* 130. 165–176.

Frazier, Lyn – Clifton, Charles Jr. – Carlson, Katy 2003. Don't break, or do: Prosodic boundary preferences. *Lingua* 1. 1–25.

Gerebenné Várbíró Katalin – Gósy Mária – Laczkó Mária 1992. *Spontán beszédmegnyilvánulások szintaktikai elemzése DSS technika segítségével.* Kézirat. Budapest.

Gósy Mária 1981. A beszédhang kialakulása a gyermeknyelvben. *Magyar Fonetikai Füzetek* 7. 67–91.

Gósy Mária 1998. A beszédtervezés és a beszédkivitelezés paradoxona. *Magyar Nyelvőr* 122. 3–15.

Gósy Mária 2003a. Virtuális mondatok a spontán beszédben. *Beszédkutatás 2003.* 19–44.

Gósy Mária 2003b. A spontán beszédben előforduló megakadásjelenségek gyakorisága és összefüggései. *Magyar Nyelvőr* 127. 257–277.

Gósy Mária 2004. *Fonetika, a beszéd tudománya.* Osiris Kiadó, Budapest.

Gósy Mária 2005. *Pszicholingvisztika.* Osiris Kiadó, Budapest.

Gósy, Mária – Kovács, Magdolna 2007. Virtual sentences of spontaneous speech: boundary effects of syntactic-semantic-prosodic properties. In Gardner-Bonneau, Daryle – Blanchard, Harry E. (eds.): *Human factors and voice interactive systems.* Second edition. Springer, New York, 361–379.

Horváth Viktória 2006. A spontán beszéd és a beszédfeldolgozás összefüggései gyerekeknél. *Beszédkutatás 2006.* 134–146.

Horváth Viktória 2009. *Funkció és kivitelezés a megakadásjelenségekben.* Doktori disszertáció. ELTE, Budapest.

Juhász József – Szőke István – O. Nagy Gábor – Kovalovszky Miklós (szerk.) 1987. *Magyar értelmező kéziszótár.* Akadémiai Kiadó, Budapest.

Kassai Ilona 1998. *Fonetika.* Nemzeti Tankönyvkiadó, Budapest.

Keszler Borbála 1983. Kötetlen beszélgetések mondat- és szövegtani vizsgálata. In Rácz Endre – Szathmári István (szerk.): *Tanulmányok a mai magyar nyelv szövegtana köréből.* Tankönyvkiadó, Budapest, 164–202.

Kohler, Klaus J. 1983. Prosodic boundary signals in German. *Phonetica* 40. 89–134.

Kugler Nóra 2000. A mondattan általános kérdései. In Keszler Borbála (szerk.): *Magyar grammatika.* Nemzeti Tankönyvkiadó, Budapest, 369–394.

Lanstyák István 2009. *A magyar beszélt nyelv sajátosságai.* Stimul, Pozsony.

Lee, Laura L. – Canter, Susan M. 1971. Developmental Sentence Scoring: A clinical procedure for estimating syntactic development in children's spontaneous speech. *Journal of Speech and Hearing Disorders* 36. 315–340.

Lehiste, Ilse 1979. Perception of sentence and paragraph boundaries. In Lindblom, Björn – Öhman, Sven (eds.): *Frontiers of speech communication research.* Academic Press, London–New York–San Francisco, 191–201.

Lengyel Zsolt 1981. *Tanulmányok a nyelvelsajátítás köréből.* Akadémiai Kiadó, Budapest.

Levelt, Willem J. M. 1989. *Speaking. From intention to articulation.* The MIT Press, Cambridge, MA.

Markó Alexandra 2005. *A spontán beszéd néhány szupraszegmentális jellegzetessége. Monologikus és dialogikus szövegek összevetése, valamint a hümmögés vizsgálata.* Doktori disszertáció. ELTE, Budapest.

Markó Alexandra 2009. Stigmatizált hanglejtésforma a spontán beszédben. *Beszédkutatás 2009.* 88–106.

Markó Alexandra 2010. A prozódia szerepe a spontán beszéd tagolásában. *Beszédkutatás 2010.* 82–99.

S. Meggyes Klára 1971. *Egy kétéves gyermek nyelvi rendszere.* Akadémiai Kiadó, Budapest.

S. Meggyes Klára 1981. A mondat- és szövegalkotás néhány sajátossága hatéves gyermek képleírásában. *Pszichológia* 1. 541–551.

Moosmüller, Sylvia 1997. The revelance of fundamental frequency contour for interruptions: A case study of political discussions in Austria. In Kotthoff, Helga – Wodak, Ruth (eds.): *Communicating Gender in context.* John Benjamins, Amsterdam, Philadelphia, 401–420.

Neuberger Tilda 2011. Gyermekek spontán beszédének szerkesztettsége és folyamatossága. *Beszédkutatás 2011.* 83–95.

Piaget, Jean 1970. *Válogatott tanulmányok.* Gondolat Kiadó, Budapest.

Swerts, Marc – Collier, René – Terken, Jacques 1994. Prosodic predictors of discourse finality in spontaneous monologues. *Speech Communication* 15. 79–90.
Vallent Brigitta 2005. A spontán beszéd ötven éve és ma. Esettanulmány. *Beszédkutatás 2005*. 99–111.
Váradi Viola 2008. A virtuális mondatok műfaji meghatározottsága. *Beszédkutatás 2008*. 134–147.
Váradi Viola 2009. A spontán beszéd mondatai a műfajok és az életkor tekintetében. In Keszler Borbála – Tátrai Szilárd (szerk.): *Diskurzus a grammatikában – grammatika a diskurzusban*. Tinta Könyvkiadó, Budapest, 65–73.
Vértes O. András 1953. *A gyermek nyelve*. Kézirat. Budapest.
Wacha Imre 1988. Élő nyelvi (spontán) szövegek megnyilatkozásainak (szintaktikai) vizsgálati szempontjaihoz (a gazdagréti kábeltelevízió élő nyelvi felvételei alapján). In Kontra Miklós (szerk.): *Beszélt nyelvi tanulmányok*. MTA Nyelvtudományi Intézet, Budapest, 102–158.

A BESZÉLŐVÁLTÁSOK SAJÁTOSSÁGAI ÓVODÁSKORÚ GYERMEKEK TÁRSALGÁSÁBAN

Szabó Ágnes – Tóth Andrea

Bevezetés

A társalgás a legáltalánosabb beszédhelyzet, amelyben megjelenik a beszédpartnerhez/beszédpartnerekhez történő alkalmazkodás szükségessége. A nyelvi interakciók jellemzője, hogy a folyamatban részt vevők az átadó és a befogadó, illetve szóbeli kommunikáció esetén a hallgató szerepét töltik be. A nyelvi folyamat tehát percepciós és produkciós mechanizmusok egyidejű mozgósítását jelenti. Szóbeli interakció esetén az üzenet átadásának folyamata a következőkben határozható meg. 1. A beszélő beszédprodukciója során a makro- és mikrotervezési fázisokon keresztül eljut a gondolattól a nyelvi megformálás szakaszáig, amelyet aztán artikulációs szerveinek segítségével meghangosít. 2. A beszélő mint címzett beszédpercepciós mechanizmusának segítségével fölismeri, beszédhangokként, fonémákként azonosítja, majd nagyobb egységekké állítja össze a hallott akusztikus jeleket, ezáltal megtörténik a közlés dekódolása (Gósy 2005: 120). E két folyamat egy időben zajlik, miközben a befogadó és az átadó személye változik. Tekintettel arra, hogy ez kétoldalú művelet, sikerességéhez szükséges az átadó és a befogadó együttműködése.

A nemzetközi szakirodalomban a beszédpartnerhez való alkalmazkodást már számos nézőpontból vizsgálták. Bell (1984, idézi: Bata 2009a) szerint a stílus változása a résztvevőktől függ, vagyis a beszélő igazodik a címzetthez, sőt beszédére még azok a személyek is hatással vannak, akik nem címzettjei az üzenetnek, de valamilyen módon meghallják azt. A címzetthez történő igazodás vetette fel azt a kérdést, hogy hogyan alkalmazkodnak a felnőttek a gyermekekhez a beszéd során. Az eddigi kutatások (Sundberg 1999; Liu et al. 2003; Koós 2010) rámutattak, hogy a beszélők másképpen beszéltek a gyermekekhez, mint a felnőttekhez, tehát nemcsak a beszédpartner, de a beszélő életkora is hatással van a beszédalkalmazkodásra.

A társalgás során az aktuális beszélők folyamatosan váltják egymást, ezt beszélőváltásnak (turn-taking, Turnübergabe) nevezzük (Sacks et al. 1974; Brown–Yule 1989; Iványi 2001; Bata–Gráczi 2008). A beszélőváltás tehát a beszélők meghatározott szabályok szerinti váltakozása (Sacks et al. 1974). A társalgás alapegysége a beszédlépés vagy forduló (turn), amely egy beszélőtől származó megnyilatkozás (Boronkai 2009). Ennek során a társalgás egyik résztvevője beszél addig, amíg át nem adja, vagy át nem veszik tőle a

megszólalás jogát. Az aktuális beszélő pragmatikai, szintaktikai, prozódiai és egyéb (nonverbális) jelekkel jelezheti partnere számára a beszédlépésének végét. A beszélőváltás történhet: 1. önkiválasztással (szóátvétel), ilyenkor az nyer jogot a beszédlépés megtételére, aki először kezd el beszélni; 2. külválasztással (szóátadás), amikor az aktuális beszélő jelöli ki a következő beszélőt; 3. ha sem önkiválasztás, sem külválasztás nem történik, akkor az eredeti beszélő folytatja a beszédlépését (Boronkai 2009).

A társalgások fonetikai szempontú vizsgálatáról számos nemzetközi és hazai publikáció született. Bosch és munkatársai (2005) kutatásukban például a társalgások temporális sajátosságait vizsgálták telefonbeszélgetések és személyes társalgások elemzésével. Foglalkoztak a szünetek és az egyszerre beszélések jellegzetességeivel, elemzésükben a társalgás típusától függően eltéréseket találtak a vizsgált jelenségeknél: a személyes kommunikációban ritkábban fordult elő egyszerre beszélés és hallgatás, mint a telefonbeszélgetésekben. A beszélőváltások időzítési kérdéseivel foglalkozott Heldner és Edlung (2010) is. Elemzésükben kísérletet tettek a beszédlépések között eltelő minimális szünetidőtartamok jellemzésére.

Gósy (2005) kutatásában a néma szünetek elemzésének problémáira hívja fel a figyelmet. A szünetek előfordulását két okra vezetik vissza, a lélegzetvételre és az értelmi tagolásra. A pszicholingvisztikai szakirodalom több egyéb szünetfajtát is említ, a gondolkodási szünetet, a hatásszünet; amelyek más-más funkciót tölthetnek be a beszédprodukcióban. Markó (2006) kutatásában egyetemi hallgatók beszélőváltásait elemezte, főként fonetikai szempontból. Négy résztvevős társalgást vizsgált, melyet narratívákkal is összevetett. Eredményei rámutattak, hogy a szünetek (a társalgási egységen belüli beszédjel-kimaradás) és a hallgatások (a társalgási egységek határán lévő beszédjel-kimaradás) között statisztikailag igazolható különbségek vannak a társalgásban. Elemezte az úgynevezett „hosszú szüneteket" és a beszélők egyéni „toleranciaküszöbét". Vizsgálta továbbá az egyszerre beszélések formai sajátosságait is. Evellei (2009) háromfős társalgások segítségével elemezte a szünetek szerepét. Hipotézise, amely szerint a szünetek szerepet kapnak a társalgás beszélőváltási mechanizmusaiban, s ez a szerep szünettípusonként változó, részben igazolódott. Bata (2009a; 2009b) a társalgás fonetikai jellemzőinek alakulását a beszédpartnerek életkorának, illetve az ismeretségnek a függvényében vizsgálta. Kutatásában a beszédlépések és a beszélőváltások jellegzetességeit elemezte háromfős társalgások feldolgozásával. Eredményei igazolták, hogy az egyszerre beszélések mennyisége nem függ a beszédpartnerek életkorától, és ezek nagy része együttműködést feltételez, tehát nem versengés miatt jön létre.

A társalgási szabályok elsajátítása már gyermekkorban megkezdődik, hiszen a gyermek szocializációja során környezetétől tanulja a megfelelő mintákat. Az óvodáskorú gyermekek már rendelkeznek az alapvető társalgási szabályok ismeretével (Boronkai 2009).

A jelen kutatásban célunk az óvodáskori beszélőváltások elemzése, főként fonetikai aspektusból. Kutatásunkban mindemellett arra is keressük a választ, hogy a különböző beszélőváltási formákhoz milyen funkciók kapcsolhatók, illetve hogy a beszélőváltások sajátosságai milyen alapvető társalgási stratégiákra utalnak. Hipotézisünk szerint a gyermekek társalgási stratégiái bizonyos fokig eltérnek a felnőttekétől (Markó 2006), hiszen az óvodás gyermekek az anyanyelv-elsajátítás intenzív szakaszában vannak.

Anyag, módszer, kísérleti személyek

Kutatásunkban tíz óvodás lánygyermek társalgását elemeztük, életkoruk 4–5 év közötti volt. Mindnyájan magyar anyanyelvűek, ép hallásúak, anyanyelv-elsajátítási folyamatuk tipikus. A felvételeket óvodai fejlesztő szobában rögzítettük. A társalgások résztvevője minden esetben ugyanaz a (nő) felvételvezető (életkora 22 év) és két gyermek volt. A felvételvezető alapvetően a társalgást volt hivatott segíteni, résztvevői stratégiájának célja nem a folyamatos aktivitás, sokkal inkább a megakadó társalgás továbblendítése volt. A párok kiválasztása véletlenszerűen, számlapok húzásával történt. A hanganyagok bemutatkozásból, bábos szerepjátékból, hétköznapi témájú, életkorspecifikus társalgásból állnak. A felvételvezető napi programjukról, családjukról, szabadidős tevékenységükről kérdezte a gyermekeket. A bábjátékok alkalmával ismert meséket játszhattak el a gyermekek. A hanganyag azon részeit, amelyek nem spontán beszédet (mese visszamondása, bábjáték) tartalmaznak, kizártuk az elemzésből. Összesen 2 óra 40 percnyi hanganyagot elemeztünk, amelynek elkészítettük kézi átiratát, majd annotáltuk a Praat 5.2 szoftver segítségével (Boersma–Weenink 2010). Meghatároztuk a társalgásban részt vevők beszédének időtartamát, a hallgatások, az egyszerre beszélések és a szünetek időtartamát. Méréseket végeztünk a beszélőváltások előfordulási arányaira vonatkozóan, ill. osztályoztuk a beszélőváltásokat forma és funkció szerint.

Elemzésünkben a társalgásokat összetartozó egészként kezeljük, ezért eredményeink a társalgás mindhárom résztvevőjének beszélőváltásait tartalmazzák. Az elemzett csoportoknál azonban zárójelben (GY-vel jelölve) megadjuk, hogy az aktuálisan elemzett adatok hány százaléka köthető a gyermekekhez.

A beszélőváltásokat **formájuk** (az időzítés) alapján három fő csoportba soroltuk:

1. egyszerre beszélés;
2. hallgatást követő váltás (ha a két beszédlépés között eltelik valamennyi idő);
3. azonnali váltás (ha az új megszólaló az előző beszélő beszédlépésének végén kezdi saját beszédlépését).

Az egyszerre beszéléseket formailag további négy csoportba soroltuk:

a) vége előtti megszólalás (ha a beszélő korábban elkezdi saját beszédlépését, mint ahogy az előző beszélő befejezte volna);

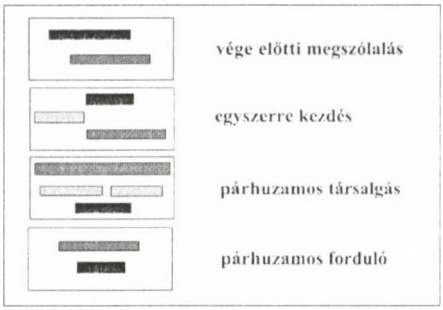
b) egyszerre kezdés (ha egyidőben két vagy több önkiválasztás történik);

c) párhuzamos társalgás (két beszélő egymással való társalgása, a megelőző beszélő beszédlépése alatt);

d) párhuzamos forduló (ha az új megszólaló átmenetileg párhuzamosan beszél beszédpartnerével, és a beszédlépését ő fejezi be előbb).

A beszélőváltások **funkciója** szerint öt csoportot különítettünk el:

1. visszacsatolás;
2. kérdés (ill. kérés, felszólítás);
3. válasz;
4. új téma kezdése;
5. régi téma folytatása (egy korábban elkezdett, de megszakított beszédlépés folytatása), vö. 1. ábra.



1. ábra
Az egyszerre beszélések lehetséges időbeli elrendeződése

Az egyszerre beszélések funkcióit Adda-Decker és munkatársai (2008) nyomán a közbeszóló szándéka szerint a következőképp csoportosítottuk:

a) „backchannel" – visszacsatolás, ide tartozik pl. az ilyen funkciójú hümmögés, az *igen, aha;*

b) „complementary" – kiegészítő szóátvétel, a közbeszóló segítő célzatú beszédlépése;

c) „anticipated turn taking" – együttműködő szóátvétel, amikor a beszélő egyértelműen jelzi beszédlépésének végét, leggyakrabban kérdés-válasz forma esetén;

d) „turn stealing" – erőszakos szóátvétel, amelynek célja a szó elragadása.

Az elemzésben a példák bemutatásához használt jelölések a következők:

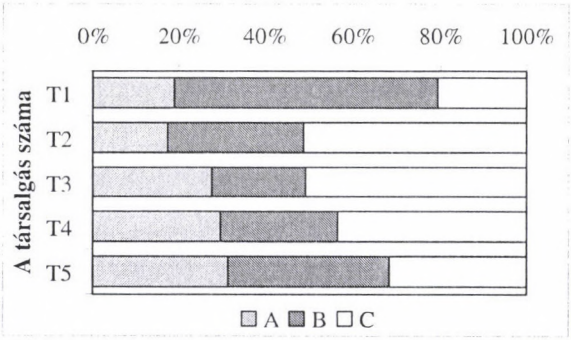
A: a felvételvezető közlése; B, C: az óvodás gyermekek közlései; az egyidejű beszédrészeket szögletes zárójelek jelzik; a beszédfordulók között előforduló szünetek időtartamát zárójelben közöljük.

A statisztikai elemzésekhez az SPSS 17.0 programot használtuk, a normál eloszlás sérülése miatt nem parametrikus tesztet (Kruskal–Wallis-próba) alkalmaztunk.

Eredmények

A beszélőváltások

Összesen 2491 db beszélőváltást adatoltunk, percenként átlagosan 18 beszélőváltás történt. A váltások mintegy fele, 50,3%-a volt olyan, amelyet óvodás kezdeményezett. A társalgások beszédidőit mutatja a 2. ábra.



2. ábra
Beszédidők aránya a társalgásokban
(A: felvételvezető; B, C: óvodások; T1–T5 a társalgás számát jelöli)

Formailag a leggyakoribb a hallgatást követő váltás volt (47,4%; GY: 71,0%), vagyis a két egymást követő beszédlépés között eltelt valamennyi idő. Például:

(1) A: most szombaton (1,5 s)
C: vasárnap meg megyek Bogiékhoz

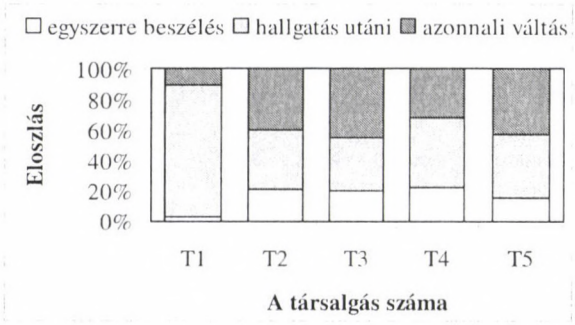
35%-ban (GY: 69%) a beszédlépések között nem volt jelkimaradás, azonnal váltották egymást a beszélők, például:

(2) B: szülinapja lesz a Nórinak
A: ühüm mesélte

A beszélőváltások 17,6%-a (GY: 81%) volt egyszerre beszélés, amikor két vagy több beszélő beszédlépése egyidőben zajlott. Például (a szögletes zárójelek jelölik az egyidejű beszédrészeket):

(3) C: vasárnap meg megyek [Bogiékhoz
B: én meg] nem tudok elmenni

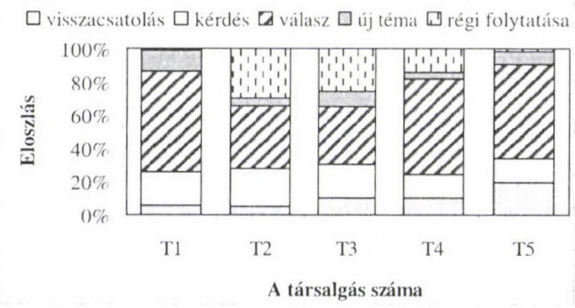
Az egyes társalgások beszélőváltásainak formai megoszlását szemlélteti a 3. ábra. Megfigyelhető, hogy T1 esetében magas a hallgatások száma, amely passzívabb társalgási részvételre utal.



3. ábra
A beszélőváltások formai eloszlása társalgásonkénti bontásban

Elemeztük a beszélőváltásokat funkciójuk szempontjából, megvizsgálva a szóátvételek valószínűsíthető okait (4. ábra). Elsőként a hallgatást követő váltások és az azonnali váltások funkcióit elemezzük. Az e típusú beszélőváltások formai megoszlásában a leggyakoribb a válasz (49,8%; GY: 25%), majd a kérdés (18,1%; GY: 88%) volt. Például:

(4) A: te miket kaptál
B: hát én kaptam amiket meséltem az a hello kittys táska meg a kutyus



4. ábra
A hallgatást követő és azonnali váltások funkcióinak eloszlása társalgásonkénti bontásban

Az együttműködés megvalósításában fontos szerepet játszik a visszacsatolás, amely a beszélőváltások 10%-át (GY: 37%) tette ki. Leggyakrabban a felvételvezető visszacsatolásai voltak ezek: hümmögések, az *igen, aha,* ill. idesoroltuk a „kérdésként" megismételt válaszokat is, például:

(5) B: Nekem nincs tesóm.
A: Nincs tesód?

A beszélő egy korábbi, megszakadt témát folytatott az esetek 14,6%-ban (GY: 64%), például:

(6) A: ó de aranyos neve van (0,7 s)
B: nekünk van papagájunk

7,2%-ban (GY: 85%) a beszélő új témát kezdett, például:

(7) B: nem ő fiú
A: mit csináltatok most fönt a csoportszobában?

Az egyszerre beszélések

Az egyszerre beszélések a beszélőváltások 17,6%-át adták. Átlagos időtartamuk 1109 ms volt, terjedelmük 101–11 580 ms közötti. A legrövidebb egy visszacsatoló funkciójú hümmögés, a leghosszabb egy párhuzamos társalgás volt.

Formai megoszlásukat tekintve a leggyakrabban a vége előtti megszólalások fordultak elő (45%; GY: 78%). A beszélő ilyenkor legtöbbször már elegendő információval rendelkezik, hogy beszédlépést kezdeményezzen, gyakori kérdés-válasz esetében. Például:

(8) A: és kertészkedni is [szoktatok?
B: szoktunk] ma is fogunk

Az esetek 40%-ában (GY: 71%) egy időben két önkiválasztás történt, például:

(9) C: répát zöldséget
B: [meg salátát]
A: [locsoljátok is?]

Kevésbé volt gyakori (13%; GY: 36%) a párhuzamos forduló, például:

(10) C: van egy kis [bokrunk] tulipánbokor
B: [aha]

Mindössze a beszélőváltások 2%-ában jelent meg a párhuzamos társalgás (GY: 93%).

Funkció szerinti bontásban (5. ábra) az egyszerre beszélések között az „erőszakos" szóátvétel dominált (38%; GY: 100%). Többségében az óvodások által kezdeményezett beszélőváltások tartoztak ide. E funkció versengő

társalgási stratégiát feltételez, hiszen a közbeszóló szándéka a szó magához ragadása, az együttműködési elv megsértésével. Például:

(11) C: amikor szoktunk menni sétálni [akkor
B: van egy] kis bokrunk

5. ábra
Az egyszerre beszélések funkciói

32%-ban (GY: 30%) történt kiegészítő szóátvétel, amely alapvetően az együttműködés érvényesülését támogatja. (Jelezhet ugyanakkor türelmetlenséget is.) Például:

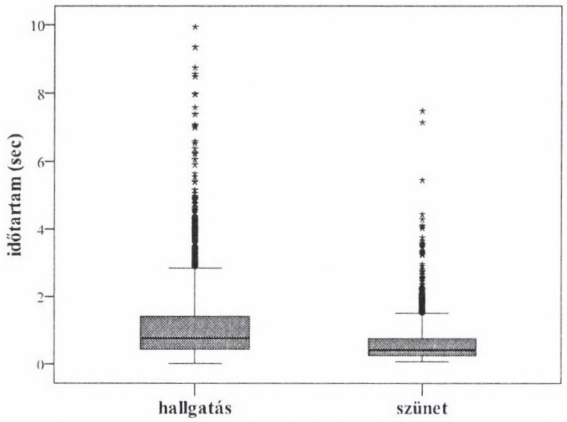
(12) C: most meg nagy [tulipán] fánk van
B: [tulipán]

Az egyszerre beszélések 22%-a (GY: 89%) funkciójában együttműködő szóátvétel, 8%-ban (GY: 49%) a funkció visszacsatolás volt.

Összehasonlítva a felnőtt korcsoport (Markó 2006) eredményeivel, látható, hogy az egyszerre beszélések száma több mint kétszeresére nőtt az óvodáskorú gyermekek társalgásában. A domináns funkciójú váltások – tudniillik az „erőszakos" szóátvételek – fakadhatnak az életkori sajátosságokból. Óvodáskorban a gyermekek gondolkodására az egocentrizmus jellemző, amely a társas – jelen esetben az együttműködés elve – helyett az egyéni érdeket helyezi előtérbe (Piaget 1993). Ebből arra következtethetünk, hogy ebben a fejlődési szakaszban a társalgási maximákat még felülírja az egocentrikus gondolkodás, vagyis azok még nem kellően kiforrottak.

A szünetek és a hallgatások

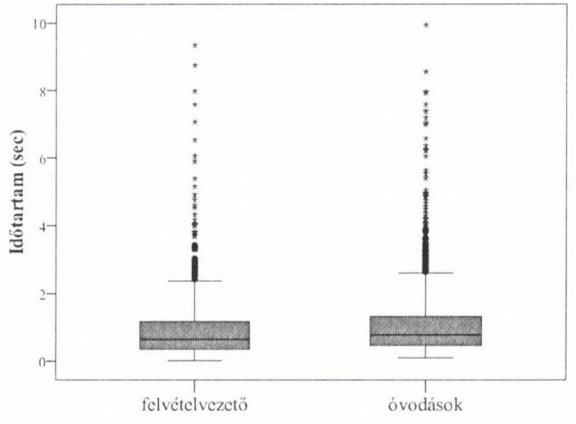
Elemeztük a társalgási egységeken belüli (szünetek) és az egységek határán levő beszédjel-kimaradások (hallgatások) időtartamát, beleértve a felnőtt beszélő megnyilatkozásait is (6. ábra). A szünetek átlagosan 730 ms, a hallgatások 1190 ms hosszúak voltak. Az időtartambeli eltérés statisztikailag is igazolható (Kruskal–Wallis-próba: $\chi^2 = 251{,}500$, df = 1, $p < 0{,}001$), a hallgatások időtartama szignifikánsan magasabb a társalgásokban.



6. ábra
A szünetek és hallgatások időtartamának eloszlása

A társalgásokban előfordultak az átlagosnál hosszabb szünetek, ezek időtartamát és formai jellemzőit szintén elemeztük. Markó (2006) 500 ms-ban határozta meg a „hosszú szünetek" időtartamát. Hanganyagunkban a szünetek 36,2%-a (379 db) volt 500 ms-nál hosszabb. Terjedelmük 500–7480 ms között szóródott, átlagos időtartamuk 1240 ms volt. Formailag e szünetek a leggyakrabban (65,3%) valamilyen szintaktikai kapcsolatban fordultak elő, például kötőszó, névelő után, szintagmán belül. 16,2%-uknál volt észlelhető valamilyen pragmatikai jegy, például a beszélő érzékeltette, hogy még nem ért a története végére. A „hosszú" szünetek 13,3%-a volt részben vagy egészben kitöltött, 7,8%-uknál emelkedő dallammenetet adatoltunk.

A hallgatások időtartama 172–9940 ms között szóródott. Megvizsgáltuk, hogy mikor válik zavaróvá a hallgatás. Feltételeztük, hogy a felvételvezető (felnőtt) és az óvodások (gyermek) ún. „toleranciaküszöbe" között szignifikáns eltérés lesz. Hipotézisünket a statisztikai elemzés igazolta (Kruskal–Wallis-próba: $\chi^2 = 18{,}919$, df = 1, $p < 0{,}001$). A toleranciaküszöb megmutatja, hogy a beszélő számára mikor válik zavaróvá a hallgatás. Kiszámításának alapja azon hallgatások időtartamainak átlaga, amelyeket a kérdéses beszélő által kezdeményezett beszédlépés követ. Az óvodások toleranciaküszöbe tehát magasabb, mint a felnőtté, a gyermekeket kevésbé zavarja a hallgatás. Jelen vizsgálatunkban a felnőtt korcsoportot csak a felvételvezető reprezentálta, ezért a toleranciaküszöbök nagyságának vizsgálatára további kutatások szükségesek. A toleranciaküszöbök értékeit a 7. ábra szemlélteti.



7. ábra
A beszélők toleranciaküszöbe

Következtetések

A jelen kutatásban óvodáskorú gyermekek beszélőváltásait elemeztük, főként fonetikai szempontból. A beszélőváltások formáját tekintve a leggyakrabban hallgatás utáni váltások fordultak elő, a váltások funkcióit tekintve a kérdés és a válasz volt domináns. A kérdés funkciójú beszédlépések többsége a felvételvezetőhöz, a válaszoké a gyermekekhez volt köthető. Az egyszerre beszélések formai megvalósulásai leggyakrabban az ún. „vége előtti megszólalás" és az „egyszerre kezdés" voltak. Az egyszerre beszélések többségét a gyermekek kezdeményezték.

Vizsgálatunkban a szünetek átlagos időtartama 730 ms, a hallgatások időtartama pedig átlagosan 1190 ms volt. A szünetek és a hallgatások közötti időtartambeli eltérés statisztikailag igazolható volt, a hallgatások időtartama szignifikánsan magasabb a társalgásokban.

Kutatási eredményeinket összevetve a felnőtt társalgások elemzésével a következő eltéréseket tapasztaltuk: a felnőtt adatközlőknél (Markó 2006) a szünetek átlagos időtartama 290 ms, a hallgatásoké pedig 648 ms. Elmondható tehát, hogy a gyermekek esetében a szünetek és a hallgatások időtartama szignifikánsan hosszabb, mint a felnőtteknél. Eredményeink igazolták, hogy az óvodások számára valószínűsíthetően kevésbé zavaró a hosszabb hallgatás.

Feltételeztük, hogy az életkori sajátosságok hatással vannak a társalgási stratégiákra. A felnőtt adatközlőknél (Evellei 2009) a megszólalások legnagyobb része visszajelzés, aránya a beszélgetésekben közel 73%. Az általunk elemzett társalgásokban a visszajelzések előfordulási aránya mindössze 8% volt. Hipotézisünk igazolódott: az óvodások társalgási stratégiáiban a versen-

gés dominált, ezt igazolja az egyszerre beszélések magas aránya (17,6%), amely a felnőtteknél adatoltak (6,87%) több mint kétszerese. A versengés nem volt jellemző a felnőtt beszédpartnerekre, Bata (2009) is rámutatott kutatásában, hogy az egyszerre beszélések nagy része együttműködést feltételez, nem pedig versengés miatt jön létre. Az óvodáskori versengés – versengő stratégia – okát pszichológiai alapon, a Piaget (1993) által egocentrizmusnak nevezett jelenséggel magyarázhatjuk. E szerint a műveletek előtti korban (1–6 éves korig) a gyermek nem képes más nézőpontját figyelembe venni és saját nézőpontjától független ítéletet alkotni. További kutatások szükségesek annak megállapítására, hogy melyik az az életkor, illetve az anyanyelv-elsajátításnak az a szakasza, amikor a gyermekek „átlépnek" a versengő stratégiából az együttműködést feltételezőbe, vagyis elsajátítják a grice-i maximákat (Grice 1975).

Irodalom

Adda-Decker, Martine – Barras, Claude – Adda, Gilles – Paroubek, Patrick – Boula de Mareüil, Philippe – Habert, Benoit 2008. Annotation and analysis of overlapping speech in political interviews. In: *Proceedings of the Sixth International Conference on Language Resources and Evaluation.* http://ftsp.mercubuana.ac.id/03/788_paper.pdf (A letöltés ideje: 2011. szeptember 1.)

Bata Sarolta 2009a. A társalgás fonetikai jellemzőinek alakulása a beszélők életkorának függvényében. In Váradi Tamás (szerk.): *III. Alkalmazott Nyelvészeti Konferencia.* MTA Nyelvtudományi Intézet, Budapest, 3–13.

Bata Sarolta 2009b. Beszélőváltások a beszédpartnerek személyes kapcsolatának függvényében. *Beszédkutatás 2009.* 107–121.

Bata Sarolta – Gráczi Tekla Etelka 2008. A beszédpartner életkorának hatása a beszéd szupraszegmentális jellegzetességeire. In Keszler Borbála – Tátrai Szilárd (szerk.): *Diskurzus a grammatikában, grammatika a diskurzusban.* Tinta könyvkiadó, Budapest, 74–82.

Boersma, Paul – Weenink, David 2010. *Praat: Doing phonetics by computer.* http://www.fon.hum.uva.nl/praat/ (A letöltés ideje 2011. január 11.)

Boronkai Dóra 2009. *Bevezetés a társalgáselemzésbe.* Ad Librum, Budapest.

Bosch, Louis T. – Oostdijk, Nelleke – Boves, Lou 2005. On temporal aspects of turn taking in conversational dialogues. *Speech Communication* 47. 80–86.

Brown, Gilliam – Yule, George 1989. *Discourse analysis.* Cambridge University Press, Cambridge–New York–Port Chester–Melbourne–Sydney.

Evellei Kata 2009. *A szünetek szerepe a társalgásban. Különös tekintettel a beszélőváltásokra.* Záródolgozat. ELTE, Budapest.

Gósy Mária 2005. *Pszicholingvisztika.* Osiris Kiadó, Budapest.

Grice, Herbert Paul 1975. Logic and conversation. In Cole, Peter – Morgan, Jerry L. (eds.): *Syntax and Semantics 3. Speech Acts.* Academic Press, New York, 41–58. (Magyarul: A társalgás logikája. In Pléh Csaba – Síklaki István – Terestyéni Tamás [szerk.]: *Nyelv – kommunikáció – cselekvés.* Osiris Kiadó, Budapest, 213–228.)

Heldner, Mattias – Edlund, Jens 2010. Pauses, gaps and overlaps in conversations. *Journal of Phonetics* 38. 555–568.

Iványi Zsuzsanna 2001. A nyelvészeti konverzációelemzés. *Magyar Nyelvőr* 125. 74–93.

Koós Ildikó 2010. A szemkontaktus-kezdeményezés és szemkontaktusbontás az anya-csecsemő kommunikációs interakcióban. *Beszédkutatás 2010.* 243–253.

Liu, Huei Mei – Tsao. Feng Ming – Kuhl, Patricia K. 2003. Speech input to infants. The acoustic-phonetic characteristics of infant-directed speech in Mandarin Chinese. In Solé, Maria Josep – Recasens, Daniel – Romero, Joachim (eds.): *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona 3–9 August 2003.* Barcelona, 901–904.

Markó Alexandra 2006. Beszélőváltás a társalgásban. Előadás. *IX. Pszicholingvisztikai és Alkalmazott Nyelvészeti Nyári Egyetem.* Balatonalmádi, 2006. május 21–24. http://fonetika.nytud.hu/letolt/ma_2.pdf (A letöltés ideje: 2011. szeptember 1.)

Piaget, Jean 1993. *Az értelem pszichológiája.* Gondolat Kiadó, Budapest.

Sacks, Harvey – Schegloff, Emanuel A. – Jefferson, Gail 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50/4. 696–735.

Sundberg, Ulla 1999. Quantity in infant-directed speech. In: *Proceedings of the International Congress of the Phonetic Sciences '99.* 2189–2191.

A BESZÉD TEMPORÁLIS JELLEMZŐI 60 ÉVVEL EZELŐTTI GYERMEK BESZÉLŐKNÉL

Menyhárt Krisztina

Bevezetés

A magyar gyermekek nyelvi fejlődése és a rájuk jellemző nyelvhasználat kutatása már több mint egy évszázados múltra tekint vissza (vö. Gósy 2005). A vizsgálatok feltérképezték a gyermekek tipikus beszédprodukciós és beszédpercepciós fejlődésének menetét, számos adatot szolgáltattak a szókincs és a grammatika elsajátításának üteméről. Az anyanyelv-elsajátítás a születéstől 6-7 éves korig több szakaszon megy keresztül, a csecsemő kifejező sírásától és gőgicselésétől egészen az iskolai oktatáshoz szükséges magasabb szintű beszédkészségek elsajátításáig. Hat-hét éves korban az olvasás és írás oktatása erősíti a beszédfeldolgozás egyes szintjeit, növeli a szókincset, 8 és 10 éves kor között erőteljesen fejlődik a nyelvi tudatosság. A 10–14 évesek beszéde és nyelvhasználata egyre jobban igazodik a tanulmányi követelményekhez, míg az ifjúság (15–18 évesek) nyelve, számos sajátossága mellett (diáknyelv, nemzedékenként változó ifjúsági szleng) temporális és szerkezeti jellegét tekintve közeledik a felnőttek beszédéhez (Gósy 2005).

Hogyan alakul a gyermekek beszédének temporális szerkezete az életkor növekedésével? Az eddigi kutatási eredményekből tudjuk, hogy a kisgyermek beszédtempója jóval lassúbb, mint a pubertáskorúaké vagy a felnőtteké – pl. Laczkó (1991) a négyéveseknél 6,4 hang/s, míg a 14 éveseknél 10,4 hang/s beszédtempót mért. Szűkebb korintervallumban kisebb különbségeket találunk: a 9-10 évesek beszédtempójának átlaga egy korábbi tanulmányunkban (Menyhárt 2002) 7,2 hang/s, a 12-13 éveseké pedig 8,3 hang/s volt (irányított spontán beszéd). Az ennél idősebb fiatalok spontán beszédének temporális sajátosságait Laczkó (2009) írja le, aki a 15 éveseknél átlagosan 8,3 hang/s, míg a 18 éveseknél 10,2 hang/s beszédtempót regisztrált. A 18 évesek spontán beszédét több szempontból is elemezte Vallent (2011), náluk az átlagos beszédtempó 124 szó/perc volt. A pubertáskorúak nagymértékű beszédtempó-gyorsulását egy brazil kutatás is megerősítette, ahol a 12–14 és a 15–17 évesek között szignifikáns tempóbeli különbséget találtak (Furquin de Andrade–Martins 2007). Bár a beszéd megakadásjelenségei nem tartoznak szorosan a temporális szerkezethez, nagyon sok mindent árulnak el a beszédtervezés és -kivitelezés folyamatáról, ami végső soron hatással van a tempóra. A magyar gyermekek megakadásairól is számos ismerettel rendelkezünk. A 6-7 éves óvodások megakadásait Horváth (2006), a 10 évesekét Menyhárt

(2003) írta le, eltérő műfajú szövegeket (mese és narratíva, az adatközlők 8 évesek) vizsgált Szabó (2008), Neuberger (2011) pedig a 6, 9 és 13 évesek megakadásait és beszédük szerkesztettségét elemezte. Az adatok nagyrészt városi gyermekektől származnak, keveset tudunk a zártabb falusi közösségekben szocializálódó gyermekek és fiatalok beszélt nyelvi sajátosságairól. Az életmódbeli eltérések alapján felmerül a kérdés, vajon van-e valamilyen különbség a városi és falusi gyermekek beszédének temporális jellemzői és megakadásai között.

A hagyományos paraszti társadalomban a gyermekek szocializációja, nevelése szigorú szabályok szerint zajlott, a zárt falusi közösségben a hétköznapi ismeretek mellett a gyermek a jellemző nyelvjárást, a tájnyelvi szókészletet és a sajátos beszédkultúrát is elsajátította. Gyakori volt a hierarchikusan szerveződő korcsoport szerinti elkülönülés, amelynek alsó rétegében a gyerekkorosztálynak megfelelő korcsoportokat találjuk (gyermekbanda, gyerekhad). A következő a serdülők korcsoportja, majd a legények és lányok, felettük pedig a felnőttek és öregek korcsoportjai álltak. Gyermeknek a születéstől kb. 12 éves korig tekintették az egyént, nevelése 6 éves korig a családban történt, majd ezután a feladatok megoszlottak az iskola, a család, a munka és a gyakran felnőtt mintára szervezett baráti társaság között (Ortutay 1977–1982; Németh 1966). Ez a viszonylag zárt és merev társadalmi struktúra a 20. században folyamatosan változott, majd az 1950-es évek társadalmi változásaival lassan felbomlott. Az életmódváltozás nyelvi vetületeit Hegedűs Lajos is regisztrálta, amikor megfigyelte a „városi" szavak megjelenését a Magyarországra áttelepített moldvai csángóknál (Hegedűs 1952).

A jelen tanulmány az 1940 és 1957 között készült Hegedűs-archívum (Gósy et al. 2011) gyermek beszélőinek hangfelvételeit dolgozza fel. Miután ebből a korból, ismereteink szerint, alig maradtak gyermeknyelvi felvételek, a hangarchívum ezen a területen is hiánypótló. Emellett a gyermekek ott található hanganyaga azért is különösen értékes, mert általában a legfiatalabb generáció beszédében jelennek meg elsőként a társadalmi változások nyelvi vetületei (pl. új szavak, kifejezések, gyorsabb beszédsebesség stb.).

A kutatás célja a 60-70 évvel ezelőtti gyermekek különböző műfajú szövegeinek temporális vizsgálata (többek között az artikulációs és beszédtempó, szünetek időtartama stb.), illetve a beszédükben megjelenő megakadások tanulmányozása. Arra a kérdésre keressük a választ, hogyan tükröződnek a régi idők gyermekeinek beszédében az életkori sajátosságok, vannak-e jellemző eltérések a mesék és a narratívák között, illetve kimutatható-e valamilyen társadalmi változás lenyomata az ötvenes évek falusi ifjúságának beszédstílusában. Hipotézisünk szerint az életkori és a műfaji sajátosságok objektív adatokkal is dokumentálhatók a 60 évvel ezelőtti gyermek beszélőknél.

Beszélők, anyag és módszer

A Hegedűs-archívumból (vö. bővebben Horváth 2008; Gósy et al. 2011) összesen 14 gyermek és fiatal beszélő hanganyagát választottuk ki és dolgoztuk fel (1. táblázat). A hanganyagok összes időtartama 42 perc, a legrövidebb minta 1'41", a leghosszabb 4'31" volt. A hangfelvételek 1950 és 1956 között készültek Baranya, Hajdú-Bihar, Nógrád és Szabolcs-Szatmár-Bereg megyében. Az adatközlők kiválasztásának szempontjai nem ismertek, de több gyermek az áttelepült moldvai csángók és bukovinai székelyek leszármazottja. Bár a felvételek az ország négy egymástól távol eső vidékén készültek, ez az összevetést nem befolyásolja, ugyanis a temporális jellemzők függetlenek a nyelvjárástól, másrészről a beszélők társadalmi helyzete és életmódja nagyon hasonló.

1. táblázat: A beszélők adatai

Gyermek	Életkor	Megye	Iskolázottság	Téma
leány1	10 év	Baranya	4. osztályos	mese
leány2	11 év	Baranya	5. osztályos	mese
leány3	11 év	Baranya	5. osztályos	mese
leány4	12 év	Szabolcs	6. osztályos	narratíva
fiú1	12 év	Baranya	6. osztályos	mese
fiú2	12 év	Baranya	6. osztályos	mese
fiú3	12 év	Szabolcs	6. osztályos	narratíva
fiú4	14 év	Baranya	nincs adat	mese
fiú5	14 év	Baranya	nincs adat	mese
fiú6	14 év	Baranya	nincs adat	mese
fiú7	14 év	Hajdú-Bihar	nincs adat	narratíva
leány5	14 év	Baranya	nincs adat	narratíva
leány6	16 év	Baranya	nincs adat	narratíva
leány7	16 év	Nógrád	nincs adat	narratíva

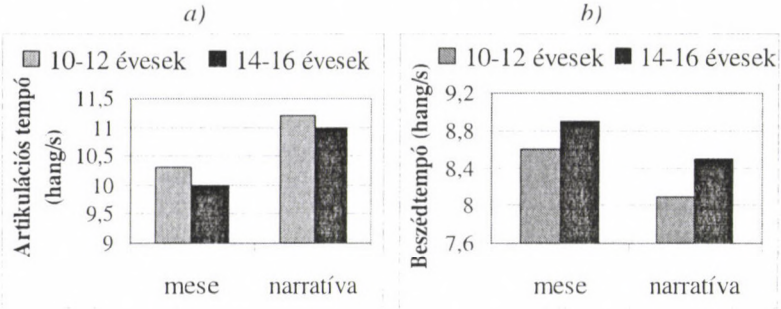
A beszélőket életkor szerint két csoportra osztottuk: 10–12 és 14–16 évesekre, mindkettőbe 7-7 gyermek került. Az adatközlők fele leány, fele fiú volt. A velük készített interjúk során 8 esetben népmesét mesélnek (pl. *Mirkó királyfi, A telhetetlen kecske, Az ördög három aranyhajszála* stb.), míg 6 esetben spontán elbeszélés hangzik el (iskolai kirándulásról, kendermunkáról, árvízről beszélnek). A 10–12 éves korcsoport hangfelvételei öt mesét és két beszélgetést, míg az idősebb beszélőkéi három mesét és négy elbeszélést tartalmaznak.

A hangfelvételeket hangzáshűen lejegyeztük, majd folyamatos akusztikai és vizuális ellenőrzés mellett Praat 5.2.30 (Boersma–Weenink 2011) program segítségével dolgoztuk fel. Adatoltuk a szünetek és a beszédszakaszok időtartamát, megállapítottuk az artikulációs és a beszédtempó értékeit. Vizsgáltuk a szünetek megjelenési helyét, típusát és átlagos időtartamát, illetve a szünetek

arányát a beszédjelhez viszonyítva. Mennyiségi és minőségi szempontból feldolgoztuk a beszédben előforduló megakadásokat is. A kapott adatokat leíró statisztikákkal és az egytényezős varianciaanalízis eredményeivel (Excel ANOVA, 95%-os megbízhatósági szinten) mutatjuk be.

Eredmények

Miután a gyermekek által elmondott szövegek két eltérő műfajt (mese és narratíva) képviseltek, ezért a vizsgálat eredményeit is ebben a bontásban adjuk meg. Az adatok műfaji és életkori eltéréseket mutatnak ugyan, ezek azonban nem jelentősek, illetve statisztikailag nem szignifikánsak. Az artikulációs tempó [1. *a)* ábra] estében a spontán narratívák (1 hang/s-mal), illetve a 10–12 éves gyermekek bizonyultak gyorsabbnak, itt a különbség elenyésző (0,2–0,3 hang/s). A beszédtempónál [1. *b)* ábra] a tendencia megfordul, a mesék, illetve az idősebb gyermekek tempója volt gyorsabb, szintén minimális mértékben (0,4-0,5 hang/s-mal). A mesélés stiláris jellegénél fogva lassúbb artikulációs tempót eredményezhet, mint a spontán elbeszélés, a már ismert szöveg előhívása az emlékezetből viszont automatikusabban történik, így a mesékben kevesebb a szünetidő, és ezért gyorsul a beszédtempó. Az idősebb gyermekek összességében kevesebb szünetet tartottak, a két csoport beszédtempója között a különbség azonban csak tendenciaszinten figyelhető meg.

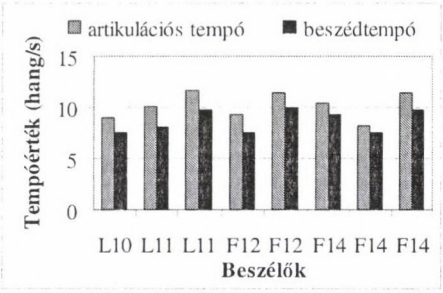


1. ábra

Az *a)* artikulációs és a *b)* beszédtempó értékei műfaji és életkori bontásban

Miután a két műfaj nem azonos arányban szerepelt a két korcsoport beszédprodukciójában, árnyaltabb képet adhat a beszéd temporális szerkezetéről az egyéni különbségek elemzése. A 2. ábra a mesék egyéni artikulációs és beszédtempóját mutatja, a beszélők nemének (leány = L, fiú = F) és életkorának jelölésével. Mesét összesen 8 gyermek mondott, a leglassúbb artikulációs tempó 8,2 hang/s volt (F14/2-vel jelölt beszélő), míg a leggyorsabb 11,7 hang/s (L11/2). Ez a 3,5 hang/s-os különbség már jelentősebb, a konkrét esetben jelezheti az artikulációs mozgások egyéni eltéréseit, illetve a szöveg

ismertségére is utalhat. A beszédtempó esetében a legalacsonyabb érték 7,6 hang/s volt, amit három gyermeknél is mértünk (L10, F12/2, illetve F14/2), a legmagasabb pedig 9,8 hang/s (L12/2 és F14/3). Itt kisebb az eltérés a leggyorsabb és a leglassúbb beszélők között (2,2 hang/s).

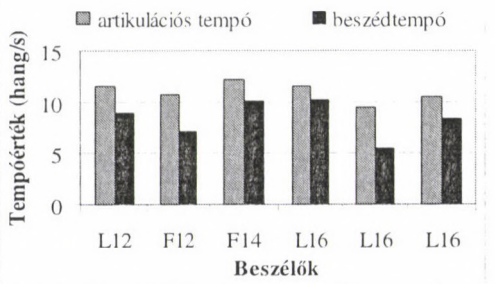


2. ábra
A mesék egyéni artikulációs és beszédtempója

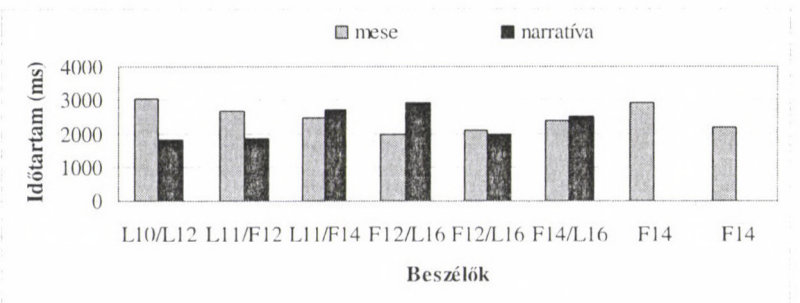
A narratívák tempóadatait (6 beszélő) a 3. ábra mutatja. A leglassúbb artikulációs tempó 9,5 hang/s volt (L16/2), míg a leggyorsabbat – 12,2 hang/s – az F14-gyel jelölt gyermeknél mértük. A két tempóadat között az eltérés valamivel kisebb (2,7 hang/s), mint a mesék esetében. A narratívák beszédtempójánál egy feltűnően alacsony értéket is találtunk: 5,5 hang/s-ot, ami ugyancsak az L16/2-vel jelölt beszélőre volt jellemző, vagyis ő volt a mintában a leglassúbb. Az ő beszéde az eleve lassú artikuláció mellett igen hosszú szüneteket is tartalmazott, aminek okát egyéni sajátosságokban kereshetjük (lehet, hogy amúgy is lassabb beszédtempóját zavarhatta a hangfelvétel ténye, illetve a sajátos téma – a kender megmunkálása – is befolyásolta). Tény, hogy ez a beszédtempóérték a legalacsonyabb az egész mintában, műfajtól is függetlenül. A leggyorsabb beszédtempót (10,13, illetve 10,14 hang/s) az F14-es és az L16/1-es beszélőnél mértük.

A szöveg tagolásáról, illetve a beszédtervezési és kivitelezési folyamatok működéséről is tájékoztatnak a beszédszakaszok időtartamadatai (4. ábra). Bár az ábra adatai jelentősebb egyéni eltéréseket mutatnak, a mesék és a narratívák beszédszakaszai átlagban csak kismértékben különböztek egymástól. A mesék átlagos beszédszakaszideje 2473 ms, a narratíváké pedig 2295 ms volt, az eltérés nem szignifikáns. A legrövidebb beszédszakaszátlag a meséknél 1999 ms (F12/1), a narratíváknál 1825 ms (F12/3), a leghosszabb pedig 3046 ms (L10) és 2925 ms volt (L16/1). A szövegekben mért legrövidebb beszédszakasz 283 ms, a leghosszabb 8290 ms, mindkettő a meséknél fordult elő, vagyis az az idő, amit a beszélő szünet nélkül beszél, igen tág intervallumok között mozog, ami függhet magától a szövegtől (pl. a mese történései-

től vagy a párbeszéd és az elbeszélés váltakozásától), annak ismertségétől, illetve a beszédtapasztalattól is.



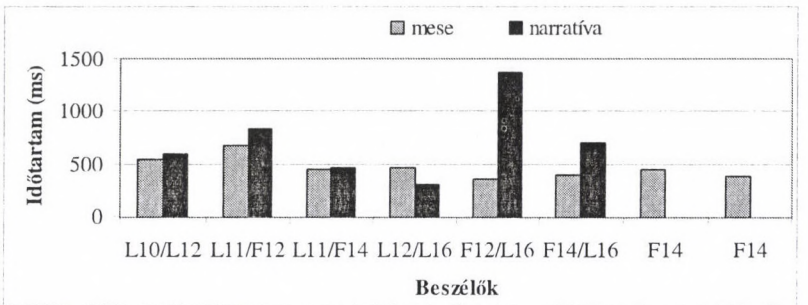
3. ábra
A narratívák egyéni artikulációs és beszédtempója



4. ábra
A beszédszakaszok egyéni értékei műfaji bontásban

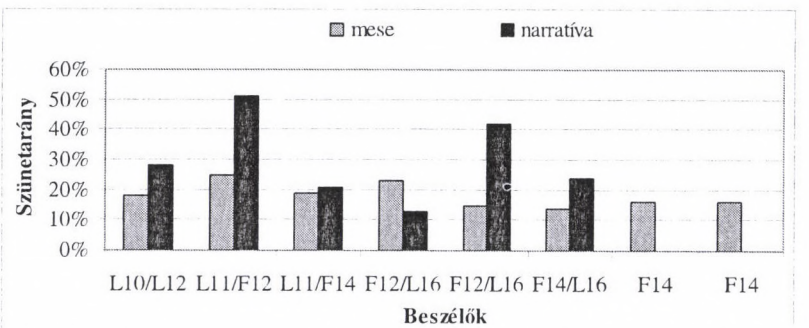
Elemeztük a néma szünetek jellemzőit, az átlagos időtartamot, arányukat a beszédjelhez viszonyítva, illetve helyüket a szövegen belül. A szövegekben előforduló kitöltött szünetekkel (hezitálásokkal) a megakadásjelenségeknél foglalkozunk. A néma szünetek átlagos időtartama a meséknél 472 ms, a narratíváknál 714 ms volt, vagyis ennél a paraméternél mértük a legjelentősebb műfaji különbséget, amely azonban nem bizonyult szignifikánsnak. A szünetek egyéni átlagai is jelentősebb eltéréseket mutattak a beszélők között (5. ábra), a meséknél a legrövidebbeket az F12/2-es számmal jelölt gyermek (371 ms), a leghosszabbakat pedig az első 11 éves kislány tartotta (675 ms). A narratíváknál a szünetátlagok sokkal tágabb határok között mozogtak, a legrövidebbek az L16/1-es számú beszélőnél voltak (310 ms), a leghosszabbakat (átlagosan 1367 ms) a minta leglassúbb beszélőjénél (L16/2) regisztrál-

tuk. A narratíváknál nagyobb különbségek figyelhetők meg az egyes beszélők között, ezt a témák eltérő volta is okozhatja.



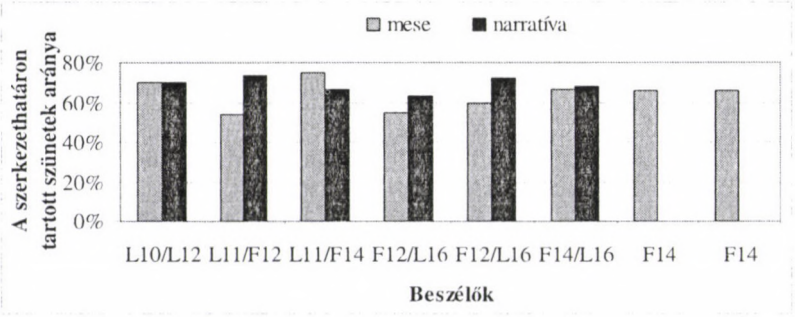
5. ábra
A szünetek átlagai

A szünetek arányát a beszédjelhez viszonyítva a 6. ábra mutatja, újra műfaji és gyermekenkénti bontásban. A meséknél a szünetek aránya átlagosan 21%, míg a narratíváknál 30% volt. Az eltérés nem szignifikáns. Az egyéni mutatók megoszlása a szünetaránynál is hasonlóan alakult, mint a szünetátlagoknál – a meséknél az egyéni értékek sokkal szűkebb határok között mozogtak, mint a narratíváknál. A meséknél a legkevesebb szünetet a beszédhez viszonyítva – 14%-ot – az F14/1-gyel jelölt gyermek tartotta, míg a legtöbbet (25%) az L11/1-es leány. A narratíváknál a legkevesebb szünetidő 13% volt (L16/1), míg a legtöbb 51% (F12/1).



6. ábra
A szünetek aránya a beszédjelhez viszonyítva műfaji és egyéni bontásban

A szünetek egyik funkciója a szöveg tagolása, éppen ezért fontos megvizsgálni, hol helyezkednek el a jelkimaradások a szövegen belül (7. ábra). A szerkezethatáron (tagoló funkcióban) tartott szünetek aránya a meséknél 64%, míg a narratíváknál 69% volt, vagyis az elbeszélések szövegét igen kis mértékben jobban tagoltnak tekinthetjük. A különbség ebben az esetben sem szignifikáns.



7. ábra
A szerkezethatáron tartott szünetek egyéni értékei

Az egyéni értékek ennél a mutatónál a narratívák esetében sokkal szűkebb határok között mozogtak, mint a korábban tárgyaltaknál: 63% (L16/1) és 72% (F12/1) között. A meséknél a két szélső érték távolabb esett egymástól – 54% (L11/1) és 75% (L11/2); érdekesség, hogy a két szélsőséget a két 11 éves leánygyermek képviseli. A mesék tagolását valószínűleg a szöveg ismeretének mértéke jobban befolyásolja, a kevéssé ismert történet összerendezése és előadása nagyobb erőfeszítést igényel, így a beszéd kivitelezésénél a tagolásra a gyermek csak kevésbé tud figyelni.

A beszéd temporális jellemzői mellett vizsgáltuk a megakadásjelenségek típusát és arányát is. Bár a megakadások, vagyis az olyan jelenségek, amelyek valamilyen módon megakasztják a beszéd folyamatosságát (vö. Gósy 2004), nem tartoznak szorosan a beszéd időviszonyai közé, fontos információkat tartalmaznak a beszédtervezés és -kivitelezés folyamatáról, ami végső soron a beszéd időszerkezetét is befolyásolja. A vizsgált hanganyagban összesen 9 megakadástípust regisztráltunk (2. táblázat), amelyek eltérő arányban fordultak elő a két műfajban, ennek oka főként a minták eltérő időtartamában keresendő (a narratívák jóval rövidebbek voltak). A leggyakoribb megakadás a hangok nyújtása volt, ezt követte a téves indítás és javítás. Viszonylag sokszor fordult elő mindkét műfajban a töltelékszó, illetve a meséknél az ismétlés (a mese szüzséjéhez tartozó, a szöveg szerves részét képező ismétléseket nem tekintettük megakadásoknak). A mesék esetében 4%-ban

hezitálások is előfordultak, míg a narratívákban 10%-ban szóban tartott szünetek is megjelentek. A meséknél a megakadások 65%-a bizonytalansági jelenség volt, ez az arány a narratíváknál 69%. A leggyakoribb hiba típusú megakadás a téves kezdés és javítás volt, amely mindkét csoportnál kitette az összes vizsgált elem egyharmadát, de mellette találtunk még grammatikai hibákat, illetve téves szótalálást is.

2. táblázat: A megakadásjelenségek előfordulásának aránya

Megakadásjelenség	Mese	Narratíva
nyújtás	33%	41%
téves kezdés, javítás	30%	31%
ismétlés	13%	3%
töltelékszó	12%	10%
hezitáció	4%	0%
téves szó	2%	0%
grammatikai hiba	3%	0%
szünet a szóban	3%	10%
újraindítás	0%	5%

Mi állapítható meg a 60 évvel ezelőtti gyermekek beszédszövegeiről? Elmondható, hogy szembetűnők voltak az egyéni különbségek mind az emlékezetből mondott mesék esetében, mind pedig a spontán elbeszéléseknél. A meséknél egyrészt érezhető a még élő mesemondó-hagyomány hatása; a gyermekek törekszenek a kifejező, élvezetes mesélésre, amelyet valószínűleg a szöveg ismertségének mértéke, illetve az előadásban szerzett eseteges gyakorlat befolyásolhat. A jól betanult mesei fordulatok előadása folyékonyan megy, minimális megakadásokkal (pl. a *Zsuzska és az ördög* című mese, 11 éves leány adatközlő):

(1) egyszer vót hol nem volt volt egyszer egy □ asszony s annak vót három lánya □ má szegény asszony nem bírta tartani a három lányt □ akkor elküldte őket hogy szerencsét próbáljanak (...) **a** □ **az** ördögnek vót három lánya □ és azt mondta az ördög **hogy ma hogy majd** □ a három lány alá tesz vánkost ezek alá meg □ köveket

(2) – hej ördög viszem ám a □ tengervágó pálcádat □ – megállj Zsuzska kutya Zsuzska **elvitted megölted két** szép lányomat most viszed a □ tengervágó pálcámat

A szöveget a párbeszédek is színesítik [vö. (2) példa], az indulatszavak használata a hallgató figyelmét hivatott lekötni. A példából azonban látszik, hogy a szünetek szerkezeten belül is megjelennek a névelők után, ami a mesélő tapasztalatlanságára utalhat. Emellett a beszédtervezés és -kivitelezés összehangolatlansága miatt a szöveg elemei össze is keverednek, amit a me-

sélő részben korrigál, az *elvitted megölted* fordulattal, viszont javítatlanul hagyja a leányok számát, holott a szöveg korábbi részéből kiderül, hogy az ördögnek is három leánya van. Találunk azonban olyan meséket is, amelyek szinte alig érthetők, hiányoznak a grammatikai alanyok, ezért igen nehéz rájönni arra, éppen ki a cselekvő személy. A következő példában hárman szerepelnek – Jankó (a főhős), az uraság és a pap (14 éves fiú adatközlő):

(3) disznóölő kést a kezébe tette □ reggel kiment a□z uraséggal a disznóknak enni adjon □ meglátta a papot a disznó hátán □ elment a □ és úgy megütte a papot □ hogy a lebukott a disznó hátáról és elment a községházához és bejelentette hogy a pap □ őnála disznót ölt □ akkor a-a □ papot eltemette...

Talán még nehezebben követhető a *Telhetetlen kecske* című mese (12 éves fiú beszélő) alábbi részlete, ahol a mesélő teljesen belezavarodik a szövegbe, a sok megakadás és szünet mellett olyan szereplőket is említ, akikről korábban szó sem volt:

(4) bement a □ be-bement a medvebarlangba aztán ott izélt bement a medvebarlangba és ott volt a nyuszi az a-a nem a nyuszi akkor a róka □ Medvebarlangba ott volt a □ az a-a r medve (...) azután meg a medve □ elment, hívta a farkast, aszongya □ gyere medve aszongya mer ott van a barlangban aszongya így □ egy egy farkas (...) aszongya ott a róka

A spontán elbeszéléseknél is találunk kisebb-nagyobb egyéni eltéréseket. A beszédszerkezet és a lényegkiemelés szempontjából igen érdekes ugyanannak az ózdi kirándulásnak a leírása, amelyet két 12 éves gyermek, egy leány (6) és egy fiú (5) mond el, kiemelve a számukra fontos mozzanatokat (a leány az elveszett ruhát, a fiú a gyárlátogatást). Az ózdi szállás mindkettőjüknek rendkívül idegen, nem is jegyezték meg, hol laktak, és ezeket a szakaszokat sok megakadással és szünettel mondják el:

(5) ott kiszálltunk az állomásban megpihentünk és jött a-a □ aki □ azzal az orvos □ jött hogy □ h-hogy menjünk a táborba □ mert még akkó' ott voltak a-a □ váliak

(6) felpakolta a holminkat □ a szekérre és elvitt ahol □ ahol lettünk □ ahova mentük

Az idősebb résztvevők spontán elbeszélései is tükrözik a régi hagyományok és az új élet párhuzamosságát – míg az egyik beszélő (14 éves leány) egy megpatkolt boszorkány történetét meséli, a másik (16 éves leány) megpróbálja összefoglalni, mit is értett a csehül beszélő filmből:

(7) Egy nénitől hallottam hogy Bukuvinában vót egy asszony aki éjjel el szokott menni nyargalni □ és vót a faluba egy szép legény □ és-s □ a-a ez az asszony elment hozzá hogy őt is megnyergelje

(8) cseh film cseh film de simán beszélt □ se írás nem vót rajta nem lehetett olvasni csak úgy az ember azér □ valahogy értette aztán

Hogyan viszonyul a régi idők gyermekeinek beszéde temporális szempontból a kortárs felnőtt beszélőkhöz, illetve a mai gyermekekéhez? A Hegedűs-archívum felnőtt beszélőinek temporális adatai korábbi kutatásainkból származnak (Menyhárt 2007; 2011). A felnőttek spontán beszédének artikulációs tempója 10,7 hang/s volt, a mesemondásé pedig 11,0 hang/s. A gyermekeknél a narratíváknál 11,1 hang/s-ot, a meséknél 10,2 hang/s-ot mértünk (összesített átlag). A beszédtempó esetében a felnőttek átlaga 8 hang/s (narratíva), illetve 8,5 hang/s (mesék) volt, míg a gyermekek esetében 8,3 és 8,8 hang/s-os értékeket találtunk. A különbségek statisztikailag nem szignifikánsak. Bár a vizsgált minta viszonylag kicsi, jól látható, hogy az archívumi beszélőknél nem figyelhető meg a beszédsebesség növekedése az életkorral. A szünetek átlagértéke a felnőtteknél a narratívák esetében 692 ms volt, a meséknél 804 ms, ez a gyermekeknél 714 ms, illetve 472 ms volt, vagyis a gyermek beszélők a mesemondáskor lényegesen rövidebb szüneteket tartottak, mint a felnőttek. A felnőtt beszélők műfajtól függetlenül szüneteik 72%-át tartották szerkezethatáron, a gyermekek ennél valamivel kevesebbet (mesék: 64%, narratíva: 69%). A szünetek aránya a felnőttek spontán beszédében 32%, a meséknél 38%, a gyermekeknél pedig 30% és 21%, vagyis a meséknél a két korosztály között jelentősebb eltérést találunk. A beszéd temporális szerkezetének elemzett adatait tekintve, úgy tűnik, hogy a régi idők beszélői a nyelvjárási különbségek és az egymástól távol eső területi elhelyezkedés ellenére viszonylag homogén közösséget alkotnak.

A megakadásjelenségek közül az archívumi felnőtteknél a leggyakrabban a töltelékszavak fordulnak elő, ezt követik a nyújtások, az ismétlések és a téves indítások (vö. Gósy–Gyarmathy 2008). A gyermekeknél is ugyanezek a leggyakoribb megakadások, azonban a gyakorisági sorrendjük eltérő (vö. 2. táblázat).

A mai gyermek beszélők adataival az összevetést csak óvatosan lehet elvégezni, hiszen a rendelkezésre álló tények zömében város gyermekektől származnak, akinek élete igen távol áll a régi idők falusi gyermekekéitől. A könnyebb összehasonlítás érdekében az archívumi gyermekek adatait is korcsoport szerinti bontásban adjuk meg. E szerint az archívumi gyermekek fiatalabb csoportjának artikulációs tempója 10,55 hang/s, beszédtempója 8,5 hang/s volt, az idősebbeké pedig 10,54 hang/s, illetve 8,7 hang/s. A mai gyermekeknél az eddigi vizsgálatok a következő eredményeket hozták: a 12-13 évesek artikulációs tempója 11,7 hang/s, beszédtempója 8,3 hang/s (Menyhárt 2002); a 15 éveseké 11,5 hang/s és 8,3 hang/s (Laczkó 2009). Ennek alapján 60 év távlatában valamiféle gyorsulás csak az artikulációs tempóban figyelhető meg, érdekes, hogy sem az archívumi, sem a mai gyermekeknél nincs eltérés a két életkori csoport között. A szünetátlag az archívumi gyermekek két csoportjánál 565 ms, illetve 586 ms volt, a mai 12 éve-

seknél ez 644 ms, a 15 éveseknél pedig 795 ms. A szünetek aránya a beszédjelhez képest a régi idők gyermekeinél 26,0%, illetve 21,0%, míg a maiaknál 31,0% és 36,5%. A mai gyermekek tehát e szerint átlagosan hosszabb szüneteket tartanak, illetve beszédük nagyobb részét teszik ki a szünetek.

A megakadások tekintetében az archívumi gyermekeknél – műfajtól függetlenül – a leggyakoribbak a nyújtások, téves kezdések, ismétlések és töltelékszavak, a mai 13 éves gyermekeknél pedig a hezitálások, a töltelékszavak, az ismétlések, illetve a téves kezdések (Neuberger 2011). A műfajok között jelentősebb különbséget mutatott ki Szabó (2008) tanulmánya, ami a jelen téma szempontjából is fontos. Nyolcéves gyermekek mesemondásában a leggyakrabban a töltelékszavak, a hezitálások, az ismétlések és a nyújtások fordultak elő, míg a spontán narratívákban az arányok kissé módosultak, és a legtöbbször hezitálások jelentek meg, ezeket követték a töltelékszavak és a nyújtások. A legnagyobb különbség a régi és a mai gyermekek között a hezitálások használatában figyelhető meg; a 60 évvel ezelőtti gyermekbeszédben hezitálások csak elenyésző mértékben fordultak elő, míg a mai gyermekeknél az összes megakadás akár egyharmadát is kitehetik (pl. Szabó 2008).

Következtetések

A tanulmány célkitűzése a 60-70 évvel ezelőtti gyermekek beszédsajátosságainak feltárása volt, ezen belül vizsgáltuk a beszéd temporális jellemzőit és a megakadásjelenségeket műfaji bontásban. A kapott adatokat emellett összevetettük a rendelkezésre álló kortárs felnőttekéivel, illetve a szakirodalomból ismert, a mai fiatalokra vonatkozó adatokkal is. A kiinduló hipotézis életkori, illetve műfaji eltéréseket feltételezett a beszélők között. A kapott adatok azonban ezt a feltételezést nem erősítették meg. A mesék artikulációs tempója lassúbb volt ugyan a narratívákénál, de a beszédtempónál ennek fordítottját figyeltük meg, és az eltérés minimális volt. A két vizsgált életkori csoport között sem lehetett tempóbeli különbséget igazolni. A szünetek esetében a narratívákat hosszabb átlagok és a megjelenésük nagyobb aránya jellemezte, de ez az eltérés sem volt szignifikáns. A megakadások esetében a két műfajnál hasonló tendencia érvényesült. Emellett a Hegedűs-archívumban talált gyermekfelvételek tempója és többi vizsgált paramétere nem különbözött jelentősen a kortárs felnőttek, illetve a mai gyermekek adataitól sem.

Az eredmények alapján két kérdéskör merül fel. Az első, hogy miért nem találunk a beszéd temporális jellemzőinél különbséget a 60 évvel ezelőtti fiatalabb és idősebb gyermekek, illetve a felnőttek között. A fiatal beszélők hangfelvételeinek relatíve kis száma miatt csak feltételezni lehet az okokat, amelyeket valószínűleg a hagyományos falusi társadalom sajátosságai is befolyásoltak. A zárt közösségek tagjainak viselkedését szigorú szabályok korlátozták és befolyásolták. Feltételezhető, hogy a megszokottól való eltérést a beszédben is gyakran megszólták, ami a fiatalokat is az íratlan szabályok követésére késztette (Paládi-Kovács 2002). A másik ok a „gyermekkornak" a

maitól eltérő értelmezése lehet. Noha a jogok tekintetében a 10–16 évesek messze nem voltak felnőttek, a mindennapi élet munka- és egyéb kötelességei rájuk is vonatkoztak. Az idősebbek a felnőttekkel egyenrangúan vettek részt a munkafolyamatokban, a nap nagy részét már a felnőttekkel töltötték. A 10–12 éveseknek elvileg az iskolában lett volna a helyük, de az 1940-es és 50-es évek néprajzi forrásai szerint dologidőben még a fiatalabb gyermekeket sem járatták iskolába (vö. Ortutay 1977–1982). Ezek a tényezők valószínűleg együtt hatottak abba az irányba, hogy a beszélők tempójának hasonlók lettek az átlagértékei.

A másik kérdéskör a régi és a mai gyermekek közötti minimális tempóeltérésre vonatkozik. Objektív adatok nélkül is egyértelmű, hogy a mai városi és a régi falusi fiatalok beszéde között óriásiak a különbségek. Eredményeink alapján ezek azonban a téma, a beszédstílus és a bizonytalanság kezelésére szolgáló stratégiák (megakadások) eltéréseiben ragadhatók meg, és nem annyira a beszéd temporális szerkezetében. Az archívumi gyermek beszélők viszonylag kis száma miatt a pontos okok nem tárhatók fel.

A Hegedűs-archívum meséi és beszélgetései egy már letűnt kor társadalmának lenyomatai, elemzésükkel nemcsak a beszélt nyelv, a beszédkultúra változásai érhetők tetten, hanem a beszélgetések témái azokat a mélyre ható társadalmi változásokat is láttatják, amelyek éppen az elmúlt század 40-es és 50-es éveitől kezdődtek. A gyermek beszélők meséi és történetei jól mutatják a hagyomány és az újítás, a régi és az új együttélését.

Irodalom

Boersma, Paul – Weenink, David 2011. *Praat 5.2.30.* www.fon.hum.uva.nl/praat

Furquim de Andrade, Claudia – Martins, Vanessa de Oliveira 2007. Fluency variation in adolescents. *Clinical Linguistics and Phonetics* 21. 771–782.

Gósy Mária 2004. A spontán magyar beszéd megakadásainak hallásalapú gyűjteménye. *Beszédkutatás 2004.* 6–19.

Gósy Mária 2005. *Pszicholingvisztika.* Osiris Kiadó, Budapest.

Gósy Mária – Gyarmathy Dorottya 2008. A nyelvhasználati változás egy jelensége. *Magyar Nyelvőr* 132. 206–222.

Gósy Mária – Horváth Viktória – Nikléczy Péter 2011. A Hegedűs-archívum mint korszerű adatbázis. In Bárth M. János – Varga Fruzsina Sára (szerk.): *Hangok – helyek.* ELTE MNyFI, Budapest, 85–103.

Hegedűs Lajos 1952. *Moldvai csángó népmesék és beszélgetések.* Közoktatásügyi Kiadóvállalat, Budapest.

Horváth Viktória 2006. A spontán beszéd és a beszédfeldolgozás összefüggései gyerekeknél. *Beszédkutatás 2006.* 134–147.

Horváth Viktória 2008. A Hegedűs-archívum (1942–1962) feldolgozásának alapelvei. In Zelliger Erzsébet (szerk.): *Nyelv, területiség, társadalom. A 14. Élőnyelvi Konferencia előadásai.* Magyar Nyelvtudományi Társaság, Budapest, 270–276.

Laczkó, Mária 1991. The interrelation of articulation rate and pauses in children's speech. In Gósy, Mária (ed.): *Temporal factors in speech*. HAS Research Institute for Linguistics, Budapest, 139–151.

Laczkó Mária 2009. Tizenévesek beszédének fonetikai és stilisztikai jellemzése. *Anyanyelv-pedagógia* 2009/1. http://www.anyanyelv-pedagogia.hu/cikkek.php?id=151

Menyhárt Krisztina 2002. A beszéd időviszonyai egynyelvű és kétnyelvű gyermekeknél. In Hunyadi László (szerk.): *Kísérleti fonetika, laboratóriumi fonológia*. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen, 55–73.

Menyhárt Krisztina 2003. A spontán beszéd megakadásjelenségei az életkor függvényében. In Hunyadi László (szerk.): *Kísérleti fonetika, laboratóriumi fonológia* a *gyakorlatban*. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen, 125–138.

Menyhárt Krisztina 2007. Rituális szövegek fonetikai jellemzői. *Beszédkutatás 2007*, 83–99.

Menyhárt Krisztina 2011. A tündérmesék prozódiája fonetikai és folklorisztikai szemszögből. In Balázs Géza – Veszelszki Ágnes (szerk.): *Nyelv és kultúra, kulturális* nyelvészet. Magyar Szemiotikai Társaság, Budapest, 274–280.

Németh Imre 1966. *A népi közösségekben élő fiatalok életmódja és kultúrája*. Arrabona, Budapest.

Neuberger Tilda 2011. Gyermekek spontán beszédének szerkesztettsége és folyamatossága. *Beszédkutatás 2011*. 83–96.

Ortutay Gyula (főszerk.) 1977–1982. *Magyar Néprajzi Lexikon*. Akadémiai Kiadó, Budapest.

Paládi-Kovács Attila (főszerk.) 2002. *Magyar Néprajz. VIII. Társadalom*. Akadémiai Kiadó, Budapest.

Szabó Kalliopé 2008. Megakadásjelenségek nyolcévesek beszédében. *Anyanyelv-pedagógia* 2008/2. http://www.anyp.hu/cikkek.php?id=56

Vallent Konstantinné 2011. A spontán beszéd sajátosságainak tükröződése középiskolások fogalmazásaiban. PhD-értekezés. ELTE, Budapest.

A kutatás a Bolyai János kutatási ösztöndíj támogatásával készült.

GYAKORISÁGI VIZSGÁLATOK SPONTÁN BESZÉDBEN

Beke András – Gósy Mária – Horváth Viktória

Bevezetés

A nyelvvel, illetve a beszéddel kapcsolatos kutatások fontos iránya a gyakorisági tényezők megismerése, adatolása, feldolgozása, elemzése. A szegmentumoknak, a szegmentumok kapcsolatainak, a szegmentumsoroknak, illetve egyes szupraszegmentumoknak az előfordulási sajátosságai, kihasználtságuk, megterheltségük egy nyelv tulajdonságainak meghatározó részét jelentik. A nyelv rendszerszerű leírásának fontos aspektusát képviselik a gyakorisági mutatók, például a különféle fonémák és a beszédhangok száma, a szótagok típusainak eloszlása, a szavak mennyisége vagy a tonális jellemzők eloszlása (pl. Rovenchak 2011). Ismereteink szerint (Laver 1995) a legkisebb számú hangkészlete a rotokasz (kelet-pápuai nyelv) és a mura (Brazília) nyelveknek van, mindössze 11 beszédhang. A legtöbb beszédhangot, 141-et a !Xu nyelvben (Angola) találjuk. A magán- és a mássalhangzók aránya is jellemez egy nyelvet; a haidában (Kanada, Alaszka) például 46 mássalhangzó mellett mindössze három magánhangzó található, míg az ausztronéz pajvan nyelvben 10 mássalhangzó mellett van 12 magánhangzó (Maddieson 1986). Az eddig elemzett nyelvekben a leggyakoribb magánhangzók az *i, e, a, o, u* fonémák[1] realizációi, ugyanakkor például a *f* mássalhangzó csak a nyelvek 40%-ában található meg. Noha ezek az adatok kiváló lehetőséget biztosítanak bizonyos összevetésekre, sőt alapot adnak univerzális megállapítások kialakításához, arról nem nyújtanak információt, hogy milyen az adott szegmentumok előfordulási gyakorisága, azaz megterheltsége. Egy nyelv szavainak, szótagjainak előfordulási gyakorisága, az őket alkotó fonémák száma, összefüggéseik, a szófajok eloszlása, egyes mondatszerkezetek előfordulása mind-mind hozzájárulnak a nyelv kvantitatív jellemzéséhez (Martindale et al. 1996).

Ezeknek és a hasonló kérdéseknek a vizsgálata szövegkorpuszok elemzését teszi szükségessé. A szövegek sajátosságai hatással vannak azok statisztikai jellemzőire. Szegmentumokra, illetve szavakra vonatkozó statisztikai elemzéseket végeztek írott szövegeken, szótári szavakon, és található néhány olyan adathalmaz, amely beszéd lejegyzett változatain alapul (pl. Weiss 1961; Sigurd 1968; Altmann 1993; Crystal 1997; Tambovtsev–Martindale 2007; Sandoval et al. 2008; Renwick 2011). A kvantitatív adatok tényei segíthetik

1. A fonémákat dőlt betűjelekkel szemléltetjük.

az univerzálékutatást éppúgy, mint például a modern beszédtechnológiai alkalmazásokat.

A nyelv statisztikai megközelítése hosszú időre vezethető vissza, az alexandriai grammatikusok vagy a bibliai konkordanciák munkálataihoz, illetve később a különböző gyakorisági szótárak készítéséhez (vö. Szende 1973). Az ilyen típusú vizsgálatok azonban sporadikusak maradtak, nem állítható, hogy a gyakorisági sajátosságok iránti érdeklődés folyamatos lett volna vagy széles köröket érintett volna. Kimondható, hogy inkább a gyakorlati szükséglet, mint a tudományos érdeklődés vezette a szakembereket, hogy nyelv- és beszédstatisztikai kérdésekkel foglalkozzanak (pl. gyorsírás, nyelvoktatás, klaviatúrák betűkiosztása, idegen nyelvek oktatása, távközlés, informatikai felhasználás, beszédtechnológia). Az első, tudományosnak tekintett nyelvstatisztikai munka a német beszédhangokról szólt több mint másfél évszázaddal ezelőtt (Förstemann 1846).

A magyar nyelvtudományban Simonyi Zsigmond nevéhez köthető az első, gyakorisággal kapcsolatos elemzés, amelyben a szerző az adatok alapján a nyelvhasználatra vonatkozó megállapítást fogalmazott meg (1896). Simonyi az *íra* alak előfordulását vizsgálta és hasonlította össze két költő (Vörösmarty és Arany) művében. A kapott adatok alapján a nyelv változására vonatkozóan tesz megállapítást, miszerint ennek a formának a használati gyakorisága csökkenni látszott. Később egyre többen kezdtek foglalkozni olyan kérdésekkel, amelyek a nyelvstatisztika területére estek (pl. Tolnai 1906; Tarnóczy 1942; Vértes 1955; G. Varga 1967; Fodor 1960; Petőfi S. 1967; Papp 1967; Kelemen 1967; Dömölki et al. 1967; Fónagy 1962; Nagy 1972; Szende 1973; Keszler 1983; Gósy 2004).

Az elmúlt évtizedben fokozottan fordultak a kutatók a nyelvi statisztikai információk irányába. Feltételezik, hogy a nyelvhasználati gyakoriság a nyelvi szerkezet elemzésének és magyarázatának a tényezője (Bybee 2006a). Az a tény például, hogy az egyén milyen grammatikai, illetve szintaktikai szerkezeteket, avagy szavakat használ, hatással van a belső nyelvi reprezentációira. Ezek a reprezentációk egyfelől a beszélőnek a nyelvre vonatkozó konvencionális ismereteit tükrözik, másfelől jelzik a nyelvi változatosságot és a nyelvváltozást magát (Bybee 2006b). A nyelvhasználati sajátosságok ilyen módon magyarázni képesek a szinkrón nyelvi változások természetét vagy az egyéni nyelvi különbségeket. A legkülönfélébb nyelvi elemek, lexikai egységek vagy szerkezetek előfordulási gyakoriságának szisztematikus elemzése nagy hatást gyakorol a diskurzuselemzésre, a korpusznyelvészetre, a fonológiára, a fonetikára, sőt a nyelvtörténetre is. Nélkülözhetetlen az alkalmazott nyelvészeti kutatások jó részében, az alkalmazott beszédkutatásokban, avagy a beszéd mesterséges felismerésében, a bűnügyi fonetikában.

A módszertani lehetőségek sokfélesége, a számítógépes elemzések és a nagy adatbázisok új lehetőséget nyújtanak a nyelvstatisztikai vizsgálatok végzéséhez. Minél nagyobb az adatmennyiség, annál relevánsabbak az ered-

mények, és ezért annál megbízhatóbban használhatók fel általánosításokra. Noha a nyelvhasználat kvantitatív jellemzőinek fontossága nem megkérdőjelezhető, a kvalitatív vizsgálatok meghatározóbbnak tűnnek mind a külföldi, mind a hazai szakirodalomban.

A spontán beszéd jellemzésében nagy szerepe van a különféle tényezők előfordulási gyakoriságának (pl. néma és kitöltött szünetek, nyújtások, ismételt szavak, redukciók, egyszerre beszélések, hangsúlyhelyek). Kutatásunk a szegmentális szintre irányult, nagy mennyiségű spontán korpuszon végeztünk statisztikai elemzéseket. Vizsgálataink célja elsősorban a fonémák, a trifónok előfordulási gyakoriságának, valamint a szóhosszúság alakulásának a megismerése volt. A szótövek, a tartalmas és a funkciószavak előfordulásáról kívántunk továbbá mennyiségi információkat szerezni. Azt feltételeztük, hogy a gyakorisági eloszlások nem fognak jelentős különbséget mutatni Szende mintegy négy évtizeddel korábbi adataihoz képest, valamint hogy a szóhossz és az előfordulás ismert összefüggése, az ún. Zipf-törvény (Zipf 1949) kimutatható lesz a spontán beszédben.

Kísérleti személyek, anyag, módszer

Kutatásunkban a BEA-adatbázis 100 beszélőjének (42 nő, 58 férfi, átlagéletkoruk 41 év) spontán narratíváit (interjúk és érvelések) használtuk fel a lejegyzések alapján. Közel 26 órányi spontán beszédet elemeztünk, ebben a szavak száma 151 161, a fonémák száma 748 099 volt. A lejegyzéseket a statisztikai feldolgozásokhoz előkészítettük, a *ly* betűt a megfelelő *j* fonémára változtattuk, a nonverbális jelekre vonatkozó utalásokat (pl. nevetés, krákogás) a szövegekből töröltük. Az így kialakított spontánbeszéd-korpusz lehetővé tette a gyakorisági elemzések statisztikai alapú elvégzését. Az elemzések során figyelembe vettük a helyesírási szabályokból adódó betű–fonéma megfeleléseket. Mindenekelőtt tisztáznunk kell, hogy mit is értünk fonémán, vagyis mik azok a szegmentumok, amiket számolni fogunk. A generatív fonémafogalom ('mögöttes szegmentum') értelmében (vö. (Siptár–Törkenczy 2000) nincsen hosszú *zzs, vv, hh* és *ddzs,* mivel ezek nem (illetve igen ritkán, pl. *pech, hodzsa*) fordulnak elő morfémán belül: csak hasonulás révén, morfémahatáron jönnek létre (*rizzsel, szívvel, ahhoz, briddzsel*). A *dz* pedig vagy egyáltalán nem létezik, mint önálló fonéma, vagy csak hosszú van belőle (pl. *pedzi, madzag*). A kérdést egyszerűsíti, hogy az általunk vizsgált anyagban az említettek közül csupán a hosszú *vv* és *hh* mássalhangzók fordultak elő (0,2%-ban). Noha elfogadjuk a generatív felfogást, amely szerint ezek nem fonémák (hanem levezetett alakok), a statisztikában feltüntetjük ezeket a (ritka) előfordulásúakat, mint (főként) hasonulás következtében létrejött szegmentumokét (az 1. táblázatban *-gal jelölve). A célunk ezzel elsősorban az, hogy ne maradjanak ki a gyakorisági táblázatból, és így az adatok összevethetők legyenek korábbi munkákkal. Két azonos rövid fonéma egymás utáni előfordulását – például, ha közöttük szóhatár (vagy összetételhatár) volt –

annak ellenére a hosszú fonémák közé számítottuk, hogy voltaképpen két rövid mássalhangzó egymás mellé kerüléséről van szó (pl. *hattól* típusú szavak intervokális mássalhangzója)[2]. A szavak esetleges rövidebb ejtését (pl. *tát* a *tehát* helyett vagy *akko* az *akkor* helyett) nem vettük figyelembe, minden esetben a teljes fonémasort adatoltuk. Megvizsgáltuk *a)* az összes fonéma, *b)* a hármas beszédhang-kapcsolatok (trifónok) előfordulásának gyakoriságát, *c)* a szavak hosszúságát, *d)* a szótövek és a szavak előfordulási sajátosságait, valamint *e)* a tartalmas és a funkciószavak arányát. A trifónok elemzése a fonémák hangkörnyezetének leírása szempontjából volt meghatározó. Minden egyes szóban meghatároztuk a lehetséges trifónokat, függetlenül a szótaghatártól. Például a *traktor* szó esetében a trifónok a következők: *tra, rak, akt, kto, tor.* A fonémák és a trifónok gyakoriságának, valamint a szóhosszúságnak az elemzéséhez MATLAB környezetben írt saját programokat használtunk, míg a szótövezéshez a HUNMORPH szoftvert alkalmaztuk.

Eredmények

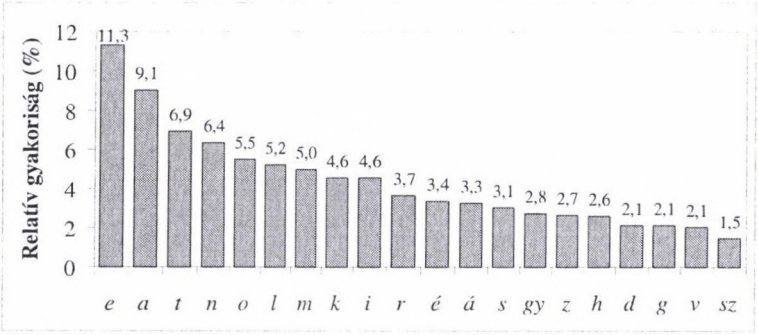
1. A nyelv szegmentumainak statisztikai leírásában az egyik legfontosabb jellemző a magánhangzók és a mássalhangzók aránya. A vizsgált magyar spontánbeszéd-korpuszban a magánhangzók és a mássalhangzók aránya (V/C) 0,683%. A magánhangzók 40,6%-ban, míg a mássalhangzók 59,4%-ban fordultak elő. Ez az eredmény közel megegyezik Szende mintegy négy évtizeddel ezelőtti közleményével (1973), ahol a V/C arány 0,727% volt. A magánhangzók előfordulása az általa elemzett anyagban 42,1%, a mássalhangzóké 57,9% volt. Noha a különbségek minimálisak, annyi megállapítható, hogy napjaink spontán beszédében a magánhangzók aránya kissé alacsonyabb, a mássalhangzóké kissé magasabb. Minthogy 14 magyar magánhangzó-fonéma és 25 mássalhangzó-fonéma van, amelyek aránya 0,56%, ezért a spontán beszédben adatolt előfordulásuk azt mutatja, hogy a magánhangzók gyakoribbak, mint a mássalhangzók.

Elemeztük az adatokat a teljes fonémaállományt tekintve (ekkor ez volt 100%), valamint a magánhangzók és a mássalhangzókat külön-külön véve is. Valamennyi adat esetében azt feltételeztük, hogy várhatóan a leggyakoribb fonéma az *e* lesz, és a gyakorisági sort a hosszú mássalhangzó-fonémák fogják zárni (1. ábra).

Az ábrán a húsz leggyakoribb fonéma (1,49%-ig) előfordulását tüntettük fel. Közülük 6 magánhangzó, ezek részesedése 42,85%, 14 mássalhangzó, a részesedés 28,57%. A leggyakrabban használt húsz fonéma az összes fonéma előfordulásának 71,42%-át adja; a magánhangzók ebben nagyobb mértékben vesznek részt. A fonetikailag is párt alkotó rövid-hosszú magánhangzó fonémák közül az első húszban csak a rövidek szerepelnek, a hosszú magánhang-

2. A szerzők megköszönik Siptár Péternek a fonémákkal kapcsolatos megjegyzéseit és tanácsait, amelyek alapvetőek voltak a statisztikai elemzések kialakításához.

zókat (a fonetikai pár nélküli) *á, é* képviseli. Teljesen hiányoznak a felső nyelvállású labiális palatális és veláris magánhangzók. A gyakori artikulációs gesztusok tehát a szájüreg középső és alsóbb részeiben jellemzők. Hosszú mássalhangzók egyáltalán nem fordulnak elő. Megjelenik csaknem az összes alveoláris helyen képzett mássalhangzó, egy-egy bilabiális és labiodentális fonéma, egy-egy posztalveoláris és palatális, a két veláris és a laringális *h*. A képzési helyet tekintve tehát a leggyakoribb mássalhangzók meglehetősen arányosan oszlanak meg, az alveolárisok fokozott előfordulása mintegy az artikuláció gazdaságosságát is mutatja. A magyar mássalhangzók képzési módját tekintve csupán az affrikáták összetett artikulációja nem fordul elő a gyakori fonémák között. Jóval nagyobb mértékű a zöngések előfordulása, ami a *m, n, l* és a pergőhang gyakoriságával magyarázható. Ha csak az obstruenseket nézzük, akkor nincs különbség az előfordulásban a zöngésség tekintetében. A húsz leggyakoribb fonéma képzési konfigurációja egyértelműen az ejtési gazdaságosságot, illetőleg a relatíve könnyebb artikulációt szemlélteti.

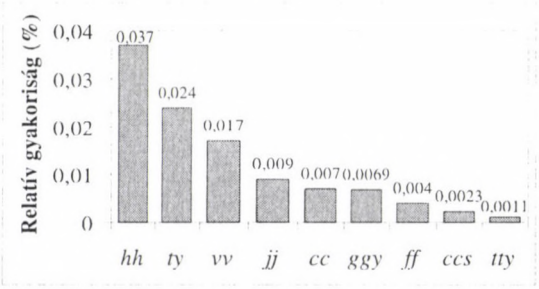


1. ábra
A leggyakoribb fonémák előfordulása a vizsgált anyagban

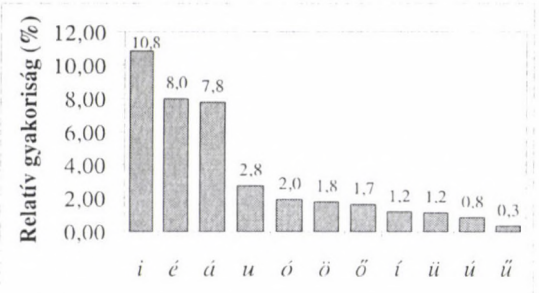
Szende négy évtizeddel ezelőtti elemzéseinek anyagát 79 399 fonéma alkotta, ez a jelen kutatásnak csupán 10,6%-a. Ha összevetjük a kétféle korpusz eredményeit, a tíz leggyakoribb fonémát tekintve alig látunk különbséget, a továbbiakban az eltérés az egyes fonémák sorrendiségét érinti (vö. 1973: 30). Úgy tűnik tehát, hogy a korpusz nagyságától, illetve az időkülönbségtől függetlenül a magyar spontán beszédben a legnagyobb mértékű megterhelése változatlanul ugyanazokat a fonémákat érinti, azonban az egyes típusok mutatnak (minimális mértékű) különbségeket.

Megnéztük, hogy a magyar spontán beszédben mely fonémák mutatják a legritkább előfordulást. Az adatok igazolták azt a feltételezésünket, hogy a legritkábbak a nyelvileg hosszú mássalhangzók lesznek. A korpuszban a magánhangzók közül legkisebb arányban a felső nyelvállású, labiális magán-

hangzók fordultak elő (3. ábra). A mássalhangzók közül hosszú, zöngétlen mássalhangzók (*ff, ccs, tty*) zárják a gyakorisági sort (2. ábra). A legritkábbak között az egyedüli rövid fonéma a *ty*. Szende adatai csak a mássalhangzók tekintetében mutatnak némi eltérést. Sorrendjében az utolsó helyeken a *hh, jj, ggy* mássalhangzókat tünteti föl (1973: 30, 31).



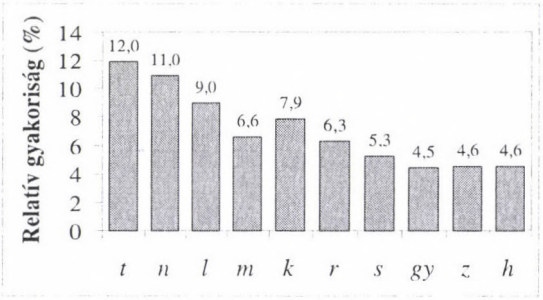
2. ábra
A tíz legritkább fonéma előfordulása a vizsgált anyagban



3. ábra
A magánhangzó-fonémák előfordulása a vizsgált anyagban, az összes magánhangzó 100% (a fonémák betűjelekkel szerepelnek)

A magánhangzókon és a mássalhangzókon belül külön-külön is elemeztük az egyes fonémák előfordulását. Az összes magánhangzót véve 100%-nak, az eredmények azt mutatják, hogy az *e* 27%-ban fordult elő, a második leggyakoribb, az *a* pedig 21,5%-ban (3. ábra). 10% és 15% közötti az *o* és az *i* előfordulásának aránya, az összes többié 10% alatti. A rövid magánhangzófonémák aránya anyagunkban 78,18%, a hosszúaké 21,82%, ezeknek az aránya csaknem megegyezik Szende adataival (78,49%, ill. 21,51%). A rövid magánhangzók megterheltsége jóval nagyobb, mint a hosszúaké, és ez a tendencia nem változott az elmúlt négy évtizedben.

A mássalhangzók esetében a leggyakoribb a *t* volt 13,44%-kal, majd az alveoláris nazális (*n*) következett 11,11%-kal (4. ábra); további 5 mássalhangzó fonéma pedig 5%-nál gyakrabban fordult elő anyagunkban. A rövid mássalhangzó-fonémák aránya anyagunkban 95,57%, a hosszúaké 4,43% volt, Szende adatai hasonlók (96,73%, ill. 3,27%). A rövid mássalhangzók megterheltsége jóval nagyobb, mint a hosszúaké. Az 1. táblázat valamennyi mássalhangzó-fonéma gyakorisági adatait tartalmazza.



4. ábra
A 10 leggyakoribb mássalhangzó-fonéma előfordulása a vizsgált anyagban (az összes mássalhangzó 100%)

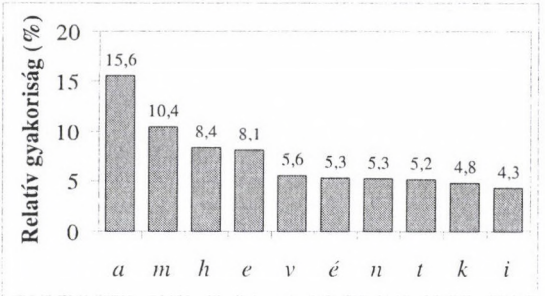
2. Egy nyelv leírásában, különösen a fonotaktikai sajátosságok, illetve szabályok megállapításához fontos jellemző, hogy a spontán beszédben előforduló szavak milyen fonémával kezdődnek és végződnek. Ezek az adatok szorosan összefüggnek a szókészleten túl a nyelv morfológiai sajátosságaival, a gyakori morfémákat felépítő fonémasorokkal és bizonyos funkciószók használati gyakoriságával. A leggyakoribb magyar szókezdő fonéma (valamennyi szókezdőt tekintve 100%-nak) az *a* (15,57%), ami a határozott névelők, a mutató és a vonatkozó névmások fonémasorával, illetve ezeknek a szavaknak a funkcionális sajátosságaival is magyarázható. Ezt követi a *m* (10,42%), amely részben a szókészlet sajátosságaiból, részben adódik a *meg* igekötő előfordulásából, valamint a *minden, még, már* szavak gyakoriságából a spontán közlésekben (5. ábra). A *h* relatív gyakorisága a *hát* és a *hogy* előfordulási sajátosságaival magyarázható. Az *e* gyakorisága főként a határozatlan névelővel, a szókezdő *v*-é pedig a létige különböző formáival hozható kapcsolatba. A szókezdő *é* helye a sorban főként az egyes szám első személyű személyes névmás megjelenését tükrözi.

Szó végén rendszerint azok a fonémák fordulnak elő gyakrabban, amelyek a toldalékokban vagy – egyéb okoknál fogva – gyakori szavakban találhatók (6. ábra). Ilyen például a *t* fonéma, szóvégi gyakorisága a múlt idő és a tárgyrag használatából adódik. Az *a* magánhangzó nagyarányú szóvégi megjelené-

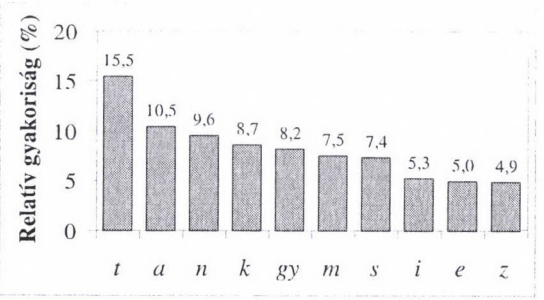
se magyarázható egyes toldalékok, elsősorban a *-ba* és a *-ra* előfordulásával, az egyes szám harmadik személyű igei személyjellel (pl. *adja, látja, mondja*), a birtokjellel (*labdája, szabadsága, kutyája*), valamint a *ha, hogyha* kötőszavakkal. Ugyancsak egyes ragok (*-ban, -ben, -ön, -en, -ön*) következtében (is) kerül a harmadik helyre az alveoláris nazális mássalhangzó, ehhez a létige egyes szám harmadik személyű formája is hozzájárul. A gyakoriságát szó végén mindenekelőtt a többes számú alakok idézik elő, míg a *gy* ötödik helyére a rangsorban főként a *hogy* kötőszó és az *egy* gyakorisága a valószínű magyarázat.

1. táblázat: Az összes mássalhangzó-fonéma előfordulásának adatai

Sorrend	A fonéma betűjele	Relatív gyakoriság (%)	Sorrend	A fonéma betűjele	Relatív gyakoriság (%)
1.	t	11,97	24.	bb	0,5200
2.	n	10,98	25.	nn	0,4000
3.	l	9,01	26.	c	0,3000
4.	m	8,59	27.	ssz	0,2000
5.	k	7,92	28.	nny	0,1560
6.	r	6,31	29.	rr	0,1500
7.	s	5,26	30.	ss	0,1400
8.	gy	4,75	31.	zz	0,1300
9.	z	4,56	32.	mm	0,0900
10.	h	4,54	33.	pp	0,0700
11.	d	3,69	34.	zs	0,0600
12.	g	3,68	35.	gg	0,0400
13.	v	3,56	36.	dd	0,0300
14.	sz	2,57	37.	hh	0,0200
15.	b	2,25	38.	ty	0,0200
16.	j	1,72	39.	vv	0,0100
17.	p	1,25	40.	jj	0,0097
18.	f	1,15	41.	cc	0,0078
19.	tt	1,14	42.	ggy	0,0076
20.	ny	0,72	43.	ff	0,0040
21.	cs	0,66	44.	ccs	0,0016
22.	ll	0,61	45.	tty	0,0004
23.	kk	0,56			



5. ábra
A leggyakoribb szókezdő fonémák előfordulása a vizsgált anyagban



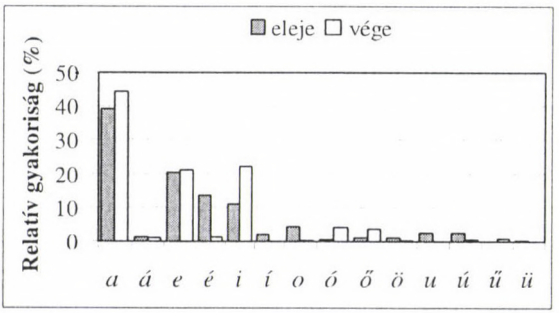
6. ábra
A szó végén előforduló leggyakoribb fonémák a vizsgált anyagban

A szó eleji és a szóvégi pozíciót összehasonlítva, azt látjuk, hogy a magánhangzók egy részénél nincs különbség az előfordulásban (7. ábra). Ilyen az *e, ö, ü* fonémák. Egy más részüknél kismértékű különbség látható. Az *i, o, u, ú* gyakoribbak a szó eleji, míg az *a, ó, ő* a szóvégi pozícióban. Jelentős különbséget adatoltunk a szóban elfoglalt helyzet függvényében az *é* és az *i* esetében; az előbbi jóval nagyobb arányban fordul elő szó elején, míg az utóbbi szó végén.

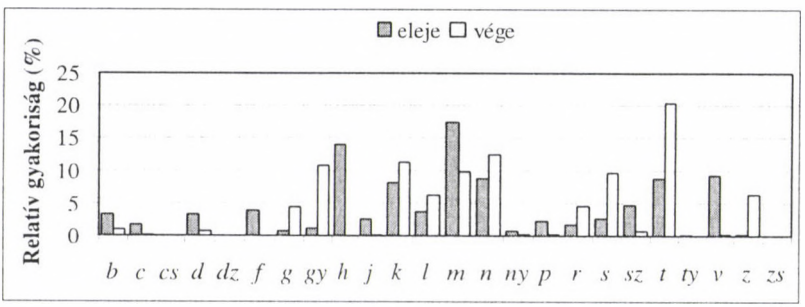
A mássalhangzókra is jellemző, hogy melyik helyzetben és milyen mértékben gyakoribbak (a hosszú mássalhangzókkal itt csekély előfordulásuk miatt nem foglalkoztunk részletesen). Az egyébként is ritka mássalhangzófonémák (pl. *cs, ty*) jellemzően nem mutatnak eltérést a szó eleji, illetve a szóvégi megjelenésben (8. ábra).

Kismértékben változó az előfordulása a pozíció függvényében 12 fonémának, közülük 7 szó eleji helyzetben gyakoribb (pl. *b, d, f, ts, s*), öt pedig szóvégiben (*g, k, n, l, r*). Hét esetben tapasztaltunk jelentős eltérést a helyzettől

függően; négy fonéma a szó végén (*t, gy, s, z*), három pedig a szó elején volt gyakoribb (*m, h, v*). A szóvégi gyakori *s* főként az *és, is* szavakkal, a *z* pedig a határozott névelővel magyarázható.



7. ábra
A magánhangzók eloszlása a szóban elfoglalt helyzet függvényében

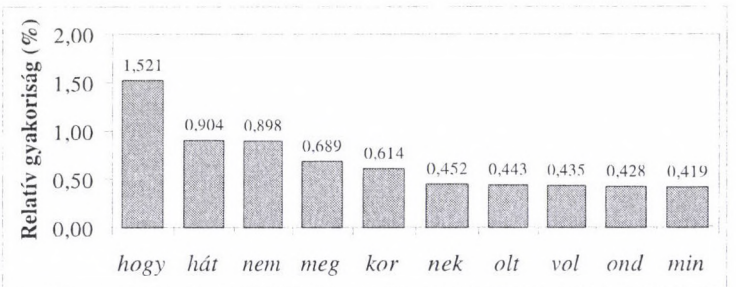


8. ábra
A mássalhangzók eloszlása a szóban elfoglalt helyzet függvényében

3. A fonémák hangkörnyezetének egyfajta leírását teszik lehetővé a hangkapcsolat-hármasok, a trifónok. A magán- és mássalhangzók 455 680 darab különböző trifónt alkottak. A fonémák esetében láttuk, hogy bizonyos szavak magyarázzák az előfordulás gyakoriságát, ez megállapítható a trifónokra is. A *hogy* gyakori előfordulása eredményezi a leggyakoribb trifónt (9. ábra).

A *hát* diskurzusjelölő, illetve töltelékszó már lényegesen ritkábban volt adatolható, aránya megegyezik a tagadószóéval. A leggyakoribb trifónok közé tartozik előre jelezhetően a *meg* (feltehetően döntően igekötőként). A *kor* trifón rangsorbeli ötödik helye elsősorban az *akkor* szóval, a *nek* pedig részben a toldalékkal, részben a *nekem* (*neki* stb.) szavakkal magyarázható. A spontán beszédre jellemzőnek tekinthető a múlt idejű létige gyakori használata, és en-

nek következtében a *vol* és az *olt* trifónok gyakorisága. A *min* valószínűsíthetően a megfelelő funkciószók előfordulási következménye. Narratívák és társalgások szavainak elemzésében azt találták, hogy a három leggyakoribb szó a *hogy*, a *van* és a *nem*, ezeket követték az *és*, az *igen* és az *én* szavak (Gósy 2012). Anyagunkban az összes trifónból 2214 darab mindössze egyszer fordult elő, ez az összes trifón 0,48%-a, például: *cku, cög, átö, éen, ejr, göm, hág, onő*.

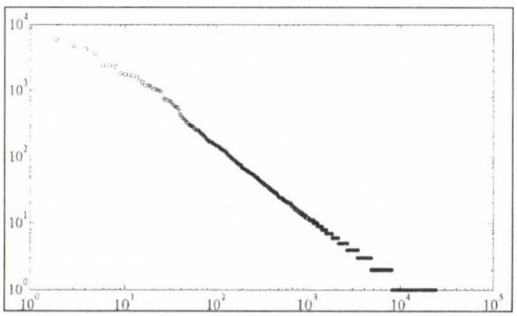


9. ábra
A trifónok gyakorisági előfordulása a vizsgált anyagban

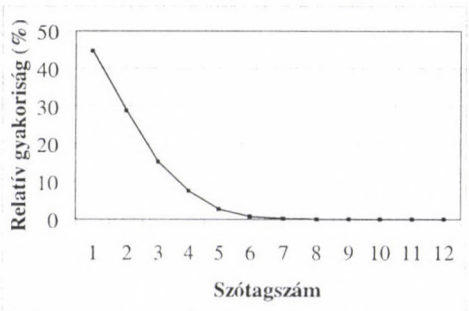
4. A szóhosszúság, vagyis a szavakat alkotó fonémák száma az egyes nyelvek egyik legfontosabb jellemzője. A portugálban például az átlagos szóhossz 4,6 fonéma (Qaresma 2008), a maninka nyelvben (Guinea) 3,02 fonéma (Rovenchak 2011), az angolban 3 körülire becsülik (Cutler–Carter 1987). A jelen korpuszban az átlagos szóhossz 4,9 fonéma volt, ami jól szemlélteti a magyar agglutináló jellegét. A szótövek átlagos hossza 4,5 fonéma. A legtöbb szó 1–14 fonémából épül fel (10. ábra). (A magyar összetett szavak helyesírásának ingadozása kismértékben, de nem számottevően befolyásolja ezeket az adatokat.)

A Zipf-törvény (1947) azt mondja ki, hogy minél gyakoribb egy szó használata, annál kevesebb fonémából építkezik. Ellenőriztük, hogy anyagunkra ez igazolható-e (vö. 10. ábra). A magyar adatok is megerősítették a törvényt, a több fonémából álló szavak használata ritkább volt, mint a kevesebb fonémát tartalmazóké. Az ábra a spontán beszédanyagban előfordult szavak számának gyakoriságát (*y* tengely) és az őket alkotó fonémák számát (*x* tengely) szemlélteti logaritmikus koordinátarendszerben. A szavakat felépítő szótagok száma is tanulságos kvantitatív mutató, amely számos ismert nyelvre igazolja azt az összefüggést, hogy minél gyakoribb egy szó, annál kevesebb szótagból áll. Szende adatai 9 szótagos szóig elemezve alátámasztják ezt (1973: 56). A leggyakoribbak az egy szótagú szavak (46,58%), a 8, 9 szótagból építkezők előfordulása 0,003%, illetve 0,002%. Saját anyagunkban 12 szótagos szavakig

vizsgáltuk a szótagok számát (11. ábra). Eredményeink nagyon hasonlók a négy évtizeddel korábbi arányokhoz. Az egy szótagból álló szavak előfordulása a legnagyobb arányú (44,69%), a két szótagosaké az előzőnek már csak alig a fele (28,83%). A három (15,18%) és négy szótagból álló szavak (7,58%) között jelentős csökkenés tapasztalható. Az öt szótagos szavak előfordulása ismét erőteljesen csökken (2,68%), az ennél több szótagot tartalmazóké pedig 1% alatti (6 szótagos = 0,709%, 7 szótagos = 0,2245%, 8 szótagos = 0,0586%, 9 szótagos = 0,0199%, 10 szótagos = 0,0046%, 11 szótagos = 0,0013%, 12 szótagos = 0,00066%). Példák: *fa, domb, anyag, dobogó, alkalmazzák, megújulóból, gorombáskodjanak, elkötelezettségem, megkérdőjelezhetetlen.*



10. ábra
A fonémák előfordulása a szavak hosszúságának függvényében a vizsgált anyagban, igazolva a Zipf-törvényt (a függőleges tengelyen az előfordulás, a vízszintes tengelyen a sorrend látható)

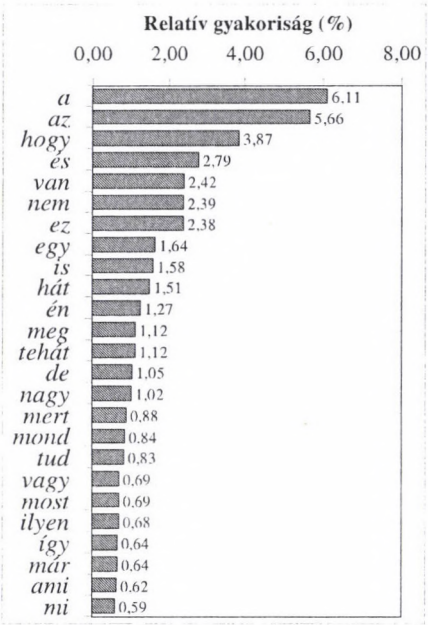


11. ábra
A szavak szótagszámának és a relatív gyakoriságnak az összefüggése

A szavak szótagjait alkotó fonémák számával kapcsolatos Menzerath (1954) hipotézise. Azt vélelmezte, hogy az átlagos szótaghosszúság csökken,

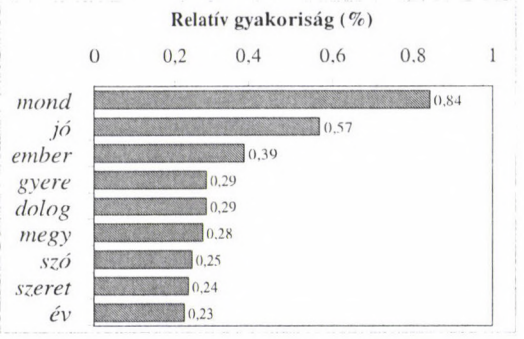
ha a szótagszám növekszik. Anyagunkban ennek ellenőrzésére az ötven leggyakoribb, minimum hat szótagból álló és az ötven legritkább 1-3 szótagból álló szón végeztünk számításokat a fonémák száma tekintetében. Az eredmények azt mutatták, hogy – bár szignifikáns különbség nem volt igazolható –, a hosszabb szavak szótagjai valóban rövidebbeknek adódtak, mint a rövidebb szavak szótagjai. Az elemzett hosszú szavak szótagjainak fonémaszáma átlagosan 2,404, míg a rövideké átlagosan 2,588 volt.

Elemezéseket végeztünk a szavak előfordulásának néhány sajátosságával kapcsolatosan. A határozott névelők megjelenése mintegy 11% (12. ábra), az adat megegyezik a négy évtizeddel ezelőttivel (Szende 1973: 54).



12. ábra
Szavak gyakorisági előfordulása a vizsgált anyagban

A gyakorisági sorrendben két kötőszó áll a névelők után. A *van*, a *nem* előfordulása gyakorlatilag megegyezik. Főnév nem található az első 25 szó között, a létigén kívül viszont további két ige is megjelenik, a *mond* és a *tud*. Külön elemeztük a szótöveket, a leggyakoribbak többsége főnév (13. ábra).



13. ábra
A kilenc leggyakoribb szótő előfordulása a vizsgált anyagban

A tartalmas szavak aránya 58,8%, a funkciószavaké 41,2%, az egymáshoz viszonyított arányuk 0,70068%. Szende elemzésében a tartalmas szavak aránya 53,62% (1973: 54), vagyis nem számottevően kevesebb, mint a jelen anyagban. A spontán angol társalgásokban a funkciószók aránya hol nagyobb (Bell et al. 2002), hol közel azonosnak adódott a tartalmas szavakéval (Cutler–Carter 1987). A különbségek számos okra vezethetők vissza, a nyelvszerkezet eltéréseitől, a beszédstíluson át a figyelembe vett szavak számáig.

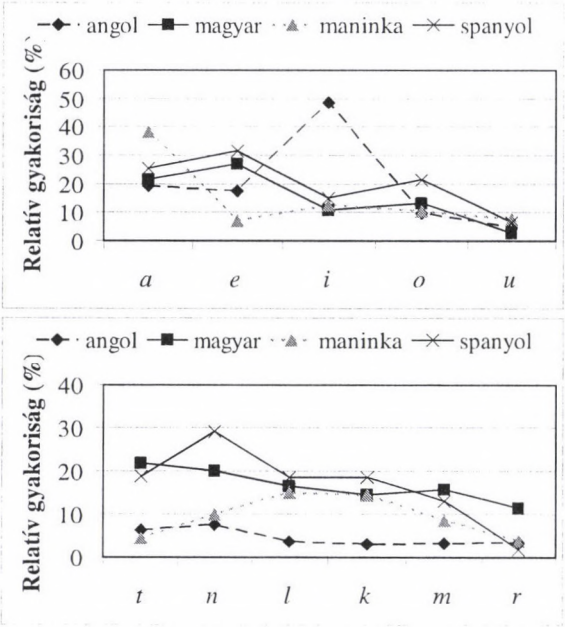
Az összes szó 18,72%-a csupán egyszer szerepelt a korpuszban; példák a tartalmas szavak közül: *akadály, agyvérzés, bérház, üvölt, tüdőgyulladás, őrjöng, szőlészet, restell, révbe, vitorlázik,* illetve a funkciószavak közül: *ama, éppúgy, valamerre, valamiképpen, akárhol, akárhány.* A korpusz nagysága, azaz a szószám (és vélhetően a téma is), meghatározó lehet az egyszeri előfordulás tekintetében. Harmincezernél kissé több szót tartalmazó korpuszban az egyszer előfordult szavak aránya narratívákban 23% körüli volt (Gósy 2012).

Következtetések

A jelen kutatás statisztikai adatai első ízben adnak pontos információt a vizsgált nyelvi jelenségek gyakorisági jellemzőiről igen nagy spontán beszédanyagon. Az adatok alapján megállapíthatjuk, hogy – legalábbis négy évtizednyi időtávban – a fonémagyakoriság nem mutat jelentős változást. Ez feltehetően annak (is) a következménye, hogy a szóhasználat változása statisztikailag, illetve a fonémasorokat illetően nem jelentős, a toldalékok, illetve a funkciószavak bizonyos használati állandósága pedig biztosítja a mérési adatok relatív változatlanságát. A fonémagyakoriságot aszimmetrikusnak tekintik abban az értelemben, hogy az új szavak általában gyakori fonémákat tartalmaznak, ezáltal ezek a fonémák még gyakoribbak lesznek (Martin 2009). (Kérdés, hogy az idegen nyelvből átvett szavak ezt a megállapítást mennyiben módosítják.)

A statisztikai adatok jellegzetesen nyelvspecifikus tendenciákat mutatnak, mivel a szavakat alkotó fonémákon túl a fonológiai szabályok (pl. magánhangzó-harmónia, fonotaktikai szabályok), valamint a morfológiai szerkezet, illetve a toldalékokat alkotó fonémák minősége nagymértékben befolyásolják a gyakorisági eloszlásokat. Az angol szótagszerkezet statisztikai megközelítése például a fonotaktikai szabályok pontosabb megfogalmazását tette lehetővé (Kessler–Treimann 1997; Cutler–Carter 987).

A kvantitatív vizsgálatok lehetőséget nyújtanak összehasonlító nyelvészeti vizsgálatokra is (vö. Martindale et al. 1996). A konkrét összevetés eredményein túl ez módot ad az univerzális és a specifikus sajátosságok megismerésére, például az egyes nyelvek hangzását vagy a fonotaktikai szabályszerűségeket tekintve. A 14. ábrán különböző nyelvcsaládokhoz tartozó nyelvek – angol (Fry 1947), spanyol (Sandoval et al. 2008), maninka (Rovenchak 2011) és magyar (saját adatok) – beszédanyagában előforduló fonémák gyakoriságát ábrázoltuk. (Az összevetés célja kizárólag a tendenciák szemléltetése, mivel a szakirodalmi adatok nem feltétlenül azonos nyelvi anyagot reprezentálnak.)



14. ábra
Négy nyelv néhány fonémájának előfordulása

Feltételezhető, hogy azok a fonémák gyakoriak egy-egy nyelvben, amelyek a (mára már elemzett) nyelvek többségében is nagy arányban fordulnak elő. A különféle képzéshelyű, zöngétlen felpattanó zárhangok igen nagy arányban képviseltetik magukat a világ elemzett nyelveiben, őket követik a nazálisok, majd a zöngétlen réshangok (Laver 1994). Ha e tekintetben megnézzük a magyar mássalhangzók gyakorisági eloszlását, akkor azt látjuk, hogy két zöngétlen felpattanót zárhang, két nazális és egy zöngétlen réshang a leggyakoribb mássalhangzók között szerepel. Az univerzális tények tehát reprezentáltak lehetnek a specifikus adatokban. Adataink egyértelműen igazolták továbbá a Zipf-törvényt.

A beszéd kvantitatív elemzésének eredményei több területen is jól használhatók. Számos artikulációs sajátosság nem magyarázható az ismert tényezőkkel, ilyenkor merül fel, hogy ezeknek a hátterében a gyakorlatlanabb artikulációs rutin állhat (pl. Lewitt–Healey 1985). A lexikális hozzáférés tanulmányozásához, a szófelismerés előjelzési stratégiáinak megismeréséhez ugyancsak nélkülözhetetlenek bizonyos gyakorisági adatok, esetenként magyarázó elvül szolgálhatnak az eredményekhez (Cutler et al. 2004). Az alkalmazott beszédkutatás is több vonatkozásban használ nyelvstatisztikai adatokat, diagnosztikai tesztek, kísérletek és nyelvi fejlesztések anyagainak összeállításakor (pl. beszédaudiometriai tesztek), illetve a mesterséges intelligencia különféle területein.

Irodalom

Altmann, Gabriel 1993. Phoneme counts: Marginal remarks on the Pääkkönen article. In Altmann, Gabriel (ed.): *Glottometrika 14*. Wissenschaftlicher Verlag Trier, Trier, 54–68.

Bell, Alan – Gregory, Michelle L. – Brenier, Jason M. – Jurafsky, Daniel – Ikeno, Ayako – Girand, Cynthia 2002. Which predictability measures affect content word durations? In: *Pronunciation Modeling and Lexicon Adaptation for Spoken Language Technology 2002*. 1–5.

Bybee, Joan L. 2006a. *Frequency of use and the organization of language*. Oxford University Press, Oxford.

Bybee, Joan L. 2006b. From usage to grammar: the mind's response to repetition. *Language* 82. 711–733

Crystal, David 1997. *A nyelv enciklopédiája*. Osiris Kiadó, Budapest.

Cutler, Anne – Carter, David M. 1987. The predominance of strong initial syllables in the English vocabulary. *Computer Speech und Language* 2. 133–142.

Cutler, Anne – Norris, Dennis – Sebastián-Gallés, Núria 2004. Phonemic repertoire and similarity within the vocabulary. In Kin, S. – Bae, M. J. (eds.): *Proceedings of the 8th International Conference on Spoken Language Processing (Interspeech 2004-ICSLP)*. Sunjijn Printing Co., Seoul, 65–68.

Dömölki Bálint – Fónagy Iván – Szende Tamás 1964. Köznyelvi hangstatisztikai vizsgálatok. *Általános Nyelvészeti Tanulmányok* II. 117–132.

Fodor István 1960. A statisztikai módszer alkalmazásának néhány kérdése. *Magyar Nyelvőr* 84. 196–205.
Fónagy Iván 1962. Beszéd és valószínűség. *Magyar Nyelvőr* 86. 309–320.
Förstemann, Ernst 1846. Über die numerischen Lautverhältnisse im Deutschen. *Germania* 7. 83–90.
Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
Gósy Mária 2012. Szóismétlések a funkció és a beszédtípus függvényében. In Bárdosi Vilmos (szerk.): *A szótól a szövegig*. Tinta Kiadó, Budapest, 75–86.
Kelemen József 1967. Lyukkártyarendszerű adatgyűjtés és feldolgozás a nyelvtudományban. *Nyelvtudományi Közlemények* 68. 343–379.
Kessler, Brett – Treiman, Rebecca 1997. Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and Language* 37. 295–311.
Keszler Borbála 1983. Kötetlen beszélgetések mondat- és szövegtani vizsgálata. In Rácz Endre – Szathmári István (szerk.): *Tanulmányok a mai magyar nyelv szövegtana köréből*. Tankönyvkiadó, Budapest, 164–202.
Laver, John 1994. *Principles of phonetics*. Cambridge University Press, Cambridge.
Levitt, Andrea – Healy, Alice F. 1985. The roles of phoneme frequency and availability int he experimental elicitation of speech errors. *Journal of Memory and Langage* 24. 717–733.
Maddieson, Ian 1986. The size and structure of phonological inventories: Analysis of UPSID. In Ohala, John J. – Jaeger, Jeri J. (eds.): *Experimental phonology*. Academic Press, Orlando, 105–125.
Martin, Andy 2009. Why are phoneme frequencies so skewed? http://www.slidefinder.net/w/why_phoneme_frequency_distributions_skewed/martin_lsa_2009_phoneme_frequencies/30594100/p2 (Letöltés: 2011. november 18.)
Martindale, Colin – Gusein-Zadeb, S. M. – McKenzie, Dean – Borodovsky, Mark Yu 1996. Comparison of equations describing the ranked frequency distributions of graphemes and phonemes. *Journal of Quantitative Linguistics* 3. 106–112.
Menzerath, Paul 1954. *Die Architektonik des deutschen Wortschatzes*. Dümmlers Verlag, Bonn–Hannover–Stuttgart.
Nagy Ferenc 1972. *Kvantitatív nyelvészet*. Egyetemi jegyzet. ELTE, Budapest.
Papp Ferenc 1961. A matematikai módszerek alkalmazásáról a nyelvtudományban. *MTA I. OK*. 17. 141–172.
Petőfi S. János 1967. A nyelvstatisztikai vizsgálatok néhány kérdése: nyelvfeldolgozás és dokumentáció. In Szépe György (szerk.): *A tudományos tájékoztatás elmélete és gyakorlata* 11. Budapest, 117–140.
Renwick, Margaret E. L. 2011. Phoneme type frequency in Romanian. *Penn Working Papers in Linguistics* 22. 195–204. http://repository.upenn.edu/cgi/viewcontent.cgi?article=1182&context=pwpl&sei-redir=1&referer=ht (Letöltés: 2011. december 12.)
Rovenchak, Andrij 2011. Phoneme distribution, syllabic structure, and tonal patterns in Nko texts. *Mandenkan* 47. 77–96.
Sandoval, Antonio M. – Toledano, Doroteo T. – de la Torre, Raúl – Garrote, Marta – Guirao, José M. 2008. Developing a phonemic and syllabic frequency inventory for spontaneous spoken Spanish and their comparison to text-based inventories. 1097–1099. http://atvs.ii.uam.es/files/2008_LREC_Moreno.pdf (Letöltés: 2012. január 4.)
Sigurd, Bernd 1968. Rank-frequency distributions for phonemes. *Phonetica* 18. 1–15.

Simonyi Zsigmond 1896. A betűk és hangok előfordulásának aránya. *Magyar Nyelvőr* 25. 325.
Siptár, Péter – Törkenczy, Miklós 2000. The phonology of Hungarian. Oxford University Press, Oxford.
Szende Tamás 1973. *Spontán beszédanyag gyakorisági mutatói.* Nyelvtudományi Értekezések 81. Akadémiai Kiadó, Budapest.
Tambovtsev, Yuri – Martinadale, Colin 2007. Phoneme frequencies follow a Yule distribution. *SKASE Journal of Theoretical Linguistics* 4. http://www.skase.sk/Volumes/JTL09/pdf_doc/1.pdf (Letöltés: 2011. 09. 23.)
Tarnóczy Tamás 1942. A magyar magánhangzók eloszlása. *Magyar Nyelv* 38. 352–357.
Tolnai Vilmos 1906. Egynéhány számadat a hangokról és betűkről. *Magyar Nyelvőr* 35. 421–425.
G. Varga Györgyi 1967. *Alakváltozatok a budapesti köznyelvben.* Akadémiai Kiadó, Budapest.
Vértes Edit 1955. Szótározott szavaink száma és a szókészlet nagysága. *Magyar Nyelvőr* 79. 329–338.
Weiss, M. 1961. Über die relative Häufigkeit der Phoneme des Schwedischen. *Statistical Methods in Linguistics* 1. 41–55.
Zipf, George 1949. *Human behavior and the principle of least effort: An introduction to human ecology.* Addison-Wesley Press, Boston, Massachusetts.

A kutatás a 78315 sz. OTKA-pályázat keretében készült.

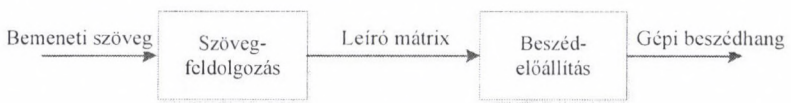
BESZÉDKORPUSZ TERVEZÉSE MAGYAR NYELVŰ, REJTETT MARKOV-MODELL ALAPÚ SZÖVEGFELOLVASÓHOZ

Tóth Bálint – Németh Géza – Olaszy Gábor

Bevezetés

A gépi beszéd-előállítás lehetőségét először Kempelen Farkas alkotta meg 1791-ben (Nikléczy–Olaszy 2004), de a számítógépes szöveg-beszéd átalakítók fejlődése is több évtizedes múltra tekint vissza mind nemzetközi (Moulines–Charpentier 1990; Hunt–Black 1996; Black–Lenzo 2000; Möbius 2000), mind pedig hazai viszonylatban (Kiss–Olaszy 1982).

A gépi szöveg-beszéd átalakítás általánosságban két fő lépésből áll: a szövegfeldolgozó és a beszéd-előállító részekből (1. ábra).



1. ábra
A gépi szöveg-beszéd átalakítás általános struktúrája

A szövegfeldolgozó modul a bemeneti szövegen végez különböző átalakításokat, majd az eredmény alapján egy leíró adatmátrixot hoz létre, melyben szerepelnek a bemeneti szöveget reprezentáló beszédhangok és a bemeneti szövegre jellemző szegmentális és szupraszegmentális információk. Ezt a leíró mátrixot kapja meg a beszéd-előállító modul, amelyből előállítja a gépi beszéd hullámformáját.

Az elmúlt évtizedekben különböző megközelítésekkel modellezték az emberi beszéd-előállítás mechanizmusát. A gépi beszéd-előállító modul két nagy csoportját különböztetjük meg: a szabály- és az adatvezérelt beszédhang-előállítást.

Szabályvezérelt beszédhang-előállítás esetén mérésekből és tapasztalati úton szerzett információk alapján felállított szabályok szerint állítjuk elő a gépi beszédhangot. Ilyen megoldás például az artikulációs (Mermelstein 1973), illetve a formánsszintézis (Kiss–Olaszy 1982; Klatt–Klatt 1990) és a hullámforma-összefűzésen alapuló diád- és triádalapú rendszerek (Moulines-Charpentier 1990; Olaszy et al. 2000).

Adatvezérelt beszédhang-előállítás esetén úgynevezett beszédkorpuszokból kinyert információk alapján állítjuk elő a gépi beszédhangot. A beszédkorpusz tartalma és felépítése alapvetően meghatározza az eredményt. Általában több órányi emberi beszédet használnak, továbbá annak fonemikus átiratát, a hanghatárokat és még további lehetséges információkat (pl. hangsúlyok). Adatvezérelt megoldás például a korpuszalapú elemkiválasztásos beszédszintézis (Möbius 2000; Németh et al. 2006) és a rejtett Markov-modell (Hidden Markov Model, HMM) alapú beszédszintézis (Yoshimura et al. 1999; Black et al. 2007). Az első, magyar nyelvre készített rejtett Markov-modell alapú szövegfelolvasó fejlesztésének részleteit Tóth és Németh (2010) tanulmánya tartalmazza.

Napjainkban a rejtett Markov-modell alapú szövegfelolvasó rendszerek a leginkább elterjedtek számos előnyös tulajdonságuk miatt. Kisméretű (1,5-2 Mbyte) beszédadatbázisból képesek jó, közel állandó minőségű, érthető beszédet előállítani, amely hordozza a beszélő hangszínezeti tulajdonságait is (Yoshimura et al. 1999). Ezen túl jelentős előny más megoldásokkal szemben, hogy lehetőség van viszonylag kicsi (5-10 percnyi) beszédkorpusz segítségével a jellemző paraméterfolyamok, és így a gépi beszédhang karakterisztikájának adott célbeszélőhöz való adaptációjára (Tamura et al. 1998; Ogata et al. 2006; Yamagishi–Kobayashi 2007). Hátránya is van az eljárásnak, nehéz a teljes prozódiai rendszert megvalósítani (kérdés, óhajtás stb.), az előállított beszéd hangkimeneti korrektsége nem hibamentes (rossz hangidőtartamok, hangkimaradások előfordulhatnak). További hátrány, hogy a felmerülő hangzási hibák korrigálása nehéz, hiszen nem látunk bele a tanulási és szintetizálási folyamat részleteibe.

Célkitűzés

Célunk az volt, hogy megállapítsuk, hogy a rejtett Markov-modell alapú szövegfelolvasás esetén javul-e a hangminőség, ha a beszédkorpusz automatikusan végzett címkézését utólagosan kézi ellenőrzésnek vetjük alá, és az esetleges címkehibák számát így gyakorlatilag nullára csökkentjük.

Rejtett Markov-modell alapú szövegfelolvasás

A 2. ábra egy általános HMM-alapú szöveg-beszéd átalakító blokkdiagramját mutatja be. Az eljárás két fő részből áll: tanítási és szintézis szakaszból. A tanítási szakaszban a tanító beszédkorpusz hullámformáiból kinyerjük a gerjesztési és spektrális paramétereket, majd ezen paramétereket és a beszédkorpuszhoz tartozó további információkat (pl. fonemikus átirat, hanghatárok, szegmentális és szupraszegmentális információk) adjuk át a HMM tanítási szakasznak. A tanító beszédkorpusz néhány órányi, lehetőleg stúdióminőségű felvételt tartalmaz egy beszélőtől, illetve a hanganyag fonemikus átiratát és minél pontosabb hanghatárjelöléseket.

Tehát először a jellemző – gerjesztési és spektrális – paramétereket kinyerjük a beszédkorpuszból, majd ezen paraméterek sokaságát generatív modellekkel helyettesítjük. A modell paramétereinek becslésére a következő képlet

szerint általában a maximum likelihood (ML) vagy ahhoz hasonló becslést alkalmaznak:

$$\hat{\lambda} = \arg\max_{\lambda} \{ p(\mathbf{O} \mid W, \lambda) \} \quad (1)$$

ahol λ a modell paramétereit, $\boldsymbol{O}$ a beszédkorpuszból származó jellemző paramétereket (tanítóadatok) és W az $\boldsymbol{O}$-hoz tartozó szósorozatot jelöli.

A címkék a hullámforma szövegének fonemikus átiratán és az időzítéseken túl számos szegmentális és szupraszegmentális információt tartalmaznak hang-, szótagmag-, szó-, mondatrész- és mondatszinten (Zen et al. 2007). Ezeket a címkéket környezetfüggő címkéknek is nevezzük.

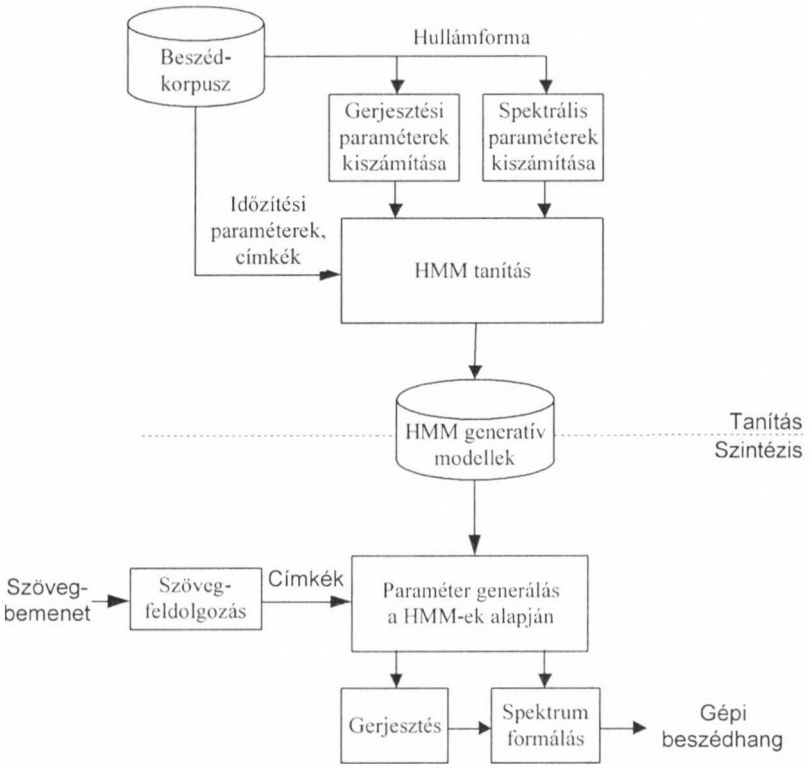
A környezetfüggő címkék lehetséges kombinációja túl nagy ($> 10^{53}$) ahhoz, hogy megfelelően reprezentatív beszédkorpuszt tudjunk hozzá készíteni, ezért a paraméterfolyamokat döntési fák segítségével csoportokba soroljuk (Yoshimura et al. 1999). Külön-külön döntési fa tartozik az egyes paraméterfolyamokhoz. Például 40 beszédhanggal számolva a kvinfón (a vizsgált hang és az előtte, ill. utána következő 2-2 hang) lehetséges változatainak száma $40^5 = 102\ 400\ 000$, melyhez hozzávéve a további környezetfüggő címkéket óriásira növeljük az állapotteret.

A HMM tanítási szakaszában a gerjesztési, a spektrális és az időzítési paraméterekhez készítünk generatív modelleket. A folytonos paraméterfolyamokat (pl. spektrális paraméterek) Gauss-eloszlásokkal közelítjük (Yoshimura et al. 1999), míg a diszkrét/folytonos paraméterfolyamokat (pl. gerjesztési paraméterek esetén diszkrét módon jelezzük a zöngétlen hangot, zöngés esetben pedig folytonos módon az alapfrekvenciát) többterű valószínűségi eloszlású HMM-ekkel (Multi-Space Probability Distribution HMM, MSD-HMM) modellezzük (Tokuda et al. 1999). Annak érdekében, hogy a hangok és a mondat ritmikáját megfelelően modellezni tudjuk, a HMM-állapotok közötti átmeneti valószínűségeket nem egy értékkel, hanem Gauss-eloszlásokkal írjuk le.

A tanítási folyamat végére előáll az úgynevezett HMM adatbázis, mely a beszédkorpusz jellemző paraméterfolyamainak generatív modelljeit tartalmazza. A szintézis során ezen modellek segítségével fogunk gépi beszédhangot előállítani a következő módon. A w szósorozathoz és $\hat{\lambda}$ becsült modell paraméterekhez tartozó o paraméterek kimeneti valószínűségét maximalizáljuk az alábbiak szerint:

$$\hat{o} = \arg\max_{o} \{ p(o \mid w, \hat{\lambda}) \} \quad (2)$$

Tehát a (2)-es egyenlet maximalizálását hajtjuk végre: a HMM generatív modellekből a bemeneti szöveg alapján előállított környezetfüggő címkehalmazra legjellemzőbb paraméterfolyamokat generáljuk. Ebből a paraméterfolyamból állítjuk elő a gépi beszédhangot olyan beszédkódoló eljárással, mint például az impulzus-zaj gerjesztésen alapuló LPC beszédkódoló.



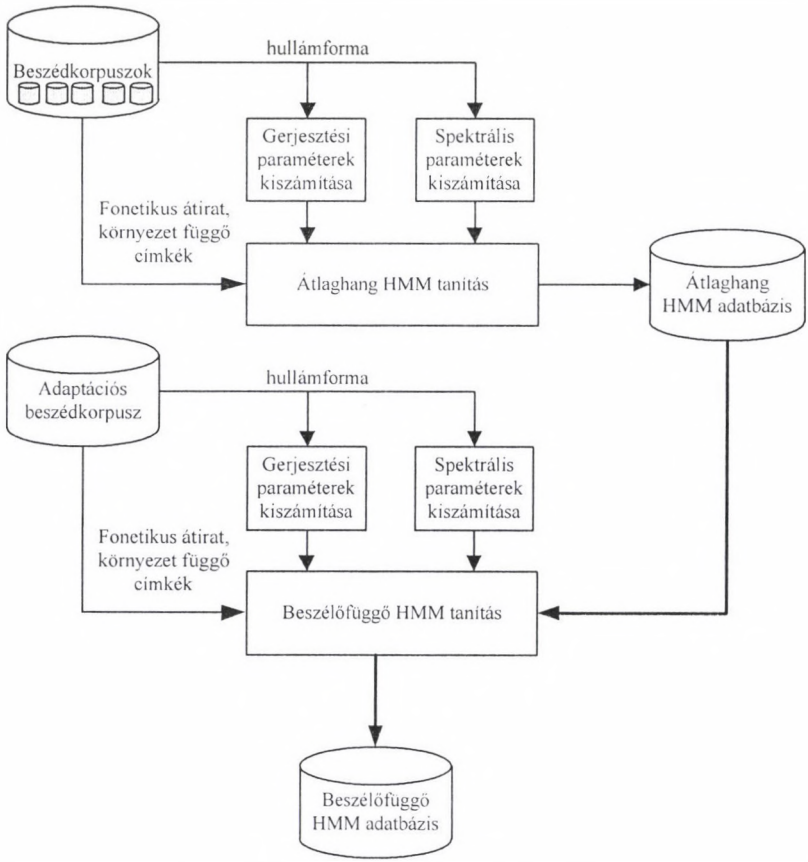
2. ábra
HMM-alapú szövegfelolvasó tanítási és szintézis fázisának blokkdiagramja

Rejtett Markov-modell alapú beszélőadaptáció

Az előző részben a beszélőfüggő tanítást mutattuk be. A HMM-alapú beszédszintézis egyik nagy előnye, hogy képes a beszélőadaptációra. A beszélőadaptáció annyit jelent, hogy a rendszer hangkarakterisztikáját képesek vagyunk egy adott célbeszélőhöz hasonlóra kialakítani. Más megoldásokkal szemben a rejtett Markov-modellek esetén előny, hogy a beszélőadaptációhoz viszonylag rövid, 5-10 perces beszédkorpusz elegendő.

A beszélőadaptált tanítás folyamata hasonló a beszélőfüggő esethez, azonban itt a tanítást két fő részre oszthatjuk: először egy átlaghangot tanítunk, melyet utána a célbeszélő hangkarakteréhez igazítunk (3. ábra). Ebben az esetben így áll elő a szintézis alapját képező HMM-adatbázis. Ezután a beszédhang előállításának módszere megegyezik a beszélőfüggő esetben használt módszerrel (2. ábra alsó része). Az átlaghang előállításához több beszélő-

től (legalább 4-5), minél hosszabb (személyenként legalább 1-1,5 óra) hangfelvételre, annak fonemikus átiratára és pontos hanghatárjelöléseire van szükség. Ezután a HMM-eket az összes beszélő adatbázisa alapján tanítjuk be az átlaghangra, melyben jelen vannak minden egyes beszélőre az alapfrekvencia, hangidőtartam és spektrális paraméterek.



3. ábra
A beszélőadaptált tanítás blokkdiagramja

Az átlaghang tanításához használhatunk férfihangot, nőit vagy mindkettőt. A gyakorlatban a kevert nemű átlaghang-előállítást célszerű választanunk, majd ebből adaptálni mind férfi, mind női hangra. Meg lehet csinálni, hogy el-

lentétes nemű átlaghangból adaptálunk női/férfihangra, azonban Isogai és munkatársai (2005) kutatásában ez jelentős minőség- és természetességcsökkenést okozott a végső hangnál a nemenkénti átlaghanghoz képest. Yamagishi és munkatársai (2007) olyan eljárásról számolnak be, mely segítségével kevert nemű átlaghangból a nemenkénti átlaghanghoz képest minimális minőség- és természetességromlás mellett lehet női és férfi hangra adaptálni.

Miután elkészültek az átlaghang HMM modelljei, a célbeszélőtől származó hangfelvételekkel tudjuk a modellt az adott személy hangkarakteréhez és beszédstílusához igazítani, adaptálni. A beszélő adaptációjára alapvetően kétfajta lehetőségünk van.

Amennyiben kevés (5-10 perc) hanganyag áll rendelkezésre a célbeszélőtől, akkor előnyös maximum likelihood linear regression (MLLR) alapú adaptációt választani (Ogata et al. 2006). Tamura és munkatársai (1998) kísérlete alapján akár már öt mondat is elegendő lehet ahhoz, hogy a célszemély hangkarakterét és beszédstílusát visszaadja a gépi beszédhang.

Amennyiben hosszabb (több mint 1 óra) adaptációs hanganyag is elérhető, akkor a maximum a posteriori (MAP) technikát érdemes használni (Yamagishi et al. 2007), mely az előzőnél jobb minőségű mesterségesen generált hangot eredményez. Ennek a technológiának az új változatai, mint például a CSMAPLR (constrained structural maximum a posteriori linear regression) közel azonos minőséget és természetességet képviselnek, mint a beszélőfüggő tanítás esetén előállított mesterséges beszéd (Yamagishi et al. 2009).

A beszélőadaptáció során MLLR eljárást használtunk. Az MLLR lineáris transzformációk segítségével az átlaghang HMM modell paramétereit a célhang „irányába" módosítja. Az állapotkimenetek ekkor a következőképp alakulnak:

$$b_j(o_t) = N(o_t; \hat{\mu}_j; \hat{\Sigma}_j) \tag{3}$$

$$\hat{\mu}_j = A_{r(j)}\mu_j + b_{r(j)} \tag{4}$$

$$\hat{\Sigma}_j = H_{r(j)}^T \Sigma_j H_{r(j)} \tag{5}$$

ahol $\hat{\mu}_j$ és $\hat{\Sigma}_j$ a j-edik állapotra jellemző kimeneti sűrűségfüggvényhez tartozó várható értékvektor, ill. kovarianciamátrix a lineáris transzformáció után. $A_{r(j)}$, $b_{r(j)}$ és $H_{r(j)}$ a várható érték lineáris-transzformációs mátrixa, a hozzá tartozó eltolásvektor és a kovariancia lineáris-transzformációs mátrixa az *r(j)*-edik regressziós osztályban.

Az adott állapotokra jellemző kimeneti sűrűségfüggvényeket regressziós fa segítségével osztályokba soroljuk, egy adott osztályban azonos lineáris-

transzformációs mátrixokat és eltolásvektort használunk. A regressziós fa méretének az adaptációs anyag mennyiségéhez való igazításával tudjuk szabályozni az adaptáció komplexitását és általánosítható képességét. Alapvetően az MLLR két fajtáját különböztetjük meg: azonos A és H lineáris-transzformációs mátrixok esetén korlátozott MLLR-ről (constrained MLLR, CMMLR), egyébként pedig korlátozás mentes MLLR-ről (unconstrained MLLR) beszélünk. A jelen cikkben ismertetett rendszer esetén CMLLR-t használtunk.

Anyag és módszer

A kutatásaink során adott szövegkorpuszokból felolvasással beszédkorpuszokat készítettünk. Ezen beszédkorpuszokat címkékkel láttuk el (szegmentálás). Megvizsgáltuk a fonématévesztés hibaarányát (phone error rate, PER), valamint a címketévesztés hatását a hangminőségre. Az eredmények szubjektív értékelése céljából meghallgatásos tesztet végeztünk.

Beszédkorpuszok

Beszédkorpuszon a következőt értjük: hanganyag, a felolvasott szöveg fonemikus átirata és szegmentálási címkék halmaza. A betanításhoz felhasznált beszédkorpuszok felolvasott beszédből készültek. Minden bemondó egységesen ugyanazt a szöveget olvasta fel, amely fonetikailag kiegyensúlyozott mondatokat (Vicsi et al. 2004) tartalmazott (kb. 2000 mondat). Öt beszédadatbázis készült, 4 férfi bemondó (életkoruk 60, 50, 30, 29 év) és egy női (33 év) hangból. A beszédkorpusz digitalizálási adatai: 16 kHz, 16 bit.

Az öt beszédadatbázishoz (1. táblázat) az eredeti szövegkorpuszok automatikus módszerrel készített fonemikus átiratát használtuk, ugyanakkor az F1 és N1 beszédadatbázison kézileg is ellenőriztük a fonemikus átirat és a szegmentálás pontosságát. A későbbiekben ez utóbbi két változatot tekintjük referenciának.

A beszélőfüggő tanítás során egyszer az F1 és N1 adatbázis automatikus módszerekkel elkészült változatát, majd a kézi ellenőrzésen is átesett változatait használtuk fel (összesen 4 darab). Beszélőadaptált esetben az átlaghangot az automatikus módszerekkel elkészített öt beszédkorpusszal tanítottuk. Ezután a beszélőadaptációt az F1 és N1 adatbázisok egy részhalmazával végeztük el (szintén a kézi és automatikus változatokkal, összesen itt is 4 darab).

Az átlaghang tanítása során a szövegek fonemikus átírása kizárólag automatikus úton történt az eredeti szövegből. Megjegyezzük, hogy az ideális eset az lenne, ha az átlaghang adatbázisai mind kézzel ellenőrzöttek lennének. Kutatásunk jelenlegi szakaszában azonban még nem állt rendelkezésre az összes beszédkorpusz kézi átirata javított hanghatárokkal, ezért a konzekvens adatbázis-építés elvét alkalmaztuk, mindegyik beszédkorpuszból az automatikus átiratot használtuk.

1. táblázat: A kutatás során használt beszédkorpuszok

Beszélő	Mondatszám	Időtartam	Feldolgozás
1. férfi beszélő (F1)	1936	190 perc	automatikus és kézi ellenőrzés
2. férfi beszélő (F2)	1938	137 perc	automatikus
3. férfi beszélő (F3)	1941	170 perc	automatikus
4. férfi beszélő (F4)	1938	214 perc	automatikus
1. női beszélő (N1)	1937	128 perc	automatikus és kézi ellenőrzés

Fonématévesztési hibaarány

Módszert dolgoztunk ki a fonématévesztések hibaarányának (PER) megállapítására. A következő hibafajtákat kezeltük:

– A bemondó mást olvas fel, mint ami a szövegben van. 1. példa: szöveg: *és*, fonetikai átirat: *és*, kimondva: *s*. 2. példa: szöveg: *lehetlen*, átirat: *lehtetlen*, kimondva: *lehetetlen*.

– A fonetikai átíró mást jelöl, mint az elhangzott elem. Példa: szöveg: *900*, fonetikai átirat: *kilencszáz*, kimondva: *kilencáz*.

A javítás módszere: a gépileg felcímkézett adatállományokat kézi ellenőrzésnek vetjük alá. Ennek során a fonemikus átiratot és az elhangzott hanganyagot egybevetjük, és a hibákat javítjuk a megfelelő helyen.

Minden esetben a kézzel javított fonemikus átiratot tekintjük referenciának. A kézzel javított átiratról feltételezzük, hogy hangról hangra azt tartalmazza, ami a beszédkorpusz hanganyagában szerepel. A fonématévesztések hibaarányának meghatározása során a referenciához képest vizsgáljuk az automatikusan készített fonemikus átiratot.

A fonématévesztések hibaarányának számítása közben a fontos jellemzők:

– Fonémák száma: az összes fonéma száma a beszédkorpuszban.

– Helyes fonémák száma: a kézi átirathoz képest mennyi fonéma azonos.

– Törlések: az automatikus átiratban a kézi átirathoz képest kitöröltünk egy fonémát.

– Helyettesítés: az automatikus átiratban a kézi átirathoz képest helyettesítettünk egy fonémát.

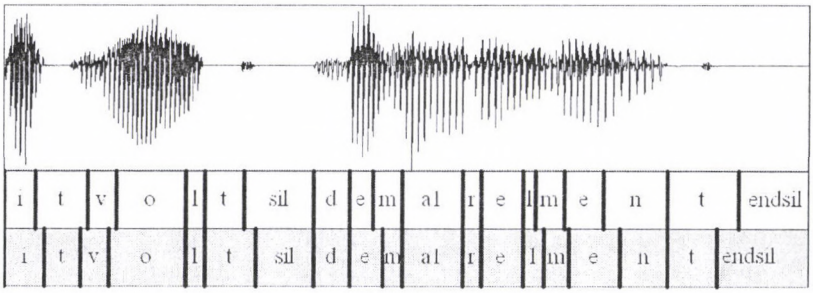
– Beszúrás: az automatikus átiratban a kézi átirathoz képest beszúrtunk egy fonémát.

A javítások száma a törlések, helyettesítések és beszúrások számának összege. A fonématévesztések hibaarányát a következőképp számoltuk (Young et al. 2006):

$$\mathrm{PER} = \frac{\mathrm{HelyesFonémákSzáma}}{\mathrm{FonémákSzáma}} \qquad (6)$$

Szegmentálás

Szegmentáláson azt a folyamatot értjük, mely során a hangfelvételek hullámformáján címkével bejelöljük minden hang és szünet kezdetét (hanghatárok). Szegmentálási szempontból kétféle címkét különböztetünk meg, a hang elejét jelző jelölést és a szünetek elején elhelyezett címkét (hangsor belseji szünet és hangsor végi). Ezeket a hanghatárokat automatikus módszerrel, úgynevezett kényszerített felismeréssel (forced alignment) határoztuk meg, ami gépi beszédfelismerési módszeren alapul (Mihajlik et al. 2002). Ekkor a beszédfelismerő a bemeneti szöveget elemezve jelöli ki a hanghullámban a hangokat és a hanghatárokat. Az eljárást azért hívják kényszerített felismerésnek, mert a szövegből adódik, hogy milyen hangok követik egymást, továbbá hogy a hanghullámban csak annyi hangot jelölhet be az algoritmus, amennyi a szövegből következik. Fontos figyelembe vennünk, hogy hibátlan fonemikus átirat esetén is lehetnek hibás hanghatárcímkék. Ezért a kényszerített felismerés szegmentálási eredményét kézi ellenőrzéssel javítottuk (4. ábra). Mivel a kézi ellenőrzés nagy emberi erőforrást igényel, ezért ezt a használt beszédkorpuszoknak csak egy részében végeztük el (lásd 1. táblázat).



4. ábra
A hanghatárok automatikus, gépi elhelyezése (fent) és a kézi javítás eredménye (lent) az *Itt volt, de már elment.* mondatban

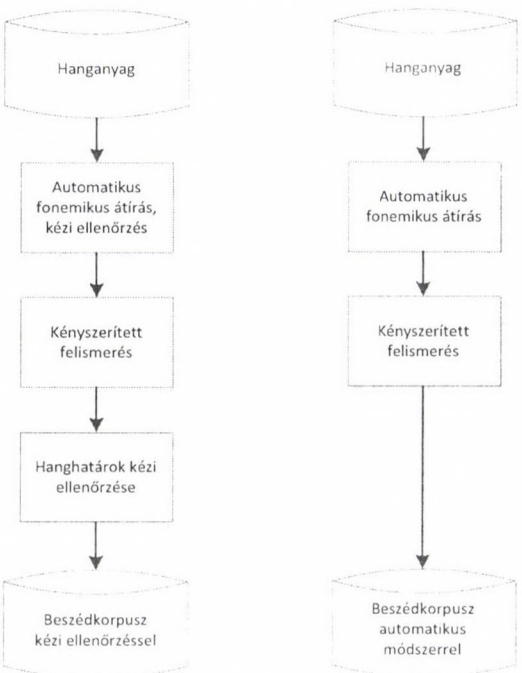
A hanghatárhibákat osztályokba soroltuk be – attól függően, hogy hány ms eltérés van a referencia és az automatikus hanghatár-meghatározás között – a következőképp: 0,1–9 ms, 10–19 ms, 20–29 ms, 30–39 ms, 40–49 ms, 50–59 ms, 60–69 ms, 70–79 ms, 80–89 ms, 90 ms-nál több.

A szegmentálás végeztével előáll(nak) a HMM-tanításhoz, illetve adaptációhoz szükséges beszédkorpusz(ok). Az automatikus és kézi ellenőrzéssel történő beszédkorpuszok előállítási folyamatát az 5. ábra szemlélteti. Az ábrán szürke dobozokkal jelöltük a tanítás/adaptáció során felhasznált elemeket.

A gépi beszéd előállítása

A HMM-szintézishez a HTS 2.1-es rendszer módosított, magyar változatát használtuk kevert gerjesztéssel (Tóth–Németh 2010). A lényegkiemeléshez

25 ms-os Hanning-ablakot használtunk 5 ms eltolással. A jellemző paraméterfolyam 39 mel-kepsztrális együtthatót, a $\log\{F_0\}$-t, a zöngésségi arányt és az ezekhez tartozó delta és delta-delta együtthatókat tartalmazta. Kutatásunk során vizsgáltuk a beszélőfüggő és beszélőadaptált eseteket is.



5. ábra
A beszédkorpuszok előállítási folyamata
kézi ellenőrzéssel (bal) és automatikusan (jobb)

Meghallgatásos teszt

Az eredmények kiértékelése céljából meghallgatásos tesztet állítottunk össze. A korábban ismertetett módon összesen 8 különböző gépi beszédelőállító rendszert hoztunk létre a teszthez: BF-F1-kézi, BF-F1-automatikus, BF-N1-kézi, BF-N1-automatikus, BA-F1-kézi, BA-F1-automatikus, BA-N1-kézi, BA-N1-automatikus. A meghallgatásos teszt két részből állt. Az első részben a tesztalanyok a hangminták hangzásának természetességét osztályozták a következő rendszerek esetén: BF-F1-kézi $\leftrightarrow$ BF-F1-automatikus, BF-N1-kézi $\leftrightarrow$ BF-N1-automatikus, BA-F1-kézi $\leftrightarrow$ BA-F1-automatikus, BA-N1-kézi $\leftrightarrow$ BA-N1-automatikus.

A meghallgatásos teszt első részében minden rendszerből 10 hangmintát (mondatot) használtunk fel, páronként a minták szövege azonos volt. Egy tesztelő a 10 mintából 2 mintapárt hallgatott meg a beszélőfüggő férfi és női, illetve beszélőadaptált férfi és női HMM-rendszerek által generált hangminták közül. A tesztalanyok ötelemű skálán osztályozhatták, hogy a mintapár mintái ugyanolyan természetesek-e, illetve hogy valamelyik mintát kicsit vagy sokkal természetesebbnek érzik-e a másik mintánál. A mintapárokat mindkét sorrendben lejátszottuk a tesztalanyoknak, hogy ellenőrizzük az ítéletalkotás következetességét. Amennyiben jelentős eltérés volt az osztályozásban a különböző sorrendek esetén, akkor ezen méréseket nem vettük figyelembe a tesztek során. Ilyen jellegű eltérés csupán néhány esetben volt észlelhető, de nem volt jellemző sem a beszélőre, sem a tesztalanyokra.

A meghallgatásos teszt második felében az eredeti beszélő természetes bemondásához hasonlították a tesztalanyok a gépi beszédet, hogy mennyire adja vissza annak hangzását. Ebben az esetben mind a nyolc rendszer részt vett a tesztben.

A meghallgatásos teszt második felében is összesen 10 hangmintát használtunk fel, egy tesztalany egy rendszerből egy hangmintát hallgatott meg, és 1-től 5-ig kellett a mintákat osztályoznia. Az 1-es osztályzat itt azt jelentette, hogy egyáltalán nem adja vissza az eredeti beszélő hangkarakterét, az 5-ös pedig, hogy a szintetizált hangminta összetéveszthető az eredeti beszélővel.

A meghallgatásos teszt mindkét részében minden tesztalany esetén másmás sorrendben játszottuk le a hangmintákat, így zárva ki az esetleges „memóriahatásokat" (van Santen 1993).

A tesztet összesen 29-en végezték el, 14 férfi és 15 nő. Az átlagéletkor 31 év volt, a legfiatalabb tesztalany 20, a legidősebb 65 éves volt. 6 tesztalany beszédszakértő volt. A teszt internetalapú volt, böngészőből lehetett kitölteni, a hangminták MP3 kódolással voltak tárolva 128 kbps, 16 bit minőségben.

Eredmények

A korábbiak alapján az eredmények ismertetését három részre bontjuk: megvizsgáljuk a fonématévesztések hibaarányát (PER) a beszédkorpuszokban, ismertetjük az automatikus és kézi szegmentálás különbségeit, illetve meghallgatásos tesztekkel vizsgáljuk a pontos címkézés hatását a gépi beszéd minőségére.

Fonéma tévesztések hibaarányának vizsgálata

A kísérleteink során összesen 8-féle beszédkorpuszt vizsgáltunk: 4 korpuszt a beszélőfüggő tanításhoz és 4 korpuszt a beszélőadaptált változathoz (ezek a korábbi korpuszokat, illetve azoknak részhalmazait tartalmazták).

Az adaptációs kis beszédkorpuszt az F1 és N1 korpuszból származtattuk, és elemeit úgy választottuk ki, hogy az adott F1 és N1-ből az összes törlés, helyettesítés és beszúrás szerepeljen benne. Ezáltal az automatikus fonemikus átiratból a legtöbb hibát tartalmazó részeket tartottuk meg, továbbá véletlen

módon kiválasztottunk hozzá annyi hanganyagot, hogy 10 perc körüli időtartam álljon rendelkezésünkre.

Az eredményeket a 2. és a 3. táblázat mutatja. BF-fel jelöljük a beszélőfüggő, BA-val a beszélőadaptált eljárásokhoz tartozó korpuszokat. A 2. táblázat alapján megállapíthatjuk, hogy a beszélőfüggő esetekben a hibák száma elenyésző a teljes korpusz nagyságához képest. A 3. táblázat pedig megmutatja, hogy az adaptációs beszédkorpusz esetén már nagyobb mértékű hibaaránnyal kell számolnunk.

2. táblázat: A beszélőfüggő tanításhoz felhasznált beszédkorpuszok

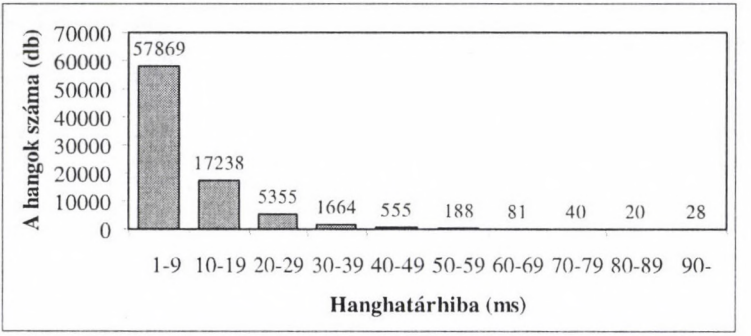
	BF-F1-kézi	BF-F1-automatikus	BF-N1-kézi	BF-N1-automatikus
Mondatszám	1936	1936	1937	1937
Időtartam	190 perc	190 perc	128 perc	128 perc
Fonémaszám	80964	81053	80893	81058
A helyes fonémák száma	80964	80380	80893	80663
Törlések	–	32	-	51
Helyettesítések	–	57	-	114
Beszúrások	–	584	-	260
A javítások száma	–	673	-	425
PER	0%	0,83%	0%	0,52%

3. táblázat: A beszélőadaptációhoz felhasznált beszédkorpuszok

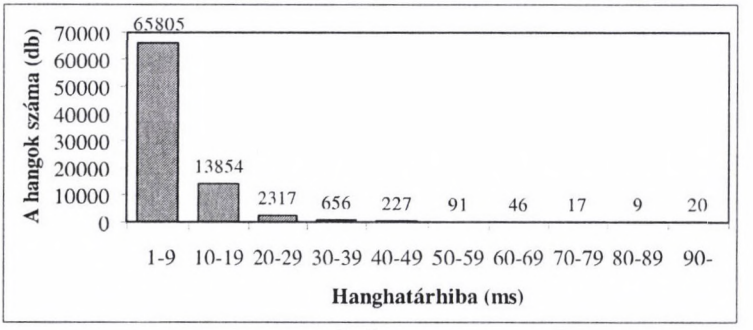
	BA-F1-kézi	BA-F1-automatikus	BA-N1-kézi	BA-N1-automatikus
Mondatszám	104	104	164	164
Időtartam	10 perc	10 perc	11 perc	11 perc
Fonémaszám	4281	4370	6934	7099
A helyes fonémák száma	4281	3697	6934	6674
Törlések	–	32	-	51
Helyettesítések	–	57	-	114
Beszúrások	–	584	-	260
A javítások száma	–	673	-	425
PER	0%	15,5%	0%	6%

A szegmentálás vizsgálata

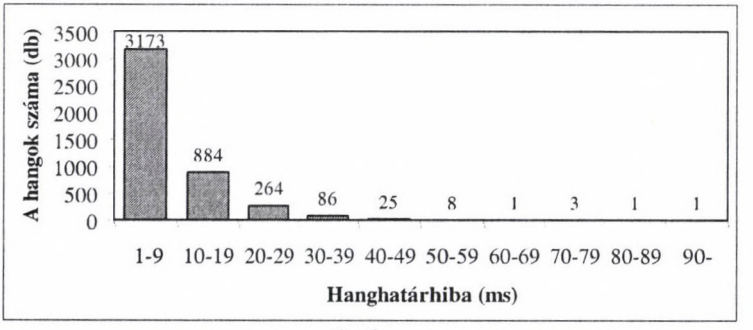
A hanghatárhibák összehasonlításából származó eredményeket beszélőfüggő tanítás során felhasznált beszédkorpusz esetén a 6., 7., beszélőadaptált esetben pedig a 8. és 9. ábrák mutatják be.



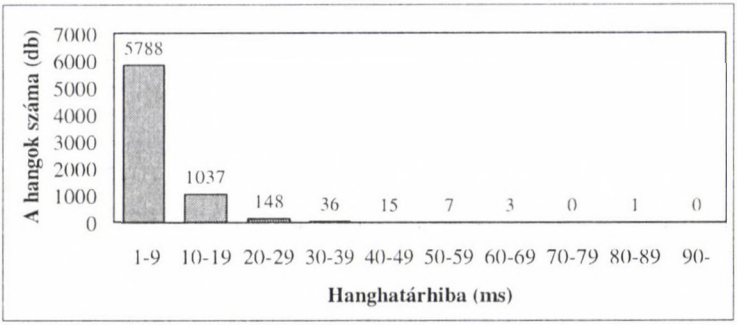
6. ábra
Az automatikus hanghatár-meghatározás pontossága BF-F1 esetben



7. ábra
Az automatikus hanghatár-meghatározás pontossága BF-N1 esetben



8. ábra
Az automatikus hanghatár-meghatározás pontossága BA-F1 esetben



9. ábra
Az automatikus hanghatár-meghatározás pontossága BA-N1 esetben

Az adaptációs beszédkorpusz esetén is megvizsgáltuk a hanghatárokat, hogy az automatikus módszer mennyire tér el a kézi javítástól. Az arányok hasonlóak voltak, mint a beszélőfüggő esetben, a hanganyag hossza miatt azonban a hibák száma jóval alacsonyabb. Az eredményeket a 7. és 8. ábra mutatja. A 2. táblázatban szereplő fonémák számát a 8. és 9. ábrákkal, illetve a 3. táblázatban feltüntetett fonémák számát a 10. és 11. ábrákkal összevetve megállapítható, hogy a beszédkorpuszban szinte az összes hanghatár módosítva lett kisebb (0,1–9 ms) és kicsit nagyobb (10–19 ms, 20–29 ms, 30–39 ms, 40–49 ms) mértékben.

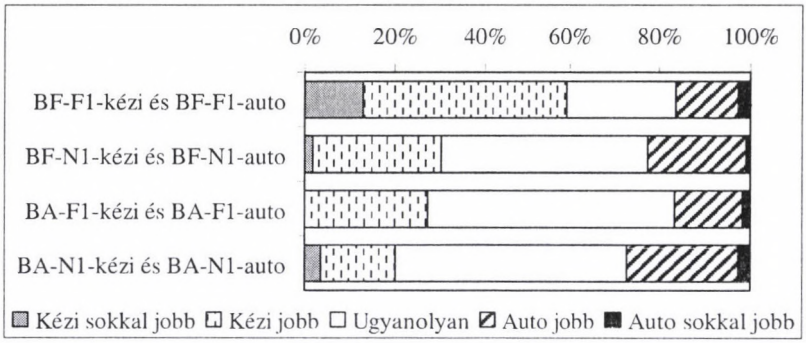
A meghallgatásos teszt eredményei

Kutatásunk célja a kialakult szövegfelolvasó rendszerek hangjának összehasonlítása volt annak a céljából, hogy megállapítsuk, hogy a kézi címkézés okoz-e minőségbeli javulást beszélőfüggő és beszélőadaptált esetekben.

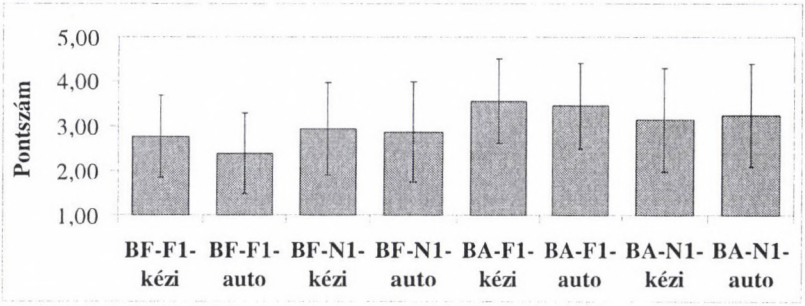
A meghallgatásos teszt eredményeit a 10. (első teszt) és a 11. ábra (második) mutatja. A 10. ábra bal oldalán jelöljük, hogy melyik rendszereket hasonlítjuk össze, a jobb oldalán pedig rendre ábrázoljuk a rendszerek összehasonlításának eredményeit (balról jobbra: az első sokkal jobb, mint a második; az első jobb, mint a második; ugyanolyan mindkét rendszer; a második jobb, mint az első; a második sokkal jobb, mint az első). Az eredmények alapján nem volt szignifikáns különbség egyik esetben sem a kézi és az automatikus címkézés között. A szignifikanciát egymintás *t*-próbával ellenőriztük. Egyedül a férfi beszélővel tanított beszélő függő rendszer esetén volt tapasztalható szignifikáns eltérés a kézi ellenőrzés javára (legalsó sor).

A meghallgatásos teszt második részében az eredeti beszélő természetes bemondásaihoz kellett a tesztalanyoknak a mintákat hasonlítaniuk. Megfigyelhető, hogy a minőség mindegyik rendszer esetén hasonló volt. Az eredményeket páronként megfigyelve (első-második, harmadik-negyedik, ötödik-hatodik, hetedik-nyolcadik oszlop) láthatjuk, hogy a természetes bemondás-

hoz képest sem volt szignifikáns eltérés a kézi és automatikus módszerek között. A szignifikanciát a várható értékre vonatkozó kétmintás párosított *t*-próbával ellenőriztük. A beszélőadaptált női hang esetén a kézi ellenőrzés minimálisan rosszabb pontokat kapott, mint az automatikus módszer. Ez a nem várt eredmény a mérést terhelő zajból származhat, nagyobb számú tesztalany esetén ez a különbség vélhetőleg megszűnne.



10. ábra
A meghallgatásos teszt első részének eredményei: páros összehasonlítás a különböző gépi rendszerek között



11. ábra
A meghallgatásos teszt második részének eredményei: páros összehasonlítás a gépi rendszerek és az eredeti beszélő között

Következtetések

Az eredmények alapján egyedül egy esetben okozott szignifikáns minőségjavulást a kézi ellenőrzés az automatikussal szemben (BF-F1), azonban a meghallgatásos teszt második része során ugyanebben az esetben már nem volt észlelhető szignifikáns minőségbeli különbség. Ez azzal magyarázható,

hogy a teszt első részében összehasonlítást végeztünk két rendszer között (relatív minősítés), ami a kis eltéréseket is jelentősen fel tudja nagyítani. A teszt második felében a rendszereket külön-külön vizsgáltuk meg (abszolút minősítés), és az így kapott eredményeket hasonlítottuk össze, és ekkor már nem mutatkozott BF-F1 esetén sem szignifikáns különbség. Az összes többi esetben (BA-F1, BF-N1, BA-N1) a kézi hanghatárjelölés és kézi fonemikus átirat nem okozott szignifikáns javulást a jelenlegi magyar nyelven működő rejtett Markov-modell alapú szövegfelolvasó rendszerben. Ezt azzal lehet magyarázni, hogy a rejtett Markov-modell alapú szövegfelolvasó generatív modelleket épít a tanító beszédkorpuszból. A beszédkorpuszban szereplő fonémák nagyszámú előfordulásának köszönhetően az automatikus eljárások által a rendszerben jelenlévő hibák a statisztikai módszerek hatására kiátlagolódnak, és így nem okoznak észlelhető minőségromlást a végső rendszerekben.

A cikkben bemutatott eredmények alapján kiderül, hogy jelenleg nem feltétlen a címkézés és a fonemikus átirat pontossága határozza meg a jó hangminőséget HMM-alapú gépi beszédkeltés esetén. Elképzelhető, hogy más, minőséget befolyásoló tényezők (pl. hangsúlyok pontos meghatározása) javítása után a fonemikus átirat és hanghatárcímkék pontossága is már szignifikáns mértékben fogja befolyásolni a gépi beszéd minőségét.

Az eredmények alapján fontos vizsgálni, hogy mi az a hibahatár, ami már jelentős minőségromlást okoz beszélőfüggő és beszélőadaptált esetben. Amennyiben a generatív modellek nagyobb hibák esetén is még megfelelő minőséget képesek produkálni, lehetséges lehet automatikus beszédfelismerő és kényszerített felismerés alapján felügyelet nélküli tanítás és beszélőadaptáció (Tóth–Németh 2011).

A HMM-alapú szövegfelolvasók minőségével a korpuszalapú elem-összefűzéses rendszerek minősége összemérhető, ezért fontos az alkalmazott módszerek elvi lehetőségét mindkét eljárás esetén megvizsgálni. A korpuszalapú elem-összefűzéses szövegfelolvasó rendszerek esetén fontos szerepet játszik a gépi beszéd minőségében a pontos fonemikus átirat és a pontos hanghatárjelölés. Az automatikus módszerekből eredő hibák az elemkiválasztás és -összefűzés során hibákat visznek be a rendszerbe, és így elkerülhetetlen, hogy amennyiben a felolvasandó szöveg a hibásan megjelölt elemek közül választ, az rosszabb minőségű gépi beszédet okoz.

Összefoglalás

Jelen cikkünkben bemutattuk a rejtett Markov-modell alapú felolvasás alapjait, és röviden ismertettük a beszélőfüggő rendszer és a beszélőadaptáció legfontosabb lépéseit. A kutatás keretein belül beszédkorpuszokat készítettünk a rejtett Markov-modell alapú szövegfelolvasó rendszerek tanításához. Elkészítettük a korpuszok automatikus címkézését, majd annak kézi ellenőrzés alapú javítását a fonemikus átiratra és a hanghatárokra vonatkozóan. Statisztikát készítettünk az automatikus eljárásokat összevetve a kézi címkézés-

sel, és nyolc különböző HMM-adatbázist hoztunk létre. Ezek alapján meghallgatásos tesztet terveztünk annak érdekében, hogy megállapítsuk, hogy okoz-e szignifikáns minőségjavulást a kézi fonemikus átirat és kézi hanghatárjelölés az automatikus módszerekkel szemben.

Eredményeink megmutatták, hogy a magyar nyelvű rejtett Markov-modell alapú szövegfelolvasó rendszerekben a kézi fonemikus átirat és hanghatárjelölés nem okoz szignifikáns minőségjavulást. Ezen eredményekre támaszkodva a jövőben a nagy mennyiségű kézi címkézés helyett használhatunk automatikus módszereket, illetve további kutatási célunk a beszédkorpusz pontossága és a HMM-alapú gépi beszéd minősége közötti összefüggések vizsgálata.

Irodalom

Black, Alan – Lenzo, Kevin 2000. Limited domain synthesis. In: *Proceedings of ICSLP*. 411–414.

Black, Alan – Zen, Heiga – Tokuda, Keiichi 2007. Statistical parametric speech synthesis. In: *Proceedings of ICASSP*. 1229–1232.

Hunt, Andrew – Black, Alan 1996. Unit selection in a concatenative speech synthesis system using a large speech database. In: *Proceedings of ICASSP*. 373–376.

Isogai, Juri – Yamagishi, Junichi – Kobayashi, Takao 2005. Model adaptation and adaptive training using ESAT algorithm for HMM-based speech synthesis. In: *Proceedings of EUROSPEECH*. 2597–2600.

Kiss Gábor – Olaszy Gábor 1982. Interaktív beszédszintetizáló rendszer számítógéppel és OVE III beszédszintetizátorral. *Magyar Fonetikai Füzetek* 10. 21–45.

Klatt, Dennis H. – Klatt, Laura C. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America* 87/2. 820–857.

Mermelstein, Paul 1973. Articulatory model for the study of speech production. *Journal of the Acoustical Society of America* 53. 1070–1082.

Möbius, Bernd 2000. Corpus-based speech synthesis: Methods and challenges. *Speech and Signals - Aspects of Speech Synthesis and Automatic Speech Recognition* 79–96.

Moulines, Eric – Charpentier, Francis 1990. Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communications* 9. 453–467.

Németh Géza – Olaszy Gábor – Fék Márk 2006. Új rendszerű, korpusz alapú gépi szövegfelolvasó fejlesztése és kísérleti eredményei. *Beszédkutatás 2006*. 183–196.

Nikléczy, Péter – Olaszy, Gábor 2004. Kempelen's speaking machine from 1791: Possibilities and limitations. (Recovering a 200 year-old technology.) *Grazer linguistische Studien* 62. 111–120.

Ogata, Katsumi – Tachibana, Makoto – Yamagishi, Junichi – Kobayashi, Takao 2006. Acoustic model training based on linear transformation and MAP modification for HSMM-based speech synthesis. In: *Proceedings of ICSLP*. 1328–1331.

Olaszy, Gábor. 1984. A phonetically based data and rule system for the real time text-to-speech synthesis of Hungarian. In: *Proceedings of the Xth International Congress of Phonetic Sciences*. Volume 2. 243–246.

Olaszy, Gábor – Németh, Géza – Olaszi, Péter – Kiss, Géza – Zainkó, Csaba – Gordos, Géza 2000. Profivox – a Hungarian TTS System for telecommunications applications. *International Journal of Speech Technology* 3–4. 201–215.

van Santen, Jan P. H. 1993. Perceptual experiments for diagnostic testing of text-to-speech systems. *Computer Speech and Language* 7. 49–100.

Tamura, Masatsune – Masuko, Takashi – Tokuda, Keiichi – Kobayashi, Takao 1998. Speaker adaptation for HMM-based speech synthesis system using MLLR. In: *Proceedings of ESCA/COCOSDA Workshop on Speech Synthesis*. 273–276.

Tokuda, Keiichi – Masuko, Takashi – Miyazaki, Noboru – Kobayashi, Takao 1999. Hidden markov models based on multi-space probability distribution for pitch pattern modeling. In: *Proceedings of ICASSP-99*. 229–232.

Tóth, Bálint – Németh, Géza 2010. Improvements of Hungarian hidden Markov model-based Text-to-Speech synthesis. *Acta Cybernetica* 19/4. 715–731.

Tóth, Bálint – Németh, Géza 2011. The effects of phoneme errors in speaker adaptation for HMM speech synthesis. In: *12th Annual Conference of the International Speech Communication Association*. 2805–2808.

Vicsi Klára – Kocsor András – Teleki Csaba – Tóth László 2004. Beszédadatbázis irodai számítógép-felhasználói környezetben. In: *Second Conference on Hungarian Computational Linguistics (MSZNY 2004)*. 315–318.

Yamagishi, Junichi – Kobayashi, Takao – Nakano, Yuji – Ogata, Katsumi – Isogai, Juri 2009. Analysis of speaker adaptation algorithms for HMM-based speech synthesis and a constrained SMAPLR adaptation algorithm. *IEEE Audio, Speech, & Language Processing* 17/1. 66–83.

Yamagishi, Junichi – Kobayashi, Takao – Renals, Steve – King, Simon – Zen, Heiga – Toda, Tomoki – Tokuda, Keiichi 2007. Improved average-voice-based speech synthesis using gender-mixed modeling and a parameter generation algorithm considering GV. In: *Proceedings of ISCA SSW6*. 125–130.

Yamagishi, Junichi – Kobayashi, Takao 2007. Average-voice-based speech synthesis using HSMM-based speaker adaptation and adaptive training. In: *IEICE Transactions on Inf. & Syst.* E90-D/2. 533–543.

Yoshimura, Takayoshi – Tokuda, Keiichi – Masuko, Takashi – Kobayashi, Takao – Kitamura, Tadashi 1999. Simultaneous modeling of spectrum, pitch and duration in HMM-based speech synthesis. In: *Proceedings of Eurospeech*. 2347–2350.

Young, Steve – Evermann, Gunnar – Gales, Mark – Hain, Thomas – Kershaw, Dan – Liu, Xunying – Moore, Gareth – Odell, Julian – Ollason, Dave – Povey, Dan – Valtchev, Valtcho – Woodland, Phil 2006. *The HTK Book, Version 3.4*. 289–292.

Zen, Heiga – Nose, Takashi – Yamagishi, Junichi – Sako, Shinji – Masuko, Takashi – Black, Alan – Tokuda, Keiichi 2007. The HMM-based speech synthesis system version 2.0. In: *Proceedings of ISCA SSW6*. 294–299.

A kutatást a BelAmi: ALAP2-00004/2005, a TÁMOP-4.2.2-08/1/KMR-2008-0007, a TÁMOP-4.2.1/B-09/1/KMR-2010-0002 és CESAR (No271022) projektek támogatták.

A PEDAGÓGUSJELÖLTEK KAZINCZY-VERSENYE

A pedagógusjelöltek Kazinczy Ferencről elnevezett szép magyar beszéd versenyének 39. országos döntője

Budapest, 2011. április 29–30.

2011 tavaszán az Eötvös Loránd Tudományegyetem Tanító- és Óvóképző Kara rendezte meg a Kazinczy Ferencről elnevezett szép magyar beszéd versenyének 39. országos döntőjét. A versenyre immár három éve jelentkezhetnek nem csupán a hagyományosan pedagógusképzést folytató felsőoktatási intézmények hallgatói, hanem a jogi és hittudományi karokon tanulók is. Ebben az évben harminckilenc versenyző nyerte el a lehetőséget – az intézményi fordulón túljutva –, hogy részt vegyen az országos döntőben. Közülük öten a határon túlról érkeztek.

A kétnapos verseny hivatalos megnyitóján a Tanító- és Óvóképző Kar dísztermében a Nemzeti Fejlesztési Minisztérium Oktatási Államtitkársága nevében Rádli Katalin főtanácsadó asszony köszöntötte a versenyzőket, a kísérő tanárokat és a szép számú érdeklődő közönséget. Ezt követően a Kar dékánja, Mikonya György szólt a hallgatósághoz, majd pedig Gósy Mária, a verseny bírálóbizottságának elnöke megnyitotta az országos döntőt. Az ünnepélyes megnyitó után a jelenlevők a magyar nyelv témájához kapcsolódó tudományos előadásokat hallgathattak meg. Elsőként Gósy Mária professzorasszony számolt be a *Beszédészlelés és helyesírás* területén végzett kutatások legújabb eredményeiről. Ezt követően Bóna Judit előadását hallgathattuk meg *Életkori sajátosságok a beszédben* címmel, majd pedig *Az agyi szótár működésé*be pillanthattunk bele Horváth Viktória előadásának köszönhetően. A negyedik – és egyben záró – előadást a Mariborból érkezett Kolláth Annától hallhattuk, aki a szlovéniai nyelvoktatásról beszélt *A muravidéki kétnyelvű oktatás rendszere és tapasztalatai a kisebbségi anyanyelv (megmaradásának) kontextusában* címmel.

A délutáni program legfontosabb részeként a szabadon választott szövegek bemutatása következett. Ahogy azt a verseny értékelésekor Gósy Mária elnök asszony kiemelte, az idén nagyon jó szövegeket hallhattunk, olyan értelemben, hogy a korábbi évek tapasztalataival összehasonlítva ezúttal sokkal inkább sikerült a versenyzőknek a személyiségükhöz és a hangjukhoz jól illeszkedő szöveget választaniuk. A bírálóbizottság tagjai pedig értékelték a teljesítményeket figyelve a tiszta beszédre, a gondos artikulációra, a jó hangsúlyokra, a hangfekvésváltások meglétére – az értő és értető felolvasás minden elemére. Ebben az évben a zsűri tagjai voltak: elnökként prof. dr. Gósy Mária (ELTE BTK, Budapest); dr. Bauko János (Konstantin Egyetem, Nyitra),

dr. Bóna Judit (ELTE BTK, Budapest), dr. Előry Vilma (Eszterházy Károly Főiskola, Eger), Hajdicsné dr. Varga Katalin (Kaposvári Egyetem), dr. Havas Judit (PIM, Budapest), dr. Horváth Viktória (MTA Nyelvtudományi Intézet, Budapest), dr. Kolláth Anna (Maribori Egyetem), dr. Merényi Hajnalka (ELTE TÓK), dr. Molnár Ildikó (Mozgássérültek Pető András Nevelőképző és Nevelő Intézete, Budapest), Nábrádiné dr. Jurasek Ibolya (PPKE Vitéz János Kar, Esztergom).

A verseny első fordulója után felüdítő kikapcsolódásként a Tanító- és Óvóképző Kar énekkara adott elő egy tavaszi kánon- és dalcsokrot. Ezt követően a versenyzők, a zsűri tagjai és a kísérő tanárok az állófogadáson zárták a napot kötetlen beszélgetéssel, ismerkedéssel.

A második versenynap délelőttjén a kötelező szöveget olvasták fel a versenyzők, Latinovits Zoltán *Monológ a költészetről* című írásának egy részletét. Ezt követően a zsűri szakmai tanácskozáson összesítette a két forduló eredményeit, és meghozta döntését. A verseny ünnepélyes eredményhirdetése ezután vehette kezdetét. A bírálóbizottság tapasztalatait Gósy Mária, a zsűri elnöke ismertette. Dicsérte a versenyzők jó szövegválasztását, ugyanakkor nem hagyta említetlenül a gyakran hallott és nehezményezett beszédhibákat sem, és felhívta a versenyzők, valamint a jelen lévő kísérő tanárok figyelmét a döntőt megelőző intézményi fordulók e szempontú rostáló szerepének fontosságára. Az értékelés után Havas Judit előadóművész felolvasta a kötelező szöveget, így hallhattuk annak egy lehetséges, mintaszerű bemutatási módját.

A zsűri döntése alapján ebben az évben a Kazinczy Alapítvány jóvoltából kiosztható 15 Kazinczy-érem közül 11-et adott át a zsűri. A 2011-ben Kazinczy-éremmel jutalmazottak névsora (a pontszámok sorrendjében): Győri Zsuzsanna (Eszterházy Károly Főiskola), Göbel Ágoston (PPKE BTK), Závogyán Renáta (SzTE Juhász Gyula Pedagógiai Kar), Ménesi Sarolt (SzTE BTK), Simon Zsuzsanna (ELTE BTK), Stein Erzsébet (ELTE PPK), Dinya Dávid (Liszt Ferenc Zeneművészeti Egyetem), Spilman Nóra (ME BTK), Fritz Gergely (SzTE BTK), Horváth Viktória (PPKE BTK), Battyányi-Nagy Annamária (Pannon Egyetem).

Az érmesek nevének ismertetése után a zsűri elnöke a Kazinczy-alapítvány nevében megköszönte az ELTE Tanító- és Óvóképző Karának a kétnapos rendezvény megszervezését, lebonyolítását. Ezt követően a rendező intézmény átadta a verseny jelképét, a Kazinczy-plakettet a jövő évi szervezőknek – a staféta a 2012. évi verseny színhelyére, Kaposvárra került.

Merényi Hajnalka

Latinovits Zoltán: Monológ a költészetről (részlet)

A verseny kötelező szövege

Aki verset mond, annak költővé kell növekednie. Azzá kell lennie, hogy eljátszhassa a költő szerepét, elmondhassa a verset. Eleinte segít a fiatalság, a frissesség, a rügyfakasztó ambíció, de miként már a virágból gyümölcs lesz, a mag terebélyes fává terpeszkedik, annál nehezebb a gyökérig menni, a költő csodáját megidézni, megélni, pedig egy verset jól elmondani, akárcsak a szavakat nem szépen, de jól ejteni, tehát igazul, ízesen, hitelesen, hittel, áttételesen, pontosan és határozottan, magamból kiszakadni, a hallgatóságba belezuhanni, a választottak megkönnyebbülése. Kivételes öröm, mintha az ember a világ egy kis pontján rendet csinálna, azt hiszem, ezt nevezik gyönyörűségnek.

Illyés Gyula írja, valaha a szövegelőadásnak is voltak Paganinijei és Liszt Ferencei, aztán sokáig nem voltak, a naturalista színjátszás űzte el őket. Színészet és előadó-művészet úgy eltávolodott egymástól, hogy szinte szólás lett: a versmondó színész feledje el színészmúltját, az előadóművésznek eszébe ne jusson, hogy színjátszás is van a földön, azaz, hogy ő is játsszon, ez természetesen tévedés. Az előadóművész is színész, s neki is játszania kell, azzal a különbséggel, hogy neki nem a szöveg cselekményét kell életre keltenie. Gondolatom az, hogy az előadóművésznek magát a költőt kell megjátszania, vagyis az alkotás néha biz hosszú folyamatát kell oly röviden összefoglalnia, mint maga a vers, ismétlem a költőt kell életre keltenie, elénk varázsolnia úgy, ahogy ő a költőt a mű első élményében, az alkotás küzdelmében elképzeli. Úgy kell elmondania, ahogy a költő szavalná a versét, ha szavalóművész is volna.

A 2011-es döntő kötelező szövegének értő-értető előadása

Egy szöveg meghangosítása előtt először mindig meg kell értenünk azt. Ez az alapja a szöveg helyes interpretációjának, a hallgató biztos megértésének. A szöveg értelmezéséhez a tartalom mellett meg kell határozni a szöveg műfaját, stílusát, hangulatát stb. is. A felolvasáskor ügyeljünk arra, hogy interpretációnk élőbeszédszerű legyen, ne váljon mesterkéltté, ne változtassuk meg természetes hangfekvésünket. Alapkövetelmény a jó beszédlégzés, a pontos, tiszta artikuláció, a hibátlan olvasás, a helyes hangsúlyozás és intonáció, a megfelelő tempóváltások és a szünettartás, illetve a szövegegész gondolatainak értő-értető megjelenítése.

Latinovits Zoltán *Monológ a költészetről* című írásának idézett részletében a versmondásról, az ahhoz szükséges képességekről, illetve a színész, az elő-

adóművész és a költő viszonyáról szól. Az idézett részlet első és utolsó mondata keretet alkot. A felolvasást, a helyes interpretációt különösen megnehezítik a hosszú, többszörösen összetett mondatok, illetve az igen sok elemből álló felsorolások.

A szöveg legfontosabb gondolata az első mondatban szerepel, amely két tagmondatból áll. Hangsúlyt kapnak a *verset* és a *költővé* szavak, amelyeket a szöveg kulcsszavainak is tekinthetünk. A második mondatban nehézséget okozhat a hangsúlyok helyes megvalósítása, a főhangsúlyokat az *azzá*, az *eljátszhassa* és az *elmondhassa* szavak kapják. A bekezdés harmadik mondatának meghangosítása fokozott figyelmet igényel: a több soron keresztül olvasható hosszú felsorolások megfelelő tagolása, illetőleg a levegővel való gazdálkodás komoly kihívást jelent a felolvasónak. A tagolásban segítségünkre lehet a szünetek megfelelő alkalmazása, szünetet tarthatunk az *ambíció, terpeszkedik, megélni, belezuhanni* szavak után. A mondat utolsó tagmondata előtti felsorolásban az artikulációs tempót is fokozhatjuk, így érzékeltetve a felsorolásban megjelenő érzelmek fokozódását is. A bekezdés utolsó mondata négy tagmondatból áll, de két fő gondolati egységre tagolódik. E két nagyobb egység között szünetet is tarthatunk (a *csinálna* szó után), az intonáció pedig az első egység végén lehet ereszkedő.

A második bekezdés első mondatát megvalósíthatjuk az előzőektől eltérő hangfekvésben, tárgyilagosabban is. A többszörösen összetett mondat tagmondatait szünettartással tagolhatjuk; a dallam az első három tagmondat végén lebegő; hangsúlyt kap *Illyés*, a *szövegelőadásnak*, a *nem* és a *naturalista*. Ezután Latinovits kifejti, hogyan különült el a színjátszás és az előadó-művészet. A kettőspont után, illetve az utolsó tagmondat előtt *(ez természetesen tévedés)* – amely a mondatnak az egész előtte álló részére vonatkozik – tartsunk szünetet. A dallam lebegő a második, a harmadik, az ötödik és az utolsó előtti tagmondat végén. A következő mondatban a hangsúlyokra figyeljünk: hangsúlyt kapnak a *színész*, a *neki*, az *azzal* (a *különbséggel* szót azonos alaphangmagassággal ejtsük) és a *nem*. A bekezdés negyedik mondata ismét többszörösen összetett, a tagolásban a vesszők segítenek. A tagmondatok végét az utolsó kivételével lebegő dallammal valósítsuk meg; az értelmi tagolást kis szünetekkel is érzékeltetve. A szöveg utolsó mondatában az első két tagmondat lebegő, az utolsó ereszkedő dallammal hangosítandó meg. Hangsúlyt kapnak az *úgy*, a *költő* és a *szavalóművész* szavak.

Végezetül mindig ügyeljünk arra, hogy egy szöveg interpretálásakor soha nem a felolvasó hangjának kell megszólalnia, hanem a szövegnek magának. Egy szöveget többféleképpen is felolvashatunk; de mindig tartsuk szem előtt, hogy a meghangosítás adekvát legyen az értelmezéssel.

Bóna Judit

„NYELVBOTLÁS"-KORPUSZ

9. rész

Az első magyar valós idejű, jegyzeteléses technikával gyűjtött megakadásjelenség-korpuszt 2004 óta adjuk közre folyóiratunkban:

Beszédkutatás 2004. 19–186.: 5139 adat;
Beszédkutatás 2005. 145–173.: 761 adat;
Beszédkutatás 2006. 231–247.: 388 adat;
Beszédkutatás 2007. 187–198.: 244 adat;
Beszédkutatás 2008. 221–239.: 444 adat;
Beszédkutatás 2009. 257–267.: 176 adat;
Beszédkutatás 2010. 283–291.: 179 adat;
Beszédkutatás 2011. 149–165.: 429 adat.

Eddig tehát összesen 7760 megakadásjelenséget elemeztünk, ebben a részben pedig újabb 298 tételt adunk közre.

Értelemszerűen ez a korpusz a hiba típusúak gyűjteménye. A kezdetektől számos, különféle szempontú tanulmány, szakdolgozat, disszertáció készült ebben a témakörben. Mindez a tapasztalat oda vezetett, hogy a „Nyelvbotlás"-korpusz osztályozási kategóriáit 2009-től egyszerűsítsük. Az alapvető ok az volt, hogy a korpuszt használóknak nagyobb szabadságot biztosítsunk egy adott jelenség mélyreható elemzéséhez.

Az egyszerűsített osztályozás elve az, hogy minden esetben a felszínen tapasztalható jelenséget soroljuk kategóriákba, a finomelemzések az adott jelenség kutatójára várnak. A fő kategóriák és meghatározásuk a következők:

1. Téves szó

A megfelelő szó helyett egy másik szó megjelenése a felszínen. Ez a kategória tartalmazza a freudi elszólásokat és a malapropizmusokat is, amelyek további elemzésekkel különíthetők el.

2. Grammatikai hiba

A köznyelvi normának ellentmondó morfológiai/szintaktikai struktúra.

3. Kontamináció

Két nyelvi jel (szó, szerkezet) összeolvadása, vegyülése.

4. Téves kezdés

Beszédhang vagy beszédhangsorozat ejtése, amely azonban nem szó.

5. Nyelvem hegyén van jelenség

A szándékolt szó jelentésének és morfolgóiai struktúrájának ismerete, de a fonológiai/fonetikai szerkezet kiejtésének (részleges) gátoltsága.

6. Perszeveráció

Egy kiejtett nyelvi elem szándéktalan, ismételt megjelenése a közlés egy későbbi időpontjában.

7. Anticipáció

Egy szándékozott nyelvi elem megjelenése a közlés egy korábbi időpontjában is.

8. Metatézis

Nyelvi elemek sorrendjének felcserélődése.

9. Egyszerű nyelvbotlás

A beszédtervezés artikulációs tervezési szintjén, ill. az artikulációban létrejött hiba, amely semmilyen egyéb okkal nem magyarázható. Altípusai: *a)* betoldás, *b)* helyettesítés, *c)* kiesés

10. Több típusba sorolható jelenségek

Azok az adatok kerülnek ebbe a kategóriákba, amelyek létrejöttében a felszíni elemzés alapján több rejtetten működő folyamat hibája is feltételezhető.

1. Téves szó

Megakadás	Szándékolt közlés
A bundáskenyeret csak bele kell mártani a tejbe ööö tojásba.	
A filológia eredete különböző pragmatikákból ered ööö paradigmákból…	
A másik gyermek esetében volt hosszabb ööö gyorsabb…	
A zöngétlen vagy a zöngekezdési idő…	
Amikor Pozsonyba megéheztünk.	megérkeztünk
…aminek azért van többletjelensége □ többletjelentése.	
Analóg szalagos beszélőkről van szó ö analóg szalagos rögzítőkről van szó.	
Az itt látható dia □ ábrából sem vonhatunk le messzemenő következtetéseket.	
Az óraátállítás nem is időtakarékos.	energiatakarékos
De ez úgyis csak egy szeánsznyi tévedés.	nüánsznyi
Ebben a rövid időintervallumban sikerült bepillantást nyernem…	nyújtanom
Ebben az étteremben értelemben…	
Ehhez tudatos, rendszeres gyakorlás születik öö szükséges.	
Én azt gondolom, és ez tényleg csak az én szuvenír véleményem…	szuverén

Megakadás	Szándékolt közlés
Ezek a tényezők függnek a beszédstílustól a beszéd ööö a beszélőtől.	
Felhívom a Zsoltit, hogy tanult-e már Rádióműsorokra.	Zsófit
Ha lesz wifi a neten □ vagy mi, a buszon...	
Ha minden köteg szakad, akkor velünk jöhetsz.	kötél
Hát hívd fel az apját □ vagyis a bátyját.	
Háztartásának anyanyelve?	házastársának
Jönnek a cafrangos foltok □ vagy mi az, rojtok.	
Különböző kontextusokban különböző dolgozatokat □ dolgokat jelenthetnek.	
Lehetnél egy kicsit talpraállabb vagy esettebb.	
Második folyadékos tanuló vagyok.	évfolyamos
Megtöltöm csokival és reklámmal □ lekvárral.	
Mi a száma □ ööö vagyis a neve?	
...minden □ mindegyikünkbe...	
Mindenki élvezze ki a négyhetes hétvégét.	négynapos
Mindenkinek megvan a saját szubjektív joga a véleményalkotásra.	szuverén
Mindez megtalálható nemzetiségi, ja, nem, nemzetségi szinten.	
Mindketten megérdemelnétek hogy elvigyétek a finálét □ a fődíjat!	
Mondjuk az a film a végétől kezdve vicces volt.	elejétől
Most, amikor ilyen meleg □ dehogyis, hideg van, én is jobban fűtök.	
Múltkor mi is láttunk egy műsort női testszépítőkről, úgy értem testépítőkről.	
Ne legyél már antiszocialista!	antiszociális
Nekünk, amikor kellett gerinc □ vagy nem gerinc, gerendagyakorlatokat végeznünk...	
Nem érzek semmit a volt barátom után.	iránt
Nemtom, hova vigyem azt a csajt megenni □ megvacsorázni.	
Olyan nehéz ezt a nyelvet kezelni □ nyelvet? □ könyvet.	
Pezsgőt is írtak vagy ittak.	

Megakadás	Szándékolt közlés
Pontosan mikor lesznek a vizsgaidőszakok □ vizsgaidőpontok?	
Realizálódik az időjárás.	normalizálódik
Sokat szoktál fut □ kocogni?	
Tehát amikor elindítom a videót □ ppt-t...	
Tehát itt hosszú ööö messzemenő következtetéseket nem tudnak levonni.	
Tényleg elég meleg ez a fagyi.	hideg
...továbbá a morális □ hoppá □ szóval nazális hangok...	
Ugyanilyen köszönet érinti ööö illeti a hallgatóinkat.	
Voltam a legjobb értelembe... a legjobb étteremben.	
Voltunk a kilátáson □ kiállításon.	

2. Grammatikai hiba

Megakadás	Szándékolt közlés
A vigasztalás mozzanata gyakorta visszatérik.	visszatér
Abból ööö azzal mosok hajat.	
Az anális fázisból a fallikus fázisban vált át.	fázisba
Az észünkkel döntöttünk.	eszünkkel
Családon belül egyre nagyobb szerepe van magának a tévében.	tévének
Csinálok a kávét, ha békén hagysz.	csinálok kávét
Elégedett lesz □ elégedetlen lesz a többi.	
Elhitte magatokról, hogy sikerül.	magukról
Én kiszálltam a mókuskerékbe.	mókuskerékből
...és a statisztikai elemzéseket is végeztünk	és statisztikai
...és bemutattak egy olyan kísérlet, amiben én is részt vettem.	kísérletet
...és ennek van négyes formája.	négy
Ettől □ eddig is lepadlóztam.	
Ez magyarázható a kis korpusz miatt.	korpusszal
Ha hazajöttelek, akkor hívunk, jaj, ha hazajöttünk, hívlak.	
Hamarosan megérkezünk az említett határátkelőhelyről.	határátkelőhelyre
...hogy ezek a könyvek népszerűvé váljon □ váljanak.	

Megakadás	Szándékolt közlés
Kell vagy kell nem?	nem kell
Kicsit sokan vagyok □ vagyunk, bocsánat.	
Le kell ülni beszélgetni abban a kis csipet-csapattal.	azzal
Mi a véleményük az eredményektől?	eredményekről
Mibe más a szokások?	miben mások
Mindegyik beszédtípusból különbözik...	beszédtípustól
Mindez a magánhangzók realizációjának □ realizációinak az eredménye.	
Nagyon jól érez érzem magam.	
Nagyon kíváncsi voltam ahhoz, hogy beszélt nyelvi anyag bevonása...	arra
Nagyon kíváncsi voltam ahhoz, hogy ez milyen új eredményeket hoz.	arra
Nem tudom, ez az ő saját kis ügyei.	ez ... ügye / ezek ... ügyei
Nincs neked a laptopnak táskád □ nincs ennek a laptopnak táskája?	
Nyilván egy büdös szót nem értenek belőle általános iskolaként □ ööö iskolásként...	
Sejtettük, hogy az adatok megváltoztatja az eredményt.	megváltoztatják
Tovább megyünk, hogy ki tudjuk a többi játékot is gyakorolni ö vagy hogy tudjuk gyakorolni.	
...vagy miről szeretnek játszani?	mivel
Vannak olyanok önök között, akik nem táplálnak Lajoshoz gyengéd érzelmeket?	Lajos iránt

3. Kontamináció

Megakadás	Szándékolt közlés
A legtöbb járművet férfiakra tervezik, abba nem gondolnak, hogy...	arra nem gondolnak × abba nem gondolnak bele
A matematikával gaj van.	baj × gond
Csak itt ücsörgéltem.	ücsörögtem × üldögéltem
...el lehet különböztetni...	meg lehet különböztetni × el lehet különíteni
Ez a jegyzet olyan össze-kusza.	össze-vissza × kesze-kusza
Ez a két jelenség nagyon szétkülönbözik eltérő beszédhelyzetekben.	szétválik × különbözik
Ezekre szeretnék részletesebben beszélni.	ezekről... beszélni × ezekre... kitérni

Megakadás	Szándékolt közlés
Ki kellett volna férni az árba.	bele kellett volna férni az árba× ki kellett volna jönni a pénzből
Különféle leshetőségek állnak fenn.	lehetőségek × eshetőségek
Mennyi az idő kint?	milyen az idő × mennyi a hőmérséklet
Miként jelentkezik meg a nyelvben...	jelenik meg × jelentkezik
Nem bonyolódnék bele a régi magyar irodalom általános kérdéseibe, mivel ahhoz nem értek hozzá annyira	ahhoz nem értek × nem értek hozzá
Nem értem, miért nem lehet őket elengeszteni.	elengedni × elereszteni
Nem kell nagy dobra fújni!	nagy dobra verni × fölfújni
Nem lehetett véghez hajtani.	véghez vinni × végrehajtani
Nem sok vizet ártok neki.	nem sok vizet zavarok × nem sokat ártok

4. Téves kezdés

Megakadás

A jelenségek megvalósula- fizikai megvalósulásai...
A másik pedig meg azt áll- ööö határozza meg...
A tu- a toleranciaküszöb nem függ az életkortól.
Az Irhás lép- ööö az Irhás árkot keresem.
...csat- mm kapcsolódó...
...D- □ Tiszántúl déli részén...
Egy mondatot kellett rekonstruáz- rekonstruálni.
...egy tévés ööö ku- tehetségkutató verseny...
Elegem van az ajtóma- automatikus ajtókból.
Elküldted a bekül- beadandót?
Én szabad ebe- embernek születtem.
Ez halászlé pontyszeletekkel. Gyorsfagyasztott két- na készétel.
Ezek elvont nyelvi em- entitások.
...hogy a ba- beszélő...
...karizmo □ karizmatikus...
Laci, megvagy a kiszo- kisszobával?
Lemegyek a koz □ közértbe.
Lexikoló- még egyszer nekifutok, lexikográfia...
Majd megcsinálja a beott beosztott.
...már mirá □ virágozik...
Mikor az Éva kiesik a f- párbaján...
Milyen beszédszava- beszédszakaszokban választja...
Most bekerült a dokto- □ az orvosi osztály ügyrendjébe.

Megakadás

Most megbeszéljük az időpontot, akinél van dá- naptár.
Most megsérü- megsértődtél rám?
Most mondtam neki, hogy kitara- kitakarítsa a szobát.
Nem fordult elő több nyelvo- nyelvbotlás a zajban.
Ő az egyetlen, akinek van határozott elha- elképzelése.
Összesen két órá- két óra negyven percnyi anyagot vettünk fel.
Persze, majd csigatermesz- öö tenyésztő leszel.
Spontán nem spi- nem sikerült.
...szimmet- szintetikus szerkesztésű...
Tehát a kettő tökéletesen ki fogja elé- egészíteni egymást.
Tele van kele- kiütéssel.
Vannak fizo- fiziológiai változások...
...vati □ vagy mi ez □ vitatott kérdés
Vigyázz, hasr- □ hanyatt ne ess!
Voltál ma az is- egyetemen?
Voltunk virágkitá- □ virágkiállításon.

5. „Nyelvem hegyén van" jelenség

Megakadás	Szándékolt közlés
Biztos úgy állítják be, mintha az is olyan a-, na, mi az, ami nincs benne a Bibliában, hogy hívják lenne, a-val kezdődik, mi az...	apokrif
Kimentem oda □ tudod □ a csirkék meg a zöldségek...	kert
Kleopátra, vagy Klütaimnésztra □ ja igen Kalüpszó volt a neve.	
Mi az a szó, segített halál?	eutanázia
Nem hoztam több □ na nem jut eszembe □ barackot.	
Otthon gyü... mi az? Tükör.	

6. Perszeveráció

Megakadás	Szándékolt közlés
A jelzett szó a jelzó után áll.	
Az anyanyelv-elsajátítás soján során alakul ki.	
Az én cipőmbe rejtették neked a coki csokit.	
Be kellene gépelnem az anyagokat, de én most virágot fogok gépelni.	

Megakadás	Szándékolt közlés
...csüngő csürgő...	
El fogsz elesni, Aliz.	el fogsz esni
Elmegy Bégybe □ Bécsbe.	
Én is szeretnék szárgabarackot!	sárgabarackot
Jó rám az a bubi □ bubi □ bugyi, amit a mama vett.	
Kapott már egyáltalán jó film Aranypálnát?	Aranypálmát
Ki kell cserélni a zuhanyrózát.	zuhanyrózsát
Kiderült, hogy az előadásom látogatásköteletes.	látogatásköteles
Kint álltam a nyugati erkényen, amikor megjött az öcsém.	erkélyen
Ma már csak a mölcsészkaron...	bölcsészkaron
Például hajnalodik pirkalodik □ ilyen szó nincs.	pirkad
Szorosz az állás.	szoros

7. Anticipáció

Megakadás	Szándékolt közlés
...a legvönyörűbb az egész világon.	
A mi szándékunkban azaz a mi lépcsőházunkban a mi szándékaink szerint nem lesz virág.	leggyönyörűbb
...beséltük a sorrendiség...	beszéltük
Csendes, zülledt az idő.	fülledt
...Dunyakanyar...	Dunakanyar
És annyira el akarok cipőt venni na □ el akarok menni cipőt venni.	
Ez összefugg függ ugye azzal...	
Ezek közül az egyik az a képesség, hogy más □ bocsánat mentális állapotokat tulajdonítsunk másoknak.	
Felpapcsolod a villanyt?	felkapcsolod
Igen, eléggé megrémült az együvege- egyliteres üvegedtől.	
Ismeretlenéknél gyakrabban fordul elő.	ismeretleneknél
Japánban összetett gyalogátlelőhelyek vannak.	gyalogátkelőhelyek
...kaleiszdoszkóp...	kaleidoszkóp
Kerestelek, de nem igazán akartad velvenni a telefont.	

Megakadás	Szándékolt közlés
Kész szerencse, hogy nem történt vonatkasztasztrófa.	
Két ruhát fizettem, és csak kétezret fizettem.	vettem
Komolyan, annyira elegem van ebbör a word helyesírás-javítóból.	ebből
Meszterszakos hallgató vagyok.	mesterszakos
Nagyon új □ nagyon jó az új hajad.	
Nem érzi magát nőnek mellettem magát.	
Rebezlit ettem reggel.	ribizlit
Szemirárium dolgozatot kell írni.	szemináriumi
Természetesen kívánni ö folytatni kívánom a munkát.	

8. Metatézis

Megakadás	Szándékolt közlés
A generatív nyelvészet a nyelvet az egyén dutásaként értelmezte.	tudásaként
A sok dekedans motívum...	dekadens
Alig van olyan nyelvreílás, ami reflektál ezekre a különbségekre.	nyelvleírás
Áttöltöm a gépeket a képre □ a képeket a gépre.	
Az Éván de megin' rusnya ruha van.	megint de rusnya ruha van
Dugva volt a villamos □ tömig.	tömve ... dugig
Elmentél vegyet jenni.	jegyet venni
Elraktad a páskádba a tènzt?	táskádba a pénzt
Én magamat akarom elhitetni azzal...	magammal akarom elhitetni azt
Észre lehet benni venne...	venni benne
Ez csak a rövid mávú teóriánkban, vagyis rövid távú memóriánkban van benne.	
Földebb a zű.	zöldebb a fű
Fúúú, amikor ott maradtunk Luigi Sport Mocskájában...	Kocsmájában
...garfika...	grafika
Ha végiggondolom az agyamnak a mélyső végségét...	végső mélységét
Heréld le a csuhádat!	cseréld le a ruhádat
Hétvégén akkor lesz vittnesz, felnesz.	fitnesz, wellness
Hogy nevet a remegéstől!	remeg a nevetéstől
Hol van a pogfiszkáló?	fogpiszkáló

Megakadás	Szándékolt közlés
Jaj, ne! Csikized a szemeddel a hajam.	hajaddal a szemem
Jól sikerült a tegnapi kikapolás □ kipakolás.	
...kutyatáp és tojókaja...	kutyakaja és tojótáp
Magával sodorva ledúzulhat a hegyoldalon.	lezúdulhat
Menjünk el az önszikolgáló étterembe!	
Mikor mostál utoljára fagot?	fogat
Mikor tudod fogni a jegyet?	fogod tudni
Neki van viszketegségi feltűnése.	feltűnési viszketegsége
Nem akrasz kimenni?	akarsz
Nem igazán akarok vele sájszebészetre menni.	szájsebészetre
Nem lehet megváltoztatni a látogatás elő-adás-kötelességét.	előadás látogatáskötelességét
Nem volt semmi konpecpi- na koncepció.	
Nézzék egy 17. ember szá 17. századi ember szemével.	
Nincsen ma csúszós kint idő?	Nincsen ma csúszós idő kint?
Olyan kis töppöm ez a lámpás.	pöttöm
Ött ó volt □ ott ő volt a legjobb.	
Teljesen fritt és fiss vagyok.	fitt és friss
Tökre nem értem, kinek mell- minek kellett ezt csinálni.	
Vegyünk milkás nyoki csuszit	csoki nyuszit
Vettem egy feteke szoknyát.	fekete

9. Egyszerű nyelvbotlás

Megakadás	Szándékolt közlés
A beszélők osztályozsása a spontán társalgásban.	osztályozása
A dolgozat pozitívusza pozitívuma a következő...	
A HÖK honlapján keresztem egy űrlapot.	kerestem
A jóta, vagyis a nóta...	
A keljfeljancit de szerettem.	keljfeljancsit
A következő héten már napszütés várható.	napsütés
Betegségükből felgyógyulva zarándoljanak zarándokoljanak el a városba.	
Egyesül vagyok.	egyedül
...elfecérelni az időt...	elfecsérelni
...értelmesen alakítsa a mondatok szószerkeit az olvasó.	szószerkezeteit

Megakadás	Szándékolt közlés
És amit már láttál, az ötvenegy néggyetménégyzetméteres lakás a Cordia felajánlásával.	
Milyen variobilitást mutat?	variabilitást
Néhány példát hadd mutattya mutassak.	
Neki nincs kopetenciája.	kompetenciája
Nem érdemes előnyben részíteni részesíteni a narratív előadást.	
Nem változtathatnak szőt szót.	
Randevű feladatok modellálása…	randevú
Tehát igazából nem érincséget téged.	érint téged
Tehát összesen három tizennégy, három tizennégy, na, háromszáztizennégy.	

10. Több típusba sorolható jelenségek

Megakadás	Szándékolt közlés
A 20. század első felét a nyelvtudományban a strukturális szemléletű irányzatok utalták… vagyis uralták.	
A mentális lexikonról már szok szó esett.	sok szó
A miénk, a Zacsi, egy zacskóku □ egy tacskóutya □ tacskókutya, na.	
A részeges kimaradás, pardon, részleges kimaradás okozhatja.	
A tizenhárom versszakos versszakok □ szótagos versszakok…	
Amit nem is lehet és nem is érdemes tekintetbe tenni.	venni
Át kell gondolnunk, hogy azok a feladatok, amik a BEA-ban volt □ vannak …audiovizá □ audiovizuálisan jól megjeleníthető.	
Az első előadás előadóihoz szeretnék kérdezni.	előadóihoz szólni × előadóitól kérdezni
Az implicit attitűdről fogok megbeszélni.	beszélni
Az irodalom, amivel az anyanyelv-elsajátítás foglalkozik…	ami az anyanyelv-elsajátítással foglalkozik
Az volt a kérésem □ az volt a kérés hozzám.	
De néznem kell a leírást, mert már régen horkoltam.	horgoltam

Megakadás	Szándékolt közlés
Dolgoztok, pihentek, mit csináltok? Adjatok magatokról életjegyet □ életjelet!	
Egy nagyon magas színvonalú beszédkonf Beszédkutatás konferenciát zárunk.	
Ehhez a munkálatokban hozzá lehetne nyílni.	nyúlni
Én majd veszek lekvárat	lekvárt
...esetről van só...	szó
És hát gyakorlatilag neki ennek a lényege...	
És nagyon jobban tanulnak nyelveket.	nagyon jól × sokkal jobban
És tök ciki volt, mert akkor szétfoslott az anyag.	szétfeslett × szétfoszlott
...etypö □ ettől egy kicsit	
Ez a konfigurácia nem felel meg a célnak.	konfiguráció
Ez mondjuk logos.	logikus × jogos
Fonjuk be egymás hasát	haját
...gyakran vonződnak a szőke nőkhez...	vonzódnak, nőkhöz
Ha kiválasztottad azt a szart, akkor most miért nem esszük meg?	vesszük
Hogyha le lehet nyomni a kollégák tortán torkán...	
Hűtsd le a kívánságodat!	kíváncsiságodat
Itt bent is nagy a baj a feszület miatt.	feszültség
Jól van a feled?	fejed
...képzettársísá □ képzettársítás...	
Kitku □ kitúrom a kukát.	
Lépj a futót ide!	lépj a futóval × tedd a futót
Még egy dolog eszembe jutott, mielőtt odaadom átadom a szót.	
Mi legyen a számonkérés módja, szemináriumi dolgozat? Esetleg valami más megszopott dolog? ööö bocsánat, megszokott.	
Miért van szükség a szétvételükre □ részvételükre?	
Most nincs idő a csücsörgésre.	csücsülésre × ücsörgésre
Nagyon ööö □ kiszabadult, vagy mi ez? Felszabadult voltam.	
Nem felejtettelek el, remélem, te is.	sem
Nem fog látszani, hogy az utolsó hétet összecsaptad?	hetet

Megakadás	Szándékolt közlés
Nem látszik, szerintem csak te látszod □ te látod.	
Nem mondom el, csak ha levágod hét fejezet.	fejemet
Nem szerették egymást, de az általuk vezetett csoportok nem álltak számban egymással.	szemben
Nem tudom, kíváncsiak vagytok rá □ Nem tudom, kíváncsiak vagytok-e rá.	
Oly bű gyotri…	bú gyötri
Otthon, a saját nyelvi közelében magabiztosan tud fellépni	közegében
Összecsengtek a kutatások az elmélet megállapátásá □ megállapításaival.	
…összeszett szavak…	összetett
Remélem, nem ez a hülye hisztéria nyeri meg.	hisztérika
Rossz a kifejtésed angolból.	kiejtésed
van egy agyéni egyéni hangszínezet	
Van egy szünet, ami némának szűnik, úgy értem tűnik.	
Van itthon szigetlépe olaj vagy mi az szigetszépe na ligetszépe olaj.	
Van nálam vásványvíz.	
Várj, még le □ be kell töltenie.	

A MAGYAR ÉS A NEMZETKÖZI TUDOMÁNYOS ÉLET ESEMÉNYEI

17th International Congress of Phonetic Sciences

Hong Kong, Kína, 2011. augusztus 17–21.

A nemzetközi fonetikai kongresszust négyévente rendezik meg, a 2011-es esemény színhelye Hong Kong volt, első ízben Ázsia. A kongresszusra jelentkezők és az elfogadott előadások száma folyamatosan növekszik, a mostani kötetében (http://www.icphs2011.hk/ICPHS_CongressProceedings.htm) 579 tanulmány szerepel: 7 plenáris, 31 tematikus, 311 szóbeli és 231 poszter.

A plenáris előadások a fonetikai kutatások széles körét fedték le. Klaus Kohler a szegmentális és szupraszegmentális variabilitásról beszélt a kommunikációs funkciók függvényében. Sara Hawkins a percepciót speciális szemszögből tárgyalta, a feladatalkalmazkodás, a jellemzőmegfigyelés és az interakciós modellek aspektusából. Jan Maddieson a nyelvek fonológiai szerveződéseit elemezte, Daniel Recasens a nyelvészeti fonetikának a jövőbeli lehetőségeiről és feladatairól, szükségszerű összefüggéseiről nyújtott összegző képet. Jacqueline Vaissière arról értekezett, hogy a magánhangzók akusztikai szerkezetének és percepciós mintázatainak egyes nyelvek közötti eltérése, avagy hasonlósága mentén hogyan értelmezhető a jonesi kardinális magánhangzók és az IPA magánhangzó-táblázata, valamint hogy mindennek hátterében milyen artikulációs viszonyok állnak. Louis Pols a beszéd dinamikájának többirányú megközelítését adta, egy-egy alkalmazás (pl. emberi vagy gépi felismerés) szempontjából integráltan tárgyalva a témát.

A tematikus szekciók hét témakör köré rendeződtek. Az első a fonológia és az aerodinamika szoros kapcsolatát hangsúlyozta. A második témakör a terepmunkának a fonetikában betöltött szerepeit és új lehetőségeit mutatta be. A további témák között szerepelt a fonetika oktatásának kérdésköre, a mintázatok rendeződése és észlelése a tonális és az intonációs vizsgálatokban, a fókusz prozódiai jellemzőinek vizsgálata, avagy a fonetika története. Az egyedi, témakörspecifikus előadások és poszterek a fonetika tudományának számos különféle témához csatlakoztak, és számtalan további kérdéskörben mutatták be a legfrissebb eredményeket. A fonetika és társtudományainak szinte minden területéről hallhattunk előadásokat. A magyar fonetikusok eredményeit négy előadás képviselte.

A gyönyörű és vengégszerető Hong Kong után a következő ülésszakot négy év múlva Glasgow-ban rendezik.

Gráczi Tekla Etelka

INTERSPEECH 2011

Firenze, Olaszország, 2011. augusztus 27–31.

Az International Speech Communication Association (ISCA) nemzetközi konferenciáját Firenzében, Olaszországban rendezték, 2011. augusztus 27. és 31. között. Az INTERSPEECH a világ számos területéről érkező, a beszédtechnológiával foglalkozó kutatóknak, oktatóknak, hallgatóknak nyújt lehetőséget arra, hogy legújabb eredményeiket bemutassák. A konferenciát minden évben más ország rendezi meg, például Japán, Anglia, Ausztrália, Belgium, USA, Portugália, Korea, Svédország.

Az INTERSPEECH történetében a 2011-es firenzei konferencián megdőlt a jelentkezők számának rekordja. Összesen 1439 tanulmányt küldtek be, ennek csupán 57%-át fogadták el szóbeli, illetve poszterelőadásra. A világ egyik legrangosabb konferenciája, hasonlóan a fonetikai világkongresszushoz (ICPhS), igen sok területet ölel fel a beszédpercepciótól a beszédszintézisen, a többcsatornás adatbázisok elemzésén, a beszédfelismerés megoldási lehetőségein, a jellemzőkinyerési technikákon át a beszélőfelismerésig és a beszélődetektálásig. A beszédtechnológiában egyre jobban előtérbe kerül az emberi társalgás gépi modellezése, a dialógusok automatikus elemzése, a beszélőváltások automatikus osztályozása. A magyar kutatásokat az MTA Nyelvtudományi Intézet, a BME TMIT és az MTA Pszichológiai Kutatóintézet munkatársai mutatták be.

A 2012-es INTERSPEECH konferencia az USA-ban, Portlandben, a 2013-as pedig Franciaországban, Lyonban lesz.

Beke András

BESZÉDKUTATÁS 2011 konferencia
Spontán beszéd: elemzések, adatbázisok, alkalmazások

Budapest, 2011. október 27–28.

2011 októberében immáron tizedik alkalommal került megrendezésre a Beszédkutatás konferencia az MTA Nyelvtudományi Intézet nagyelőadójában. Az ünnepi alkalomhoz méltó módon kiemelkedően színes és magas színvonalú program várta a hallgatóságot. Ezúttal a spontán beszéd állt a kétnapos előadássorozat középpontjában. 32 szóbeli előadás hangzott el, továbbá 3 poszter került bemutatásra.

Bánréti Zoltánnak, az MTA Nyelvtudományi Intézet igazgatóhelyettesének megnyitója után az első nap plenáris előadását Bunta Ferenc (Houstoni Egyetem) tartotta a kétnyelvű gyermekek fonológiai fejlődéséről. Az előadás során

ízelítőt kaphattunk az előadó és munkatársai által használt módszerekből, és bemutatásra kerültek a legfrissebb eredmények is. Kutatásaik fontos eredményeként Bunta Ferenc kiemelte a kétnyelvű gyermekeknek az egyes nyelvelsajátítási folyamatokban tapasztalható előnyét. Az első napi plenáris témájához további előadások is kapcsolódtak a kétnyelvűség kérdésének sok szempontú kutatásában. Simon Orsolya kisiskolásokkal végzett szófelismerési kísérleteket, míg Navracsics Judit és Sáry Gyula a nyelvhelyességi döntéseket helyezte kutatása középpontjába. Az előadások során sok szó esett a kétnyelvűséggel járó előnyökről, továbbá a hallgatóság megismerkedhetett a kétnyelvű beszédfolyamatok bonyolult működéseivel is.

A második napon a konferencia szervezője, Gósy Mária (MTA Nyelvtudományi Intézet) és Bóna Judit (ELTE) közös nagyelőadását hallgathattuk meg. A dadogással foglalkozó esettanulmány olyan részletes elemzéseket közölt, amelyekkel a dadogás és a beszédhelyzet összefüggései váltak adatolhatóvá. Az előadók megállapították, hogy a beszédtervezési működések eltérései miatt az olvasás, a mondatismétlés és a narratíva jelentősen elkülönül az interpretált beszédtől és a társalgástól. Eredményeik gyakorlati szempontból is fontosak, és hozzájárulnak a dadogók motoros kontrolljának és tervezési folyamatainak jobb megismeréséhez, ezáltal pedig akár a fejlesztésben is segítséget nyújthatnak. Az itt elhangzottakhoz kapcsolódott – az alkalmazás oldaláról – a dadogásterápia egy igen érdekes, hazánkban még nem túlzottan elterjedt terápiás technikájáról szóló beszámoló. Gerliczkiné Schéder Veronika és Vári Nikolett (Nyíregyházi Főiskola) előadásában bemutatta a lovasterápia mint módszer tudományos hátterét, gyakorlati alkalmazását, és egy hatéves kisfiú esetén keresztül a terápia hatékonyságát is demonstrálták. Gyarmathy Dorottya (MTA Nyelvtudományi Intézet) az *ízé* akusztikai megvalósulásait elemezve vizsgálta a töltelékszó nyelvi szerepeit. A pszicholingvisztika és a mérnöki alkalmazások együttműködéséről számolt be Beke András (ELTE) és Gyarmathy Dorottya (MTA Nyelvtudományi Intézet) közös munkája. Az ismétlések és újraindítások gépi osztályozására tett kísérlettel a szerzők a nyelvészet eredményeit felhasználva a spontán beszéd mesterséges felismerésének javításához kívántak hozzájárulni.

A konferencia két napján többek között számos, a beszéd szegmentális szerkezetét, a beszédhangok szintjét vizsgáló kutatást is meghallgathattak az érdeklődők. Markó Alexandra (ELTE) az irreguláris zönge határjelölő funkciójának hiánypótló vizsgálatáról számolt be. Eredményei szerint a glottalizáció megjelenésében elsősorban a prozódiai struktúra és az abban betöltött pozíció játszik szerepet, de a vizsgált jelenség erősen egyéni ejtési sajátosság. Bóna Judit (ELTE) az idősek és a fiatalok spontán beszédprodukciójának összevetését végezte el a nyelvi rövid-hosszú fonológiai oppozíció szempontjából. Amint sok másban is, az idősek és a fiatalok beszéde ebben a paraméterben is eltérni látszik, hiszen úgy tűnik, az idősek lassabb artikulációs tempójuk következtében kisebb spektrális különbséggel valósítják meg a vizsgált

fonológiai párok tagjait, ugyanakkor a fizikai időtartam tekintetében ők is elkülönítik a vizsgált magánhangzópárok tagjait. Mády Katalin (MTA Nyelvtudományi Intézet) szintén a rövid-hosszú oppozíció vizsgálatával a nyelvi redukció mint nyelvi változás kérdését taglalta. Eredményei szerint az alacsonyabb nyelvállásfokú magánhangzók az egyre csökkenő különbségek ellenére továbbra is markánsabb elkülönülést mutatnak a felső nyelvállásúakhoz képest. Hiánypótló statisztikai adattár készült Beke András (ELTE), Gósy Mária és Horváth Viktória (MTA Nyelvtudományi Intézet) munkája során 80 magyar anyanyelvű beszélő magánhangzó- és mássalhangzó-fonémáinak elemzésével. Az egyedülálló eredmény számos felhasználási lehetőséget rejt magában nemcsak a nyelvészeti, de a mérnöki alkalmazások számára is. A spontán beszéd és az olvasás beszédhang-realizációinak összevetését végezte el Deme Andrea (ELTE), Gráczi Tekla Etelka, Horváth Viktória (MTA Nyelvtudományi Intézet) és Markó Alexandra (ELTE). Eddigi eredményeik arra hívják fel a figyelmet, hogy további vizsgálatot érdemel az a szakirodalomban elterjedt nézet, amely szerint a spontán beszédre az alulkonfiguráltság, míg a felolvasásra a pontosabb artikuláció a jellemző. A beszédhangok törlődését elemezte Váradi Viola (ELTE) a spontán beszédben. Az előadó részletesen bemutatta a mássalhangzóknak a magánhangzókhoz viszonyított gyakoribb törlődésének tendenciáit, továbbá ezek összefüggéseit az egyes szupraszegmentális tényezőkkel. További vizsgálatot érdemlő konklúzióként elhangzott, hogy a fonológiai jelenségeket a fonetikai megvalósulások nem támasztották alá minden esetben. S. Tar Éva (ELTE) a fonológiai fejlődésről értekezett gyermekek beszédprodukcióját vizsgáló kutatása kapcsán. Eredményei alapján azt a következtetést fogalmazhatta meg, hogy bár a kezdetekkor tapasztalhatók különbségek a gyermekek nyelvfejlődésében a nemek tekintetében, de ezek valamivel hároméves kor után gyakorlatilag teljesen eltűnnek: nincs tehát szó valós elmaradásról, inkább csak a fejlődés tempójában megmutatkozó különbségekről.

A pragmatika tárgykörében is számos előadás hangzott el. Dér Csilla Ilona (KGRE) a diskurzusszervezést vizsgálva a beszélőváltás szerepű diskurzusjelölőkkel foglalkozott. Hasonló témában hangzott el a konferencia egyetlen hallgatói előadása Szabó Ágnes és Tóth Andrea (ELTE, mesterszak) bemutatásában is. Ők a gyermekek társalgásában megfigyelhető beszélőváltásokat, azok fonetikai paramétereit elemezték. A diskurzusjelölőket választotta témául Schirm Anita (SzTE) is. Az előadó a beszélt nyelvi és az írott nyelvi különbségeket próbálta feltérképezni a *szóval* és a *hát* szavak funkciói alapján. Németh Zsuzsanna (SzTE) a spontán párbeszédekben fellépő hibajavítási mechanizmusok működését elemezte. Erdős Klaudia (ELTE) a humor és ezen belül is elsősorban az irónia artikulációs megjelenéséről beszélt. Az előadó azokat a prozódiában megjelenő kulcsokat mutatta be, amelyek feltehetőleg segítik a hallgatót a humor felismerésében. Nemesi Attila László (PPKE) a verbális arculatmunkát, azon belül is főként az implikatúra közmédiabeli

megjelenését elemezte. A vizsgálat középpontjába a meggyőzés, a benyomáskeltés, az udvariasság és az implikált jelentések kerültek.

A beszéd szupraszegmenális szintje több szempont szerint került terítékre az ülésszak során. Neuberger Tilda (ELTE) a gyermekek spontán beszédének tagolását a beszéd grammatikai komplexitásával összefüggésben vizsgálta. A kutatás azt látszik bizonyítani, hogy a komplexitás mértékének fokozódásával csökken az egyöntetű tagolás lehetősége, ugyanakkor a gyermekek és felnőttek által használt tagoló funkciójú elemek nagyrészt megegyeznek. Menyhárt Krisztina (MTA Nyelvtudományi Intézet) szociolingvisztikai szemléletű elemzésében az '50-es évek gyermekbeszélőinek beszédsajátosságairól számolt be a Hegedűs-archívum felvételei alapján. Adataira a gyermekeknek az '50-es években jelentkező szociológiai státuszváltozásában talált feltételezhető magyarázatot.

Fontos előrelépésekről számoltak be az előadók az új adatbázisok és beszédkorpuszok építése és a már meglévők egyre szélesebb körű felhasználhatósága kapcsán. Imre Angéla (ELTE) egy régóta időszerű gyermeknyelvi adatbázis megalkotására tett első lépéseket mutatta be. Szabó Tamás Péter (MTA Nyelvtudományi Intézet) egy korszerűen feldolgozott, nyelvi és metanyelvi kutatást egyaránt lehetővé tevő, szintén gyermeknyelvi adattárat ismertetett. Olaszy Gábor (BME), Abari Kálmán (DE) és Rácz Zsuzsanna (BME) egy magyar nyelvű formánsadatbázis alapján készített átfogó elemzést a magyar magánhangzók formánsmozgásairól. Varga Zoltán (Nemzetbiztonsági Szakszolgálat) a nyelvi adatbázisoknak a kriminalisztikai alkalmazásokban betöltött szerepéről beszélt, és hangsúlyozta a nyelvészeti kutatások és a bűnüldözés egymást segítő tendenciáit. Az előadó rámutatott arra is, hogy a konferenciát szervező Kempelen Farkas Beszédkutató Laboratóriumban évek óta fejlesztett BEA beszédadatbázis használata hatalmas előrelépést jelentett a kriminológusok munkájában. Szintén a beszélőazonosítás problémájának megoldásához járult hozzá Tatár Zoltán (Nemzetbiztonsági Szakszolgálat) kutatása, amelyben az akcentusjelenségek felismerésének lehetőségét vizsgálta.

A műszaki alkalmazások oldaláról meghallgathattuk Tóth Bálint és Németh Géza (BME) előadását, amely rejtett Markov-modell alapú beszélőadaptált beszédszintetizátor fejlesztését mutatta be. Az eddig elsősorban a beszédfelismerésben használt eljárás újításként hat a beszédszintézisben, az előadók fejlesztése Magyarországon eleddig egyedülállónak számít. Szaszák György (BME) Beke Andrással (ELTE) a mesterséges beszédfelismerés hatékonyságát prozódiaalapú szintaktikai elemző segítségével kívánta javítani. A kísérlet ígéretesnek mutatkozott mind a beszédtechnológiai alkalmazások, mind pedig a kutatási célú beszédelemzés automatizálása szempontjából.

Az orvostudomány és a nyelvészeti kutatások együttműködésének eredményei is bemutatásra kerültek. Götz Gergely a fogorvoslásban kamatoztatta a fonetikai elemzési technikákat. Svindt Veronika az afáziás betegek nyelvi és

nem nyelvi kommunikációs stratégiáit vizsgálta. A széleskörű tematikai változatosságot példázza, hogy a második napon egy a tanári magyarázatról szóló módszertani előadás is elhangzott Antalné Szabó Ágnestől (ELTE).

A poszterszekcióban egyetemi hallgatók és doktoranduszok mutathatták be munkájukat. Gál Georgina (ELTE) életkori nyelvhasználati jellegzetességeket vizsgált. Mészáros Katalin Edit (ELTE, mesterszak) azt elemezte, milyen paraméterek mellett tekinthető megakadásjelenségnek a néma szünet. Krepsz Valéria (ELTE, mesterszak) pedig a nők és férfiak közti nyelvi különbségek szocializációs beágyazottságáról beszélt.

Az első napot ünnepi fogadás zárta, melyen a konferencia – Bánréti Zoltán igazgatóhelyettes szavaival élve – „keresztanyja", Gósy Mária mondott köszönetet az elmúlt évek szervezésben nyújtott segítségért munkatársainak, egyetemi hallgatóinak, továbbá sommázva az elmúlt évek tapasztalatait örömmel nyugtázta, hogy a konferencia a beszédkutatás értékes fórumává lett.

A Beszédkutatás konferencia tízéves történetében rengeteg jelentős előadás hangzott el a beszédkutatás számos területéről. Az ülésszak kimagasló erénye, hogy résztvevői között sosem csak nyelvészek kaptak és kapnak helyet, továbbá nem csak nyelvészeti szemléletmódok érvényesülnek. A minden második évben megrendezett konferencián ugyanis mindig nagy hangsúly helyeződik arra, hogy az előadások a nyelvészettől távolabb álló, a nyelvészek számára sokszor ismeretlen, ugyanakkor igen hasznos diszciplínákat és azok eredményeit is megismertessék a szakmai és laikus közönséggel egyaránt. A konferencia célja továbbá, hogy az elméleti kérdések vizsgálatán felül olyan alkalmazási lehetőségeket és területeket is bemutasson, melyek képesek lehetnek építő kapcsolatot teremteni a gyakorlati felhasználás és az elméleti kutatások között. Ennek szellemében nemcsak kutatók, de gyakorló orvosok, mérnökök, pszichológusok, tanárok és további szakemberek is szót és teret kapnak munkájának bemutatására. Az előadók között évek óta képviselteti magát a Nyelvtudományi Intézet több osztálya, az ELTE, a BME, a SOTE, a Károli Gáspár Református Egyetem, de Pécsről, Debrecenből, Szegedről, Veszprémből is minden alkalommal köszönthetünk résztvevőket.

A Beszédkutatás konferenciákon felvetődő témák híven tükrözik a hazai és a nemzetközi tudományos élet fejlődését. A program összeállításának és az előadások kiválogatásának szigorú követelményei pedig lehetővé teszik azt is, hogy csak igazán színvonalas és a nemzetközi eredményekkel összevetésben értelmezett kutatások kerüljenek a hallgatóság elé. A Kempelen Farkas Beszédkutató Laboratórium munkatársai Gósy Mária irányításával két év múlva ismét megszervezik a Beszédkutatás konferenciát. A jubileumi alkalom elmúltával pedig nem is kívánhatunk mást, mint további kitartást és erőt a szervezőknek, hogy az ülések az eddigiek sikereihez méltó módon folytatódhassanak a jövőben is.

Deme Andrea

A KÖTET SZERZŐI

Név	Intézmény	E-mail
Abari Kálmán	Debreceni Egyetem Pszichológiai Intézet	abari.kalman@arts.unideb.hu
Beke András	ELTE Fonetikai Tanszék	beke.andras@gmail.com
Bóna Judit	ELTE Fonetikai Tanszék	bona.judit@btk.elte.hu
Bunta, Ferenc	University of Houston, Houston, Texas	fbunta@uh.edu
Deme Andrea	ELTE Fonetikai Tanszék	andrea_deme@hotmail.com
Dér Csilla Ilona	KGRE Magyar Nyelvtudományi Tanszék	csillader@gmail.com
Gósy Mária	MTA Nyelvtudományi Intézet	gosy.maria@nytud.mta.hu
Gráczi Tekla Etelka	MTA Nyelvtudományi Intézet	graczi.tekla.etelka@nytud.mta.hu
Gyarmathy Dorottya	MTA Nyelvtudományi Intézet	gyarmathy.dorottya@nytud.mta.hu
Horváth Viktória	MTA Nyelvtudományi Intézet	horvath.viktoria@nytud.mta.hu
Markó Alexandra	ELTE Fonetikai Tanszék	marko.alexandra@btk.elte.hu
Menyhárt Krisztina	MTA Nyelvtudományi Intézet	menyhart@nytud.hu
Merényi Hajnalka	ELTE TÓK Magyar Nyelvi és Irodalmi Tanszék	mhajnal@yahoo.com
Nemesi Attila László	PPKE Kommunikációtudományi Tanszék	nemesi.attila@btk.ppke.hu
Németh Géza	BME Távközlési és Médiainfomatikai Tanszék	nemeth@tmit.bme.hu
Németh Zsuzsanna	SZTE Nyelvtudományi Doktori Iskola	nemethzs7@gmail.com
Neuberger Tilda	MTA Nyelvtudományi Intézet	neuberger.tilda@nytud.mta.hu
Olaszy Gábor	BME Távközlési és Médiainfomatikai Tanszék	olaszy@tmit.bme.hu
Schirm Anita	SZTE BTK Magyar Nyelvészeti Tanszék	schirmanita@gmail.com
Szabó Ágnes	ELTE Fonetikai Tanszék	agicaa89@gmail.com

Név	Intézmény	E-mail
Tóth Andrea	ELTE Fonetikai Tanszék	toandi@mailbox.hu
Tóth Bálint	BME Távközlési és Médiainfomatikai Tanszék	toth.b@tmit.bme.hu
Váradi Viola	ELTE Fonetikai Tanszék	varadiviola@gmail.com