BESZÉDKUTATÁS 2013

BUDAPEST

BESZÉDKUTATÁS 2013

Szerkesztette:
Gósy Mária

MTA Nyelvtudományi Intézet
Kempelen Farkas Beszédkutató Laboratórium
Budapest, 2013

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

Az egyes tanulmányokat szakterületi kompetenciával rendelkező szakemberek lektorálták.

Technikai szerkesztés: Markó Alexandra, Gráczi Tekla Etelka

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.



Olaszy Gábor 70 éves

Tíz évvel ezelőtt adtuk hírül a *Beszédkutatás*ban, hogy kollégánk, Olaszy Gábor 60 éves lett. Szinte nem is hihető, hogy eltelt ez az évtized, és immár a hetvenedik születésnapján köszöntjük. Hihetetlen, hiszen Gábor ugyanolyan töretlenül dolgozik, kutat, alkot, mint az elmúlt évtizedekben mindig. Hihetetlen, mert sem a munkatempóján, sem a megjelenésén nem hagytak nyomot az évek. Számos publikációval gazdagította a magyar fonetikát, újabb és újabb gondolatok, tudományos megoldások, fejlesztések jelentek meg a munkáiban. A Nyelvtudományi Intézet Fonetikai Osztályának munkatársaival szoros a kapcsolata, csakúgy, mint a beszédkutatással más munkahelyeken foglalkozó kollégákkal (ELTE Fonetikai Tanszék, Budapesti Műszaki Egyetem). Folyamatosan részt vesz a tudományos eseményeken, segít a fiataloknak az eligazodásban; mindig, minden helyzetben, mindannyian számíthatunk rá.

Méltán mondhatjuk, hogy Olaszy Gábor mintegy négy évtizede a magyar beszédtudomány meghatározó alakja. A folytatáshoz valamennyien jó egészséget, munkakedvet és sok-sok örömöt kívánunk nagy szeretettel!

TARTALOM

Gráczi Tekla Etelka – Siptár Péter: Degemináció? 7

Váradi Viola – Beke András: Az artikulációs tempó variabilitása a felolvasásban 26

Markó Alexandra: Kérdő megnyilatkozások a spontán beszédben 42

Bóna Judit: A beszédszünetek fonetikai sajátosságai a beszédtípus függvényében 60

Baditzné Pálvölgyi Kata – Font-Rotchés, Dolors: The intonation of Spanish yes-no questions in spontaneous speech and in a didactic material 76

Gráczi Tekla Etelka: Explozívák és affrikáták zöngésségének időviszonyai 94

Gósy Mária – Bóna Judit – Beke András – Horváth Viktória: A kitöltött szünetek fonetikai sajátosságai az életkor függvényében 121

Horváth Viktória: Temporális szerveződés kilencéves gyermekek spontán beszédében 144

Neuberger Tilda – Gráczi Tekla Etelka: Az alveoláris zöngétlen explozíva variabilitása 160

Gósy Mária – Gyarmathy Dorottya: A progresszív afázia akusztikai-fonetikai vonatkozásai (esettanulmány) 173

S. Tar Éva: A zöngésségi kontraszt elsajátításának mintázata atipikus nyelvfejlődés esetén 194

Krepsz Valéria: Koraszülött gyermekek nyelvhasználati jellemzőiről 210

Bereznai Anja: Sprachperzeption im bilingualen Kindergarten 218

Bartha Krisztina: A beszédfeldolgozás jellemzői magyar domináns kétnyelvűeknél 233

Kiss Gábor – Vicsi Klára: Akusztikai hangosztályok felismerésén alapuló, nemlineáris idővetemítés megvalósítása a mondathanglejtés és a szóhangsúlyozás oktatásához 247

Olaszy Gábor: Precíziós, párhuzamos, magyar beszédadatbázis fejlesztése és szolgáltatásai 261

Götz Gergely: A fogszabályozás és a fogszabályozó készülékek hatása a hangképzésre 271

A pedagógusjelöltek Kazinczy-versenye 278
Fóris Ágota: A felsőoktatási intézmények Kazinczy Ferenc Szép magyar beszéd versenyének 40. országos döntője 278

A magyar és a nemzetközi tudományos élet eseményei 280
Beszédtudomány. Az anyanyelv-elsajátítástól a zöngekezdési időig • *Németh Mátyás* 280
14. Nemzetközi Klinikai Fonetikai és Nyelvészeti Konferencia • *Tar Éva* 282
Workshop Fluent Speech • *Bóna Judit – Váradi Viola* 284

A kötet szerzői 286

DEGEMINÁCIÓ?

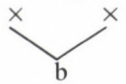
Gráczi Tekla Etelka – Siptár Péter

Bevezetés

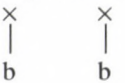
Közkeletű vélekedés a magyar mássalhangzó-hosszúság felszíni eloszlásáról, hogy szó elején vagy egy másik mássalhangzó előtt vagy után soha nem fordul elő gemináta, vagyis hosszú mássalhangzó csak intervokális helyzetben vagy megnyilatkozás végén állhat. Ez a hagyományos nézet azonban túlságosan leegyszerűsített (l. Siptár 2001, 2012; Polgárdi 2008; Pycha 2009, 2010; a geminátákról általában: Davis 2011; Ringen–Vago 2011). Ahhoz, hogy legalább a megfigyelés szintjén helytálló legyen, többféleképpen is pontosítanunk kell. Ez a pontosítás és a jelenség fonetikai és fonológiai elemzése a jelen tanulmány célja.

A magyar degeminációt (mássalhangzó-rövidülést) tárgyaló cikkében Nádasdy (1989) különbséget tesz **mögöttes** (a szótárban is ekként szereplő) és **levezetett** (a fonológiai levezetés során előálló), illetve **balról** és **jobbról szegélyezett** geminátáknak, a levezetetteken belül pedig **valódi** és **álgeminátáknak** között: az előbbiek hasonulás révén jönnek létre, pl. *lábbal*, és ugyanúgy elágazó szerkezetűként ábrázolhatók, mint a mögöttes geminátáknak, l. (1a), míg az utóbbiak (morfémahatár két oldalán álló) azonos rövid mássalhangzók egymásutánjai, pl. *lábból*, l. (1b), és a valódi geminátáktól fonetikailag megkülönböztethetetlen felszíni megvalósításuk az **ikeredés**nek nevezett művelet révén alakul ki.

(1a) Valódi gemináta



b) Álgemináta

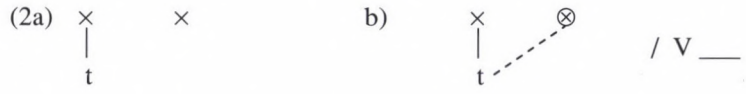


c) Rövid mássalhangzó

Mögöttes geminátáknak

A jobbról szegélyezett mögöttes geminátáknak (pl. *jo**bb**ra* [jobrɔ]) többékevésbé az imént körvonalazott módon viselkednek (vagyis szabályosan degeminálódnak), azzal a megszorítással, hogy szóhatáron keresztül (pl. *jo**bb** lesz* [job(ː)lɛsː]) a degeminációt **fakultatív:** a beszédstílus és a határ jellege szerint változó gyakoriságot mutat (Dressler–Siptár 1989, 1998: 40–41; Ács–Siptár 1994: 576–577). Minél formálisabb a regiszter és minél erősebb a gemináta meg a szegélyező mássalhangzó közötti morfológiai/szintaktikai határ, annál kevésbé valószínű a degeminációt.

Balról szegélyezett mögöttes geminátáknál általában nem fordulnak elő, hisz ehhez az kellene, hogy valamely morféma hosszú mássalhangzóval kezdődjék. Van azonban egy lehetséges jelölt egyetlen geminátából álló morfémára: a múlt idő *-tt* jele (Vago 1980; Stiebels–Wunderlich 1999; Trón–Rebrus 2005). Erről Siptár és Törkenczy (2000: 249–261) elemzése azt állítja, hogy „degeminációs hatást" mutat ugyan, de anélkül, hogy valójában degemináción esne át: a toldalék szótári ábrázolása egy olyan /t/-ből áll, amely egyetlen időzítési egységhez (×-hez) van társítva, ám ez utóbbit üres (szegmentális tartalommal nem bíró) időzítési egység követi, l. (2a). Az elemzés feltételez egy *t*-terjedési szabályt, amely akkor és csak akkor alkalmazódik, ha a (2a) konfigurációt magánhangzó előzi meg, l. (2b).

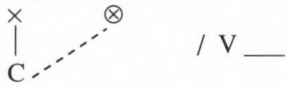


Ezek szerint egy olyan alakban, mint például *fél-**t,*** nincs semmi, ami degeminálódhatna. (Az üres ×, amelyre tehát mássalhangzó után nem tud ráterjedni a *t*, a fonetikai megvalósítás számára hozzáférhetetlen marad, vagyis nincs hatással az alak kiejtésére.) Ezzel szemben a *lát-**ott,** ad-ott* típusú esetekben végbemegy a (2b) szabály, és levezetett valódi geminátát hoz létre. (Ha egy ilyen alakot mássalhangzó kezdetű szó követ, jobbról szegélyezett levezetett valódi gemináta esete áll elő, és ugyanaz történik, mint más hasonló esetekben, l. alább az (5) szabályt.)

Levezetett valódi geminátáknál

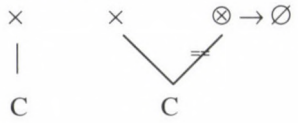
Egy másik toldaléktípus, amelynek tagjai a nélkül mutatnak degeminációs hatást, hogy szükségképpen degemináció menne végbe rajtuk, az ún. váltakozó *v*-kezdetű toldalékok csoportja (az eszközhatározói *-val/vel* rag: *só-val, méz-zel,* és a transzlatívuszi *-vá/vé*: *só-vá, méz-zé*). Ez a típus (ha degemináció volna benne) a balról szegélyezett levezetett valódi geminátáknál példája lenne: *domb-bal* [mb] (*[mb:]), *vers-sel* [rʃ] (*[rʃ:]), *lánc-cá* [n͡ts] (*[n͡ts:]), *férj-jé* [rj] (*[rj:]). Siptár–Törkenczy (2000: 269–274) erre az esetre adott elemzése a *t*-terjedés szabályának általánosításán alapul, l. (3). Az általánosított szabály pl. a *méz-zel* alakban működésbe lép, de pl. a *domb-bal* alakban nem, miáltal minden további nélkül létrejön a kívánt degeminációs hatás.

(3) C-terjedés



Siptár (2001) számos további degeminációs(nak látszó) folyamatot elemez a fentihez hasonló módon, vagyis a gemináció elmaradásaként (korlátozásaként). Vannak azonban olyan esetek is, amelyek – úgy tűnik – semmiképpen nem elemezhetők degeminációs szabály nélkül. Például a zöngésségi hasonulás kimenete adott esetben gemináta lesz (mégpedig akkor, ha előzőleg a két szegmentum csupán a zöngésség tekintetében különbözött egymástól, pl. *adta* [ɔtːɔ]; egyébként nem jön létre gemináta: *dobta* [doptɔ]). Az *adta* esetében bekövetkező KKE-hatás, mivel nem szabály, hanem automatikus konvenció eredménye (a Kötelező Kontúr Elvéről l. Durand–Siptár 1997: 144–146), nem írható le olyan terjedési műveletként, amely csak magánhangzó után megy végbe. Így a *küldte* [kyltɛ] típusú esetekhez (/ldt/ → ltː → [lt]) külön degeminációs szabályt kell bevezetnünk. Informálisan (4) szerint írhatjuk fel ezt a szabályt:

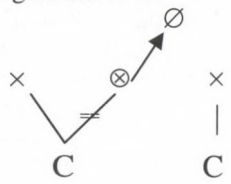
(4) Degemináció I



A (4) szabály lexikális (szószintű) szabály, ami egyebek között annyit jelent, hogy csakis szóalakon belül alkalmazható, ott viszont szigorúan kötelező. Látszólag ugyan valami hasonló történik a *küld tehát, volt dolga* típusú esetekben, vagyis akkor, ha a zöngésségi hasonulás révén létrejövő levezetett valódi gemináta két fele szóhatárt fog közre. Valójában azonban ez utóbbi típusú degemináció egyrészt fakultatív (szabadon választott), másrészt **fokozatos** jellegű (azaz nem igen/nem, hanem többé/kevésbé típusú), harmadrészt pedig érzékeny a szegélyező mássalhangzó minőségére (a közkeletű feltételezés szerint bekövetkezése valószínűbb, hatása pedig teljesebb akkor, ha a szegélyező mássalhangzó zörejhang, pl. *hozd tehát,* mint akkor, ha likvida, pl. *hord tehát*). Egyszóval ez a típus teljesen az álgeminátákhoz hasonlóan viselkedik (ez utóbbiakról l. Siptár 2001: 299–301).

A jobbról szegélyezett valódi geminátáknak esetében Siptár (2001) elemzése szerint a (4) szabály fordítottjára (tükörképére) van szükség – legalábbis a jobbról szegélyezett **mögöttes** geminátáknak (pl. *hall-gat* [hɔlgɔt]), továbbá az előző alpont végén már említett típusú jobbról szegélyezett levezetett geminátáknak (pl. *hallott felőle* [hɔlːotfɛlølɛ]) kezeléséhez. Ezért a következő degeminációs szabályt állíthatjuk fel az összes jobbról szegélyezett valódi gemináta kezelésére (akár mögöttes, akár levezetett geminátáról van szó):

(5) Degemináció II



A (4) és (5) szabályt akár össze is vonhatnánk egyetlen tükörképszabállyá („Töröld a valódi geminátá k két ×-e közül az egyiket, ha még egy mássalhangzó **előzi meg vagy követi** őket"). De ez nem lenne helyes: a (4) szabály csak szón belül alkalmazható, és mindig szigorúan kötelező, míg az (5) a nyelvtan posztlexikális fonológiai összetevőjében lép működésbe, és a beszédstílus (a fesztelenség szintje) meg a beszédtempó, illetve a határerősség fokozatainak megfelelően változó valószínűséggel megy végbe (Siptár 2001). Ezenkívül az érintett morfémák vagy szóalakok előfordulási gyakorisága is befolyásolhatja a szabály alkalmazásának valószínűségét.

A geminátá k időtartama és belső időszerkezete (meg a degemináció)

Az (1)-beli ábrázolások (és maga a *geminátá* terminus is) azt sugallják, mintha a rövid és hosszú mássalhangzók időtartama körülbelül 1 : 2 arányú lenne. Ez azonban természetesen távolról sincs így. Például Olaszy (2006) felolvasott szavakon végzett vizsgálata szerint a hosszú mássalhangzók időtartama átlagosan a rövidek időtartamának 165%-a; Beke–Gyarmathy (2010: 63) spontán beszédben mért adatai szerint ez az arány mindössze 131%.

A hosszú mássalhangzók (pontosabban a hosszú zárhangok és affrikáták) azonban nemcsak időtartamukban térnek el a megfelelő rövidektől, hanem belső felépítésükben is. Az ilyen komplex szerkezetű mássalhangzók, különösen az affrikáták esetében (Kovács 2002) bármely nyúlási folyamatnak nem egy, hanem két lehetséges célpontja van: a zárszakasz és a résszakasz. Elvileg lehetséges tehát, hogy a fonetikai és a fonológiai természetű nyúlási folyamatok más-más módon érintik ezeket a célpontokat. Pycha (2009) a zöngétlen magyar affrikáták fonetikai nyúlását és geminációját felolvasott anyagon vizsgálva arra az eredményre jutott, hogy a frázis végi nyúlás során az affrikátának az a része változik elsősorban, amely a frázishatárhoz (a folyamat kiváltójához) közelebb fekszik: a teljes időtartam-növekedés sokkal szorosabb korrelációt mutat a résszakasz időtartamának növekedésével, mint a zárszakaszéval. A frázis eleji nyúlás esetében viszont éppen ellenkezőleg: a teljes nyúlás mértéke jóval szorosabban korrelál a zárszakasz, mint a résszakasz nyúlásáéval. Eszerint a fonetikai nyúlás esetében mindig az érintett határhoz közelebb eső szakasz a nyújtó hatás elsődleges célpontja.

A *-val/-vel* által kiváltott fonológiai nyúlás (geminációs) esetében viszont Pycha a következő eredményeket kapta. Az *-on/-en/-ön* ragos (pl. *teknőcön*) és a *-val/-vel* ragos alakok (pl. *teknőccel*) teljes időtartama közötti arány

150%. Ennél érdekesebbek az affrikáták belső időszerkezetére kapott adatok: a sima (nem nyújtó) toldalék előtt álló affrikáták zárszakaszát a nyújtó toldalék előtt állókéval összevetve az arány 210%. A résszakaszban viszont csupán minimális nyúlás található: az arány 111%. Más szóval, a *-val/-vel* toldalék előtt az affrikáták zárszakasza kétszeresére nyúlik, míg a résszakasz időtartama lényegében változatlan marad. Tehát a fonológiai nyúlás (geminá-ció) esetében a nyújtó hatás elsődleges célpontja az affrikáta **zárszakasza,** a résszakasz csupán másodlagos célpont.

A fenti eredményeket háttérként figyelembe véve nézzük, mi a helyzet a degeminációs környezet esetében, vagyis a *kinccsel, tánccal* típusú példák-ban. Pycha (2010) mérési adatai szerint az ilyen esetekben az affrikáta teljes időtartama a megfelelő rövid affrikáták időtartamának 110%-a (míg a *ráccsal, páccal* típusúakban annak 150%-a). A degeminációs esetben tehát a hosszú mássalhangzó nem rövidül a röviddel egyformára, vagy másképp megfogalmazva: a szegélyező mássalhangzó nem akadályozza meg, csak na-gyon erősen korlátozza az affrikáta geminálódását. Ráadásul a zárszakasz időtartamának és a teljes affrikáta időtartamának aránya (a T/TS arány) nem-csak a korlátozatlan (szegélyező mássalhangzó nélküli) esetekben nő meg a geminátákban (0,39-ről 0,54-re), hanem korlátozott (szegélyező mássalhang-zót tartalmazó, degeminációs hatást mutató) esetekben is (0,26-ról 0,33-ra).

Pycha (2010: 147) a fentiekből azt a tanulságot vonja le, hogy az affrikáták geminálódásának a magyarban két különböző korrelátuma vagy „szignatúrá-ja" van, a teljes időtartam megnövekedése és a T/TS arány megnövekedése. Így a korlátozott geminációs helyzetében álló affrikáták egyszerre képesek ki-elégíteni a korlátozás és a geminálódás követelményeit. Például a *kinccsel* típusú alakok eleget tudnak tenni a geminációs igényének a belső T/TS arány növelésével, ugyanakkor a korlátozásnak is megfelelhetnek azzal, hogy teljes egészükben nem nyúlnak meg. Pycha eredményei szerint nagyrészt pontosan ez is történik (bár a teljes hosszúság csekély növekedése is megfigyelhető).

Mi következik mindebből a degeminációra nézve? Úgy tűnik, nem tartható az az elképzelés, hogy a *-val/-vel* ragos (és más hasonló) alakokban nem de-geminációval, hanem a **gemináció hiányával** állunk szemben. Azaz a (3) szabály esetében a folyamat **nem** korlátozható azokra az esetekre, amelyek-ben a bal oldali környezetet magánhangzó alkotja. A fentiek fényében az is kérdéses, hogy a (4) és (5) alatti degeminációs szabályok ebben a formában egyáltalán helytállóak-e. De akkor vajon mi magyarázza azt, hogy a hagyo-mányosan degeminációnak tekintett esetekben mégiscsak többé-kevésbé meg-rövidül (vagy nem nyúlik meg) az adott mássalhangzó?

A mássalhangzó-kapcsolatok időtartama és a degemináció

A kulcsot az adhatja meg, ha vetünk egy pillantást a különféle darabszámú mássalhangzó-kapcsolatok (CC, CCC stb.) időtartamára is. Olaszy (2007) vizsgálatában (10,5 hang/s artikulációs sebesség mellett) az intervokalikus (VCV helyzetű) mássalhangzók átlagos hossza 92 ms-nak adódott, míg a két-

tagú intervokalikus mássalhangzó-kapcsolatoké (VCCV) 162 ms-nak (tehát nem 184 ms-nak, ami az egyes mássalhangzók időtartamának egyszerű megkétszerezése lenne). Korábban már idéztük – Olaszy (2006)-ból – azt az adatot, amely szerint a hosszú mássalhangzók átlagos időtartama a rövidek átlagos hosszának 165%-a; itt pedig azt látjuk, hogy a kéttagú kapcsolatok átlagos hossza az egyes mássalhangzók átlagos hosszának 176%-a. Természetesen az átlagadatok sok egyedi különbséget elfednek, de feltűnő, hogy az arányszámok mennyire közel állnak egymáshoz. Nézzük most a CCC kapcsolatokat. Ezek átlagos hossza Olaszy (2007) anyagában 186 ms, azaz mindössze 15%-kal hosszabb, mint a kéttagúak időtartama, viszont a VCV helyzetű mássalhangzókénak szinte pontosan a kétszerese. Ebből már könnyen adódik az a hipotézis, hogy amit eddig degeminációs hatásnak hittünk, az nem más, mint a CCC kapcsolatok természetes rövidülése, vagyis hogy a mássalhangzó + gemináta kapcsolatok (bármely sorrendben) talán azért tűnnek mássalhangzó + egyeshang kapcsolatoknak, mert a teljes kapcsolat időtartama mindkét esetben szinte egyforma, de legalábbis távolról sem 3 : 2 arányú. Hipotézisünk – amelyet az alábbiakban kísérletileg megerősíteni vagy cáfolni kívánunk – tehát a következő:

Az ismert degeminációs hatások **nem geminátaspecifikusak,** hanem a CCC kapcsolatok fonetikai rövidülésének eredményei, akár jobbról, akár balról szegélyezett geminátákról, és akár valódi, akár álgeminátákról van szó.

Anyag, módszer és kísérleti személyek

A hipotézis vizsgálatához a következő kísérletet végeztük el. A kísérleti személyekkel 6 rövid szöveget olvastattunk fel. A vizsgált mássalhangzó-kapcsolatokat és fonológiai jellemzőiket az 1. táblázat foglalja össze. Kilenc ál-, három valódi levezetett és két mögöttes geminátát, tíz két-, hat három- és három négytagú kontroll-mássalhangzókapcsolatot elemeztünk.

A vizsgálandó szövegeket tíz 19 és 24 év közötti budapesti egyetemistával (5 nő és 5 férfi) olvastattuk fel háromszor. A beszélők hallása és beszédképzése ép volt. Egy női beszélő esetében csak két felolvasás eredményei voltak felhasználhatók. A többi beszélő esetében a később (az eredmények ismertetésénél) említendő ejtésváltozatok és néhány félreolvasás függvényében alakult az elemzett hangkapcsolatok elemszáma.

A felvételeket csendesített szobában rögzítettük, 16 bites, 22 100 Hz-es mintavételezéssel a Praat 5.3 szoftver (Boersma–Weenink 2012) hangrögzítőjének segítségével. Az elemzéseket ugyancsak a Praat szoftverben végeztük. Címkéztük a vizsgált mássalhangzó-kapcsolatok teljes, valamint a célmássalhangzó saját időtartamát.

A VC és CV határokon és a szonoránsokkal alkotott kapcsolatokban a magánhangzó második formánsának lecsengésénél és megindulásánál, közelítőhang-jellegű megvalósulás esetében a hangátmenet közepén, réshang esetében pedig a zörej kezdetéhez és végéhez legközelebbi nullátmeneten jelöltük

a hanghatárt. Az obstruensekkel alkotott mássalhangzó-kapcsolatokban a rés kezdetét és végét, illetve a zárfeloldás végét jelöltük meg hanghatárként. Ahol a zár alatt folyamatos réses áramlás volt megfigyelhető, a megelőző réshangra jellemző frekvencia- és intenzitásszerkezet alapján határoztuk meg a hanghatárt. Ha két zárhang találkozásakor nem volt adatolható az első zárfeloldása, nem jelöltük külön hangként.

1. táblázat. A vizsgált mássalhangzó-kapcsolatok

Kapcsolat	Hordozó szó(pár)	Jellemzők	
s#sp	*édes spenótot; sós spenótot*	ál-	jobbról szeg.
rs#sp	*nyers spenótot*	ál-	jobbról, balról
tt#sp	*főtt spenótot*	kontroll	4 C
sp, s#p	*a spenótot; és persze*	kontroll	2 C
rpp	*szörppel*	valódi	balról szeg.
rpt, rpr	*szörptől; szörpre*	kontroll	3 C
rp	*szörpöt*	kontroll	2 C
ntt	*ponttól*	ál-	balról szeg.
ntt	*ponttal*	valódi	balról szeg.
nt	*pontot*	kontroll	2 C
ntr, ntsz	*pontra; pontszám*	kontroll	3 C
lm(#)m	*filmmúzeumban, film még*	ál-	balról szeg.
lmm	*filmmel*	valódi	balról szeg.
lm	*filmet*	kontroll	2 C
lmsz	*filmszemle*	kontroll	3 C
lmkl	*filmklub*	kontroll	4 C
sz#szt	*olasz sztárok*	ál-	jobbról szeg.
ssz(#)t	*össztársadalmi, klassz társaságba*	mögöttes	jobbról szeg.
szt	*esztergályosok, olvasztárok*	kontroll	2 C
gszt	*világsztár*	kontroll	3 C
rt#szt	*mert sztárnak*	kontroll	4 C
nk#k	*magunk között, tolnánk ki*	ál-	balról szeg.
nkk	*magunkkal*	valódi	balról szeg.
nk	*senki, minket, magunk //*	kontroll	2 C

A statisztikai elemzéseket az SPSS 19.0 szoftver segítségével végeztük. A vizsgált hang abszolút, illetve a mássalhangzó-kapcsolathoz viszonyított relatív időtartamát vetettük össze az egyes kapcsolattípusok között. Az összevetendő mássalhangzó-kapcsolatok elemszámának eltérése és a minták nem normál eloszlása miatt nemparaméteres próbát alkalmaztunk (Kruskal–Wallis-próba).

Eredmények

A vizsgált hangkapcsolatok természetesen fonetikailag variábilisan jelenhetnek meg. A /l/ és /r/ bizonyos esetekben törlődhet (pl. Siptár 1994, Szende

2011), illetve az utóbbi változatos megvalósulásokat vehet fel (pl. Gósy 2008). A /r/ ejtése gyakran elmarad követő mássalhangzó esetén. A /lm/ kapcsolatokban a /l/ ejtése gyakorta elmarad, vagy (nagyrészt) a /m/-évél átfedésben történik (az artikulációs gesztusok időzítéséről l. pl. Hoole 1999). A /st/ néhány esetben nem tagolható rés + zárszakasz + felpattanás egységeire, hanem a rést azonnal felpattanás követi (vö. pl. Gráczi 2008). A veláris nazális ejtése pedig néhány esetben elmarad, csak a megelőző magánhangzó nazalizálódása érhető tetten. (A nazálisok törlését általában a folyamatos mássalhangzók előtti helyzetben tartjuk számon – vö. pl. Siptár (1994) –, az akusztikai elemzések során azonban egyes esetekben előfordult a /nk/ kapcsolat realizációjában is.) A /pt/ kapcsolatokban az első explozíva felpattanásának elmaradása is előfordult (vö. pl. Elekfi 1992). A *mert* szó pedig több lehetséges formában jelenhet meg (*mer, me*), amely egy vagy két mássalhangzó törlődését jelenti (vö. pl. Lanstyák 2009). Ezeknek a megvalósulásoknak az arányait és jelentőségét az egyes elemzéseknél részletesen megadjuk.

Az eredményeket az egyes vizsgált mássalhangzók szerint külön-külön elemezzük, mivel saját fonetikai jellemzőik is hatással lehetnek a megvalósulásokra.

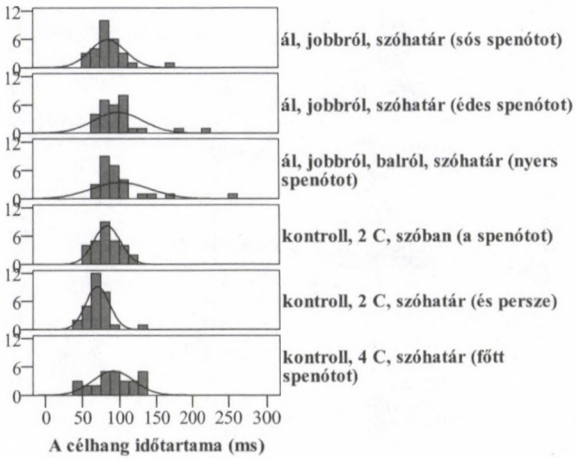
1. A (C)/ʃ(ː)p/ kapcsolatok. Négy hangkapcsolatot ki kellett zárni az elemzésből, így összesen 57 jobbról szegélyezett geminátát, 58 /ʃp/, 28 /tːʃp/ és 27 /rʃp/ kapcsolatot elemeztünk. Ezek közül nyolcszor hiányzott a /tː/ (13,8%) és 21-szer a /r/ akusztikai lenyomata (27,8%).

A két /ʃ/ + /ʃp/ kapcsolatban ejtett [ʃ] hangok átlagos időtartama 86 ± 24 ms és 99 ± 33 ms (1. ábra). Az eltérés 14%, amely az egyes beszélők ejtésében 0–31% között szóródott. A /rʃ/ + /ʃp/ kapcsolatban mért [ʃ] hangok átlagos időtartama 102 ± 39 ms. Ez mintegy 16%-kal hosszabb a rövidebb háromtagú kapcsolatban mért eredményeknél, és közel azonos a hosszabban mértekkel (2,8% eltérés). Ahol a két jobbról szegélyezett álgemináta közül az egyik hosszabb volt, ez mindig ugyanabban a szókapcsolatban volt adatolható *(édes spenótot)*. A négytagú kapcsolat ugyancsak minden beszélő ejtésében hosszabb [ʃ] hangot eredményezett, mint a rövidebb háromtagú esetében, míg a hosszabbhoz hasonlítva változó (–26–28%), hogy melyik realizáció hosszabb.

A kontrollhangsorokban kapott eredményekben 71 ± 16 ms és 92 ± 27 ms között alakult a mért célhang átlagos időtartama. A legrövidebb a szóhatáron (*és persze*), a leghosszabb pedig a *főtt spenótot* szókapcsolatban volt. Az előbbi esetében feltehetően az *és* szófaja is hozzájárul az eredmények alakulásához. Hét beszélő ejtésében 11–33%-kal hosszabb az *a spenót*-ban adatolható [ʃ], mint az *és persze*-ben található; egy ejtésben körülbelül azonos az időtartamuk (1%), és csak két ejtésben hosszabb az utóbbi 9, illetve 11%-kal. A *főtt spenótot* és az *a spenótot* hangsorok [ʃ]-ének időtartamát az egyes beszélők között összevetve négy ejtésben az előbbi hosszabb (17–32%-kal),

négy ejtésben közel azonos a két mért hang időtartama (0–3%), míg két beszélőnél az utóbbi hosszabb (16%, 36%).

A leghosszabbnak tapasztalt rövidült gemináta 10–31%-kal hosszabb, mint a kontrollhangok, a legrövidebbnek tapasztalt azonban velük közel azonos időtartamú (2,2–17%-kal hosszabb vagy 6,7%-kal rövidebb náluk). Az egyes beszélők esetében összevetve a vizsgált hangokat egyértelmű tendenciaként annyi vonható le, hogy a legtöbb ejtésben (hét beszélőnél a tízből) a legrövidebb az *és persze* kontrollhangsor [ʃ]-e, a leghosszabb valamelyik gemináta rövidülése, és a kéttagú hangkapcsolatokban mért [ʃ] hangok nyolc beszélő ejtésében rövidebbek, mint legalább az egyik rövidült álgemináta.



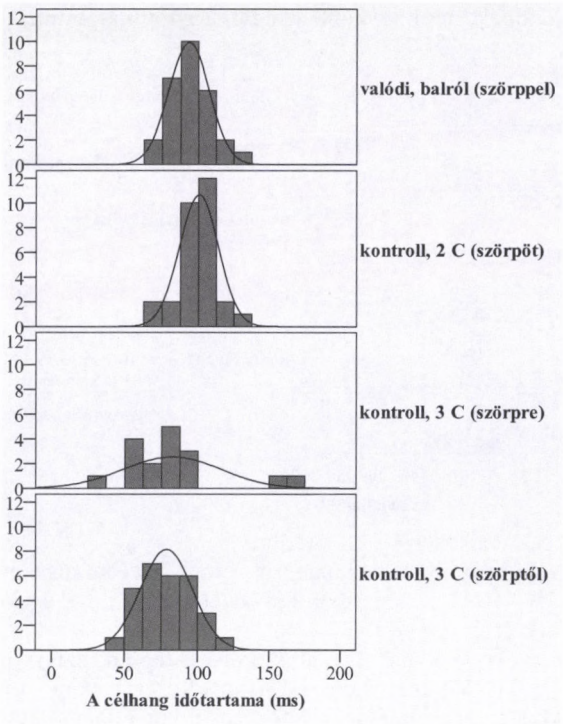
1. ábra
A [ʃ] hang időtartamának (ms) eloszlása (darabszám) a vizsgált helyzetekben

Az eredmények a Kruskal–Wallis-próba alapján szignifikáns eltérést mutatnak ($\chi^2 = 31{,}257$; $p < 0{,}001$). Ugyanakkor a hasonló típusú fonéma-kapcsolatok között tapasztalt nagy eltérések miatt általános tendenciát nem állapíthatunk meg.

2. A /rp(:)/(C) kapcsolatok. A második mért hang a [p] volt, amelyet egy valódi geminátа rövidülésében és három kontrollhelyzetben vizsgáltunk. 13 esetben nem lehetett mérni a vizsgált hang időtartamát: 11-szer elmaradt a zár utáni felpattanás a [pt] hangsorban (37,9%), egyszer a [rpr] hangsorban a zárfeloldás és a követő tremuláns nem volt elválasztható (3,4%), egyszer pedig a *szörppel* szót félreolvasta egy adatközlő (3,4%). (Egy beszélő esetében egy /rpt/ kapcsolatban nem volt külön mérhető a [p] időtartama.) A [r] 49-

szer törlődött. Ez a háromtagú fonémakapcsolatok esetében volt jellemzőbb (/rp/: 10,3%, /rp:/:75,9%, /rpr/: 65,5%, /rpt/: 17,2%).

A *szörppel* szóban a [p] időtartama átlagosan 93 ± 14 ms volt (2. ábra). A kéttagú kontrollhelyzetben (*szörpöt*) 100 ± 14 ms, a három mássalhangzóból álló két kapcsolatban (*szörptől, szörpre*) pedig 79 ± 18 ms és 82 ± 33 ms. Ebben az esetben tehát a vizsgált hang időtartama a kéttagú hangkapcsolatban mérthez hasonló (7% eltérés), míg a háromtagúakban mértnél hosszabb (11,7, ill. 15,8%-kal).



2. ábra
A [p] hang időtartamának (ms) eloszlása (darabszám) a vizsgált helyzetekben

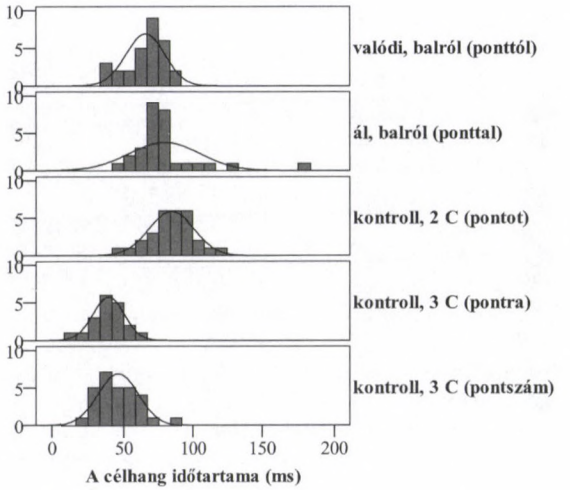
A két háromtagú kontrollhangsorban négy esetben a *szörpre* (11–40%), négy esetben a *szörptől* szóban mért [p] volt hosszabb (21–61%), kétszer pedig közel azonos volt az időtartamuk (6,7% és 7,1% eltéréssel).

A rövidült geminàta időtartama öt beszélő ejtésében rövidebb legalább 10%-kal a kéttagú hangkapcsolatban mért értéknél (11,7–15,4%), egyszer

11%-kal hosszabb, háromszor pedig közel azonos (1–9,9%). A két háromtagú kontrollhangsornál hat, illetve hét esetben hosszabb (10,7–68,1%) az ezekben mért [p] időtartamánál a rövidült gemináta, öt esetben közel azonos (0,5–6,1%), és egy esetben rövidebb (31,6%-kal).

Az eredmények a Kruskal–Wallis-próba alapján szignifikáns eltérést mutatnak ($\chi^2 = 26{,}386$; $p < 0{,}001$). Azaz a rövidült gemináta a két kontrollhelyzetben kapott értékeknél általában hosszabb időtartamban valósult meg.

3. A /nt(:)/(C) kapcsolatok. A [t] hang mérésében tizenhét realizációt kellett kizárni, mivel a *pontszám* szóban a /ts/ kapcsolat [t͡s] vagy [s] hangként jelent meg (24,1%), vagy kiesett (34,5%). Két beszélőnél minden esetben [t͡s]-ként valósult meg a *pontszám* /ts/ kapcsolata. A többi esetre kapott [t] időtartam-eloszlását az 3. ábra szemlélteti.



3. ábra
A [t] hang időtartamának (ms) eloszlása (darabszám) a vizsgált helyzetekben

A rövidült ál- és valódi gemináta átlagidőtartama 80 ± 26 ms és 66 ± 14 ms. A kéttagú kontrollkapcsolatban 84 ± 16 ms, a háromtagúakban pedig 46 ± 14 és 39 ± 13 ms. Az álgemináta tehát a valódinál hosszabb átlagos időtartammal valósult meg, lényegében a kéttagú kontrollhangsorban kapott adatokban mérttel azonos hosszúsággal, míg a háromtagú kapcsolatokban mért értékek mindháromnál rövidebbek (harmada-fele olyan hosszúak).

Ezeket a tendenciákat beszélőnként is elemezve az alábbi eredményeket kaptuk. A két háromtagú kontrollkapcsolat közül a *pontra* szóban mért [t] időtartama hat esetben hosszabb (19–36%), két esetben rövidebb (13–14%),

mint a *pontszám* szóban mért. A kéttagú kontrollhangkapcsolatban lévő [t] minden esetben hosszabb (29–62%-kal), mint a háromtagúakban lévő.

A két rövidült gemináta közül hat beszélőnél az álgemináta esetében 12–35%-kal hosszabb, négynél pedig közel azonos időtartamot (3–7%) mértünk.

A rövidült geminátákat a kontrollkapcsolatokkal összevetve a következő eredményeket kaptuk. A rövidült valódi gemináta minden esetben rövidebb, mint a kéttagú fonémakapcsolat [t]-jének időtartama (egy beszélőnél csak 7,3%-kal, a többinél 11–39%-kal); a háromtagúaknál egy eset kivételével (0,8%) hosszabb (16–51%). Az álgemináta esetében három beszélő ejtésében hosszabb időtartamot mértünk a kéttagú kontrollkapcsolatban kapotthoz képest (13–19%), a többiekében pedig rövidebbet (10–22%). A háromtagú kapcsolatokban ejtett [t]-knél minden esetben hosszabb volt (22–68%).

Az eredmények a Kruskal–Wallis-próba alapján szignifikáns eltérést mutatnak ($\chi^2 = 79{,}368$; $p < 0{,}001$). Az összes adat csoportonkénti eloszlását megfigyelve azt tapasztaljuk, hogy a három- és kéttagú fonémakapcsolatok megvalósulásaiban mért értékek közötti tartományban jelentek meg a rövidült geminátáké adatai.

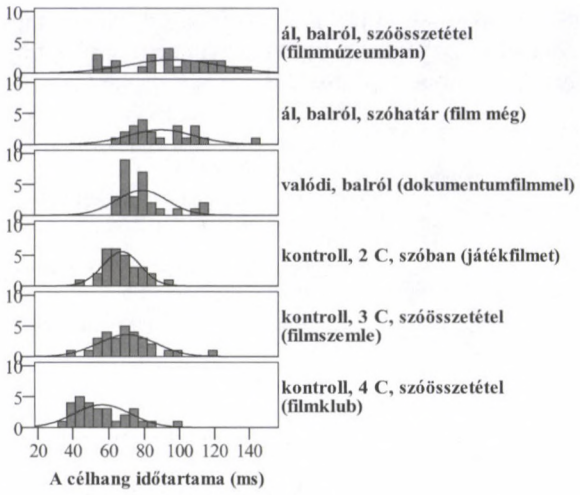
4. A /lm(:)/(C) kapcsolatok. Az álgemináták elemzéséből nyolc ejtést kellett kizárni, mivel a *film még* szókapcsolatban szünetet tartott a beszélő (27,6%). A /l/ lényegében csak kevés esetben volt bármely vizsgált hangkapcsolatban szegmentálható.

A vizsgált álgemináták átlagos időtartama 90 ± 19 ms és 97 ± 26 ms, a valódié pedig 79 ± 14 ms, vagyis a fentebbi eredményekhez hasonlóan a rövidült valódi gemináta rövidebb időtartamban realizálódott, mint a rövidült álgemináták (4. ábra).

A két álgemináta beszélőnkénti átlagos időtartama változatosan alakult. Négy esetben közel azonos volt (0–5% eltérés), négy esetben a *filmmúzeumban* (12–30%-kal), egy esetben pedig a *film még* szókapcsolat (33,1%) [m] hangja volt hosszabb. A *dokumentumfilmmel* hangsorban mért [m] hang három kivétellel rövidebb volt (13–42%-kal), mint az álgemináták, míg egy esetben közel azonos (3,8%), kétszer pedig hosszabb volt (17,6% és 28%; mindkétszer a szóhatáron átívelő geminátánál volt hosszabb).

A kontrollkapcsolatokban mért időtartamok változatosan alakultak. Némelyik beszélőnél a kéttagúban volt a leghosszabb és a négytagúban a legrövidebb a vizsgált mássalhangzó időtartama, míg másoknál a három- vagy négytagúban volt a leghosszabb.

Beszélőnként összevetve a rövidült geminátákat és a kontrollkapcsolatokat, nem kapunk egységes tendenciát. Elmondható ugyan, hogy a legtöbb esetben a *filmmúzeumban* szóban mért [m] a leghosszabb, a valódi gemináta rövidülése pedig a kontrollhangsorokban mértekkel közel azonos időtartamot eredményez, de az egyes fonémakapcsolatok megvalósulása időtartam tekintetében szinte beszélőnként más-más mintázatot mutat.



4. ábra
A [m] hang időtartamának (ms) eloszlása (darabszám) a vizsgált helyzetekben

Az eredmények a Kruskal–Wallis-próba alapján szignifikáns eltérést mutatnak (χ^2 = 61,464; p < 0,001). Ugyanakkor a geminátarövidülések és a három-, illetve négytagú kontroll-fonémakapcsolatok adatainak nagy szóródása csak óvatos következtetéseket enged levonni. A rövidült gemináták adatait a kéttagú kapcsolathoz viszonyítva elmondható, hogy az előbbiek jellemzően hosszabb időtartamban jelentek meg.

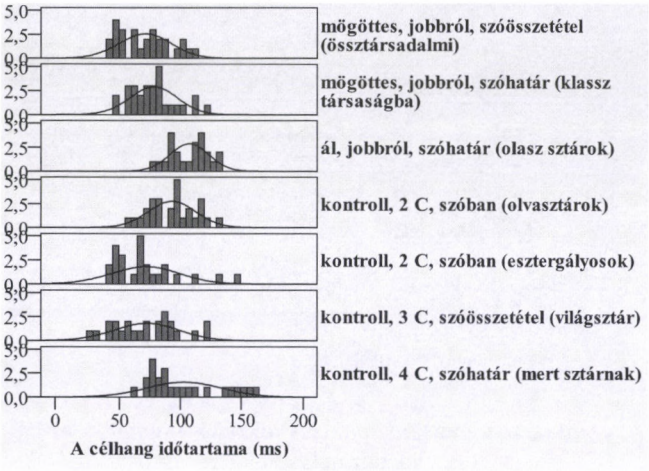
5. A (C)/s(:)t/ kapcsolatok. A /s(:)t/ kapcsolatokból kettőt kellett megakadás, négyet szünettartás, tizennégyet pedig a réshang és a felpattanó zárhang „összeolvadása" miatt kihagyni az elemzésből. A *mert sztárnak* szókapcsolatban a két első mássalhangzó csak ritkán jelent meg.

A két rövidült mögöttes gemináta időtartama 71 ± 21 ms és 78 ± 19 ms, az álgeminátáé pedig 107 ± 16 ms, vagyis 29–22%-kal hosszabb náluk (5. ábra).

Az egyes beszélők ejtésében változóan alakult a két mögöttes gemináta időtartama. Öt esetben az *össztársadalmi* szóban (15–35%), két esetben a *klassz társaságba* szókapcsolatban (11–32%) mért [s] hang volt hosszabb, míg három beszélőnél közel azonos volt az időtartamuk (< 10%). Az álgemináta csak egy beszélő esetében nem hosszabb legalább 10%-kal a mögöttes gemináták időtartamánál (13–49%).

A kéttagú kontrollkapcsolatokban az időtartam 72 ± 28 ms és 93 ± 20 ms. Az *olvasztárok* hangsor [s] hangja egy beszélő kivételével (nála azonos az időtartam) 12–41%-kal hosszabb az *esztergályosok*-ban mérthez képest. A háromtagú fonémakapcsolatban (*világsztár*) mért réshang hossza 72 ± 26 ms.

Ez azonos átlagos időtartam, mint az *esztergályosok* szóban mért, beszélőnként nézve azonban négyüknél az egyik, négyüknél a másik hosszabb, kettő esetében pedig közel azonos a [s] hosszúsága (< 10% eltérés).



5. ábra
A [s] hang időtartamának (ms) eloszlása (darabszám) a vizsgált helyzetekben

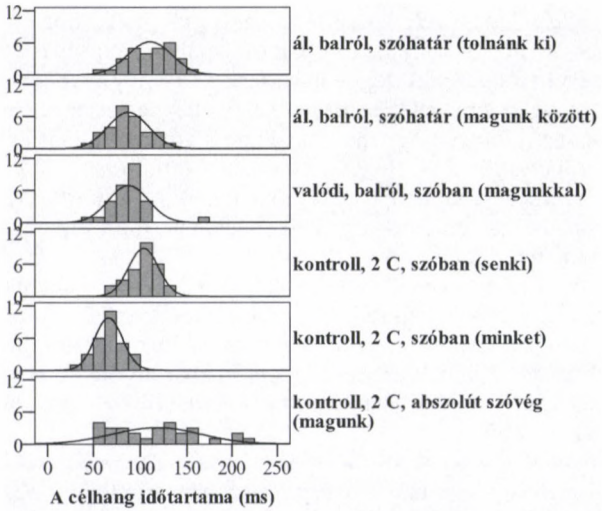
A négytagú fonémakapcsolatban (*mert sztárnak*) 103 ± 31 ms volt a réshang átlagos időtartama. Ez 12–48%-kal hosszabb, mint a rövidebb kéttagú kontrollkapcsolatban (*esztergályosok*) mért időtartama. Az *olvasztárok*-ban mért réshangnál hat beszélő ejtésében hosszabb (14–25%), kettő esetében rövidebb (13–15%), és kettő esetében közel azonos vele (< 10%). Ennek oka feltehetően a fentebb említett *(r)t*-törlődés, vagyis az, hogy lényegében nem négytagú a megvalósult hangsor. A rövidült geminátákat a kontrollhangsorokkal összevetve azt az eredményt kapjuk, hogy két beszélő ejtésében teljesül a feltételezés, hogy a rövidült geminátáák hosszabbak a háromtagú kapcsolatban mért időtartamoknál. Az álgemináta ismét több esetben követi a feltételezett tendenciát.

Az eredmények a Kruskal–Wallis-próba alapján szignifikáns eltérést mutatnak ($\chi^2 = 51{,}277$; $p < 0{,}001$). Ugyanakkor ebben az elemzésben is csak óvatos következtetések vonhatók le, amire elsősorban a két kéttagú kontrollfonémakapcsolat jelentős eltérése ad okot. Ebben a két szóban elsősorban a hangkörnyezet és a hordozó szótagok sorszáma különbözik, ami a relatíve nagy eltérést eredményezheti. További indokként esetleg az *olvasztárok* szó ritkasága merülhet fel. Több esetben is előfordult, hogy az adatközlő megállt előtte, avagy a felolvasást javítania kellett megakadás, lelassulás miatt. Mind-

emellett érdemes megjegyezni, hogy a mögöttes geminátáknak jelentősen rövidebbek az álgeminátáknál az esetek egy részében, azonban az értékek nagy szóródása miatt azzal azonos tartományban is gyakran megjelennek. Kérdés lehet, hogy ez a szóródás mivel indokolható. Az álgeminátát a háromtagú és az egyik kéttagú kontrollkapcsolatnál hosszabb időtartam jellemzi. A négytagú kapcsolat megvalósulási tartománya jellemzően az összes többi csoporttal átfedésben jelent meg. Ennek hátterében egyértelműen a *mert* ejtésváltozatai állnak.

6. A /nk(:)/ kapcsolatok. Egy kéttagú kontrollkapcsolatot kellett kihagyni az elemzésből félreolvasás miatt (3,4%), az esetek egy részében pedig a nazálist a megelőző magánhangzó nazalizációja képviselte.

A két álgemináta esetében a rövidült felpattanó zárhang időtartama 87 ± 21 és 110 ± 23 ms volt (6. ábra). A *tolnánk ki* [k]-ja nyolc beszélőnél volt hosszabb (21–35%), mint a *magunk között* [k]-ja. Egy beszélőnél az előbbi volt rövidebb (15%), egynél pedig azonos volt az időtartamuk (2,8%). A rövidült valódi gemináta átlagos időtartama ugyancsak 87 ± 21 ms volt. Három beszélő ejtésében hosszabb, háromében rövidebb, négy esetében pedig közel azonos átlagos időtartamban valósult meg, mint a *magunk között* [k]-ja.



6. ábra
A [k] hang időtartamának (ms) eloszlása (darabszám) a vizsgált helyzetekben

A három kéttagú kontrollkapcsolatban a *minket* szó [k]-ja a legrövidebb (66 ±15 ms), ezt a *senki*-é követi (103 ± 16 ms), a szóvégi helyzetben álló

magunk-é pedig a leghosszabb (118 ± 47 ms). Ez a tendencia beszélőnként hasonlóan alakult. Kivételként egy beszélőnél a *magunk* és a *minket* hasonló hosszúságú [k]-val valósult meg, négy beszélőnél pedig a *senki* [k]-ja hosszabb volt, mint a *magunk*-é.

Az átlagidőtartamok alapján a rövidült geminátáknak hosszabbak, mint a legrövidebb kontrollkapcsolatban mért felpattanók, ugyanakkor a másik két kontrollhangsorban mértnél rövidebb vagy ahhoz közeli időtartamban valósultak meg. Az egyes beszélőket egyesével elemezve is azt tapasztalhatjuk, hogy mindhárom rövidült gemináta hosszabb a *minket* szóban mért [k]-nál, míg a másik két kontrollkapcsolatban mérthez képest változó a viszonyuk.

Az eredmények a Kruskal–Wallis-próba alapján szignifikáns eltérést mutatnak ($\chi^2 = 58{,}382$; $p < 0{,}001$). Ennek ellenére az ezen fonémakapcsolatok realizációiban kapott adatok alapján a rövidült geminátáknak viselkedésére nem vonhatóak le általánosítható következtetések. A két álgemináta, illetve a két kéttagú (szó belseji) kontroll-fonémakapcsolat eltérése nagyobb mértékű, mint a rövidebb realizációkat adó álgemináta és a hosszabb realizációkat adó kéttagú (szó belseji) kontroll-fonémakapcsolat megvalósulásaié.

Összegzés

A jelen vizsgálatban azt a kérdést vizsgáltuk meg, hogy a geminátáknak további mássalhangzókkal alkotott kapcsolataiban létrejövő időtartambeli jellegzetességet (fonológiai) degeminációnak kell-e tekintenünk, vagy a hármas mássalhangzó-kapcsolatok (fonetikai) jellegzetességeit hordozzák az ilyen kapcsolatok, illetve hogy mindez függ-e a gemináta típusától.

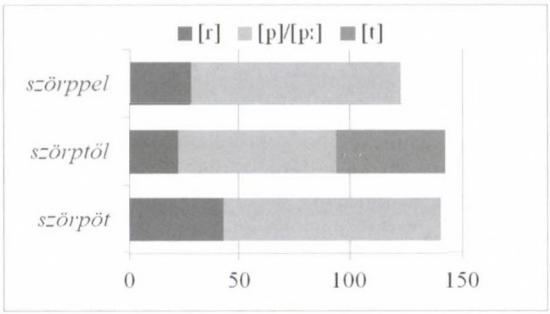
A kísérlet hat különböző fonéma vizsgálatát tartalmazta. Voltak egyszerű és összetett szerkezetű célhangok, és minden geminátatípus szerepelt a kísérleti anyagban. Az eredmények egy része alátámasztotta hipotézisünket, más része azonban ellentmondott neki.

A *szörp* és a *pont* szavak toldalékolt és összetételi alakjaiból származó adatok a „degemináció" ellen szóltak, az *össztársadalmi* és *klassz társaságba* hangsorok azonban mellette. Az alábbiakban azon megvalósulások eredményeit foglaljuk össze, amelyekben minden fonéma különálló beszédhangként valósult meg, hogy a korábban tapasztalt variabilitás egyik okát kizárjuk ebből az összegzésből.

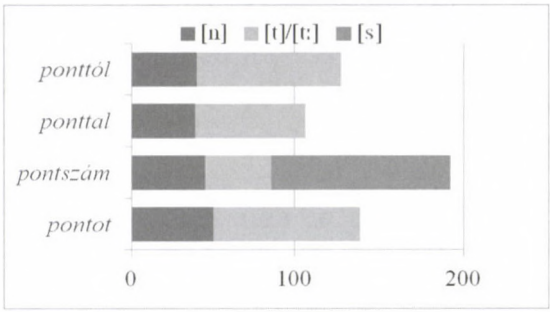
A teljes mássalhangzó-kapcsolat átlagos hossza a *szörpöt* szóban 140 ms, a *szörptől* szóban 142 ms, míg a *szörppel* szóban mindössze 123 ms volt; a szegélyező mássalhangzó (a [r]) az utóbbi két esetben egyaránt megrövidült (22 és 28 ms-ra) a kéttagú kapcsolatban mért időtartamához képest (43 ms). Ez azt sugallja, hogy a *szörppel* szóban nem a hagyományos értelemben vett degemináció, hanem a teljes kapcsolat időbeli „összenyomódása" történik, a háromtagú kapcsolathoz hasonlóan (7. ábra).

A *ponttól*-beli álgemináta és a *pontot*-beli [t] átlagos hosszúságának megegyező volta a degeminációs elemzés felé billenti ugyan a mérleget, viszont a

ponttól [n]-jének időtartama kevesebb, mint a [t] időtartamának fele, ami a degeminációs elemzés ellen szól (még akkor is, ha a *pontot* szóban is hosszabb a [t], mint a [n]; l. a 8. ábrát).



7. ábra
A mássalhangzó-kapcsolat egyes tagjainak átlagos időtartama (ms) a *szörp* szó toldalékolt alakjainak azon megvalósulásaiban, ahol mindkét/mindhárom hang adatolható volt az ejtésben



8. ábra
A mássalhangzó-kapcsolat egyes tagjainak átlagos időtartama (ms) a *pont* szó toldalékolt alakjainak és összetételének azon megvalósulásaiban, ahol mindkét/mindhárom hang adatolható volt az ejtésben

A [s]/[s:] mássalhangzók elemzésében vizsgált mögöttes gemináták rövidülésére kapott eredmények esetében azonban felmerül a tényleges rövidülés, a „degemináció" lehetősége, mivel a vizsgált gemináták realizációinak időtartama a további mássalhangzókkal alkotott kapcsolatokban jellemzően rövidebb volt az álgeminátákban mért értékeknél (85 ms és 81 ms vs. 113 ms). A kontrollhangsorok adatainak nagy eltérései miatt azonban ez a típus további vizsgálatot igényel.

A további elemzett adatok azt mutatják, hogy – a korábban említett nagy variabilitás miatt – esetükben további vizsgálati szempontok, illetve terjedelmesebb vizsgálati anyag bevonása lesz indokolt.

Következtetés

Vizsgálatunk kiinduló hipotézise az volt, hogy az ismert degeminációs hatások talán nem is geminátaspecifikusak, hanem a CCC kapcsolatok fonetikai rövidülésének eredményei, akár jobbról, akár balról szegélyezett geminátákról, és akár valódi, akár álgeminátákról van szó. Összefoglalásképpen azt állapíthatjuk meg, hogy jóllehet hipotézisünk a maga egyszerű és kézenfekvő általánosságában nem igazolódott be, annyi mindenesetre kiderült, hogy a degeminációsnak tartott esetek eddigi kategorikus (fonológiai) kezelése nem tartható fenn, s a jelenségkör valódi magyarázataként legalábbis valami nagyon hasonlóra gyanakodhatunk, mint amit a hipotézisben próbáltunk megfogalmazni. Mindenesetre számos további kísérletre, nagyobb anyagon végzett elemzésre lesz még szükség ahhoz, hogy a felvetett kérdésben valóban tisztán lássunk.

Irodalom

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

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

Boersma, Paul – Weenink, David 2012. *Praat: doing phonetics by computer.* Software package. www.praat.org.

Davis, Stuart 2011. Geminates. In van Oostendorp, Marc – Ewen, Colin – Hume, Elizabeth – Rice, Keren (eds.): *The Blackwell companion to phonology.* Wiley–Blackwell, Oxford. 1597–1621.

Dressler, Wolfgang U. – Siptár, Péter 1989. Towards a natural phonology of Hungarian. *Acta Linguistica Hungarica* 39. 29–51.

Dressler, Wolfgang U. – Siptár Péter 1998. A magyar nyelv természetes fonológiája felé. *Általános Nyelvészeti Tanulmányok* 19. 35–59.

Durand, Jacques – Siptár Péter 1997. *Bevezetés a fonológiába.* Osiris Kiadó, Budapest.

Elekfi László 1992. *A magyar hangkapcsolódások fonetikai és fonológiai szabályai.* MTA Nyelvtudományi Intézet, Budapest.

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

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

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.

Kovács Magdolna 2002. Az affrikáták időszerkezetéről. In Hunyadi László (szerk.): *Kísérleti fonetika, laboratóriumi fonológia.* Debreceni Egyetem Kossuth Könyvkiadója, Debrecen, 39–54.

Lanstyák István 2009. *A magyar beszélt nyelv sajátosságai*. Stimul, Pozsony.
Nádasdy, Ádám 1989. The exact domain of consonant degemination in Hungarian. *Hungarian Papers in Phonetics* 21. 104–107.
Olaszy Gábor 2006. *Hangidőtartamok és időszerkezeti elemek a magyar beszédben*. Nyelvtudományi Értekezések 155. Akadémiai Kiadó, Budapest.
Olaszy Gábor 2007. *Mássalhangzó-kapcsolódások a magyar beszédben*. Segédkönyvek a nyelvészet tanulmányozásához 72. Tinta Könyvkiadó, Budapest.
Polgárdi, Krisztina 2008. Geminates and degemination in Hungarian: A loose CV analysis. In Piñón, Christopher – Szentgyörgyi, Szilárd (eds.): *Approaches to Hungarian 10. Papers from the Veszprém conference*. Akadémiai Kiadó, Budapest. 127–146.
Pycha, Anne 2009. Lengthened affricates as a test case for the phonetics–phonology interface. *Journal of the International Phonetic Association* 39. 1–31.
Pycha, Anne 2010. A test case for the phonetics–phonology interface: gemination restrictions in Hungarian. *Phonology* 27. 119–152.
Ringen, Catherine O. – Vago, Robert M. 2011. Geminates: heavy or long? In Cairns, Charles E. – Raimy, Eric (eds.): *Handbook of the syllable*. Brill, Leiden, 155–169.
Siptár Péter 1994. A mássalhangzók. In Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 2. Fonológia*. Akadémiai Kiadó, Budapest. 183–272.
Siptár Péter 2001. Degemináció. In Bakró-Nagy Marianne – Bánréti Zoltán – É. Kiss Katalin (szerk.): *Újabb tanulmányok a strukturális magyar nyelvtan és a nyelvtörténet köréből. Kiefer Ferenc tiszteletére barátai és tanítványai*. Osiris Kiadó, Budapest. 291–303.
Siptár Péter 2012. Tényleg van-e a magyarban degemináció? In: Markó Alexandra (szerk.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngekezdési időig*. ELTE BTK–MTA Nyelvtudományi Intézet, Budapest. 19–34.
Siptár, Péter – Törkenczy, Miklós 2000. *The phonology of Hungarian*. Clarendon Press–Oxford University Press, Oxford.
Stiebels, Barbara – Wunderlich, Dieter 1999. Second stems in Hungarian nouns. *The Linguistic Review* 16. 253–294.
Szende Tamás 2011. Változási tendenciák a mai magyar beszédben. *Magyar Nyelvőr* 135. 107–112.
Trón, Viktor – Rebrus, Péter 2005. Re-presenting the past: contrast and uniformity in Hungarian past tense suffixation. In Piñón, Christopher – Siptár, Péter (eds.): *Approaches to Hungarian 9. Papers from the Düsseldorf conference*. Akadémiai Kiadó, Budapest. 305–327.
Vago, Robert M. 1980. *The sound pattern of Hungarian*. Georgetown University Press, Washington.

AZ ARTIKULÁCIÓS TEMPÓ VARIABILITÁSA FELOLVASÁSBAN

Váradi Viola – Beke András

Bevezetés

A 20. századi akusztikai fonetikai vizsgálatok túlnyomórészt laboratóriumi beszédet és felolvasást elemeztek. A spontán beszéd vizsgálata csak a 20. század vége felé került előtérbe, de elemzésére már korábbról is találunk néhány példát (Gósy–Horváth 2007). A felolvasott szövegek akusztikai fonetikai szempontú vizsgálata azonban még mindig napjaink egyik jelentős területe, mivel a spontán beszédben sokszor kontrollálhatatlanok az egyes tényezők. A felolvasások elemzésének nagy előnye, hogy szó szerint ugyanazon nyelvi anyag akusztikai megvalósulását vizsgálhatjuk több beszélő esetén. Számtalan hátránya is lehet, ezek közül az egyik legfontosabb, hogy az adatközlők különbözhetnek az olvasási képesség tekintetében. Ez megnehezíti a különböző beszélőktől származó hanganyagok összehasonlítását (Jacewicz et al. 2010).

A szóbeli megnyilatkozások kategorizálása számos terminológiai és meghatározásbeli nehézséget okoz mind a hazai, mind a nemzetközi szakirodalomban. A különböző beszédstílusok listája szinte végtelen, csakúgy, mint az osztályozásba bevonható szempontoké. Ha a különböző beszédstílusokat egy skála mentén helyezzük el, akkor a skála egyik végpontján a felolvasott beszéd, míg a másikon a hétköznapi spontán beszéd áll (Gósy–Bóna 2011). Laan (1997) holland nyelven végzett percepciós vizsgálatának eredményei szerint rádióhallgatás közben nagy biztonsággal eldönthető, hogy a beszélő felolvas-e egy szöveget, vagy pedig spontán beszél. A hallgató percepciós döntése a megnyilatkozás beszédstílusát illetően számos beszédjellemzőn alapszik, például grammatikai szerkesztettség, szóválasztás, beszédsebesség, szünettípusok, szünetek felépítése, hanglejtés és az artikuláció gondossága. Lindblomnak (1990) a beszédprodukció és –percepció összefüggéseit modellező H&H (hyper and hypo speech – 'túlartikulált és alulartikulált beszéd') elméletének alapfelvetése, hogy az artikuláció a feszesség tekintetében a túlés az alulartikulált beszéd két végpontja között írható le. A két legszélsőségesebb beszédstílus, a felolvasás és a spontán beszéd, ennek a skálának a két végpontja felé közelít: a felolvasás bizonyos értelemben a túlartikulált, míg a spontán beszéd az alulartikulált beszédhez áll közelebb.

Wacha Imre (1999) szerint a felolvasás és a spontán beszéd erőteljesen elkülönül egymástól. Az elkülönülés a beszélőnek a szöveghez való viszonyával és a szöveggel való szándékával magyarázható. Az élőszóban, természetes beszédhelyzetben a tervezés, a szöveg létrehozása és hangoztatása egyide-

jűleg, szimultán folyamatban zajlik le (Levelt 1999). Ilyenkor a beszélőnek nem áll módjában, hogy előzetesen megfogalmazza a gondolatait. Felolvasás esetén azonban a beszélő egy korábban rögzített, megtervezett szöveget hangosít meg, tehát nincs szüksége a gondolatok megfogalmazására, a szöveg alkotása és megszólaltatása nem egyszerre történik (Wacha 1999). A különböző beszédstílusok közötti különbségek a beszédnek mind a szegmentális (vö. Gósy 1997a; Olaszy 2000; Beke–Gráczi 2010; Gráczi–Horváth 2010; Markó et. al 2010), mind a szupraszegmentális (vö. Gósy 2000; Markó 2005, 2009; Beke 2008; Váradi 2010; Bóna 2011; Markó–Bóna 2012) jellemzőiben testet öltenek. A spontán beszédet a nem spontán beszédfajtáktól megkülönböztető prozódiai jellegzetességek (vö. Váradi 2010) a beszédprodukciós folyamat működésének eltéréseire vezethetők vissza. Felolvasás esetén mind a tagolás, mind a szünetezés a központozás által vezetett. Spontán beszéd létrehozása során a beszélőnek nem áll módjában minden esetben ezek tudatos megtervezése, sokkal inkább automatikus a szupraszegmentális elemek megvalósítása (Gósy 2004).

A beszédstílusokat jól jellemzi az is, hogy azok milyen temporális tulajdonsággal bírnak. Az artikulációs működések sebességét a fonetikai szakirodalomban az artikulációs tempóval szokás kifejezni. Az artikuláció tiszta idejére eső nyelvi jelek számától meg kell különböztetnünk a beszédtempót, amely a teljes beszédidőre eső nyelvi jelek számát jelenti (Gósy 2004). A kétféle tempó között a legfőbb különbség, hogy az artikulációs tempó kiszámításánál nem vesszük figyelembe a szüneteket, míg a beszédtempónál igen. A beszédtempó komplexebb sajátosság, ezért kevésbé informatív, mint ha külön elemezzük az artikulációs tempót, a szünetek időtartamát és gyakoriságát. Ugyanazon nyelv különböző beszélői vagy különböző anyanyelvű beszélők eltérő stratégiákat alkalmazhatnak a beszéd és a szünetezés megvalósítására (Dankovičová 1997).

Korábban azt feltételezték, hogy a beszédtempó variabilitása a szünetezés következménye, de az artikulációs tempó többé-kevésbé állandó (Dankovičová 1997). Az újabb kutatások ennek ellenkezőjét igazolták. Az artikulációs működések temporális sajátosságait számtalan tényező befolyásolja, ezért jelentős egyéni különbségek adatolhatók, és ugyanazon beszélő esetén sem jellemző a tempóértékek állandósága (vö. Quené 2004, 2007, 2008; Jacewicz et al. 2009, 2010; Gósy et al. 2011). Subosits (1990) a beszédtempót befolyásoló tényezők két nagy csoportját különbözteti meg, a belső (egyéni sajátosságok) és a külső tényezőket (beszédtéma, beszédhelyzet stb). Mind a hazai, mind a nemzetközi szakirodalomban számtalan fonetikai vizsgálat foglalkozott például a beszédtempó és az életkor (Fónagy–Magdics 1960; Oyer–Deal 1985; Laczkó 1991; Balázs 1993; Gósy 1997b; Gocsál 2000; Menyhárt 2000; Bóna 2010; Jacewicz et al. 2010), a beszéd sebessége és a nem (Oyer–Deal 1985; Tsao–Weismer 1997; Gocsál 2001; Verhoeven et al. 2004; Jacewicz et al. 2010) vagy az artikulációs tempó és a dialektus, lakóhely (Verhoeven et

al. 2004; Quené 2008; Jacewicz et al. 2009; Menyhárt 2010), a beszédtempó és a beszédmegértés (Gósy 1997b), illetve a beszédhelyzet, beszédstílus és a beszéd temporális sajátosságainak összefüggésével (Laczkó 1993; Andó 2002; Imre 2005; Markó 2005; Olaszy 2005, 2006; Váradi 2010).

Jacewicz és munkatársai (2010) az amerikai angol beszédtempó variabilitását elemezték az életkor és a nyelvjárás függvényében mondatfelolvasásokat és spontán beszédet tartalmazó korpuszon. Az északról származó adatközlők beszédtempója mind a felolvasásban, mind a spontán beszédben szignifikánsan gyorsabb volt, mint a délieké. Keresztmetszeti vizsgálatuk eredményei szerint a beszédtempó a felolvasásokban és a spontán narratívákban az életkorral változik: először az életkorral együtt emelkedik, a negyvenes évek közepe táján éri el a maximumértéket, majd az életkor előrehaladtával csökken. A férfiak spontán beszéde szignifikánsan gyorsabb volt, mint a nőké. A beszédszakasz (szünettől szünetig terjedő szöveges egység) hossza összefüggést mutatott az artikulációs tempóval: az adatközlők rövidebb beszédszakaszokat hoztak létre, ha gyorsabb tempóval beszéltek. Az északról származó beszélőknél rövidebb időtartamú beszédszakaszokat adatoltak, mint a délieknél, és az idősebb beszélők is rövidebb beszédszakaszokat hoztak létre, mint a fiatalabbak.

Az elhangzó beszéd prozódiai szerkezetének leírására mind a nemzetközi, mind a hazai fonológiai és fonetikai szakirodalomban többféle elképzelést olvashatunk. Az egyik legelfogadottabb elmélet szerint a prozódiai szerkezet a szintaktikai szerkezethez hasonlóan, egymásra szinteződve épül fel (fastruktúraszerűen) (Gussenhoven 2004; Roca–Johnson 2005; Varga 1994; Hunyadi 2004). Felülről lefelé haladva a következő szintek különíthetőek el: megnyilatkozás (utterance), intonációs frázis (intonational phrase), fonológiai frázis (phonological phrase), fonológiai szó (phonological word), láb (foot), szótag (syllable). A fonológiai elemzéseken alapuló hierarchikus modellek elsősorban a szintaktikailag kifogástalan felépítésű szerkezetek jellemzésére alkalmazhatók. Természetesen a különböző prozódiai egységek a spontán beszédben is értelmezhetők, de szerveződésük jelentősen eltér a nem spontán beszédformákétól.

A prozódiai és a szintaktikai szerkezet közötti összefüggéssel számos nyelvészeti kutatás foglalkozott. A legelfogadottabb álláspont szerint a prozódiai és a szintaktikai szerkezet szorosan összefügg, de közöttük a kapcsolat nem egyértelműen meghatározott, és természetét tekintve sem teljesen feltárt (pl. Savy–Voghera 2010). Szaszák és Beke (2012) olvasott és spontán beszédben elemezte statisztikai módszerekkel a beszédjel szintaktikai és prozódiai szerkezetét. Kutatásuk fő kérdése az volt, hogy lehetséges-e a szintaktikai szerkezet feltárása a beszédjel prozódiai elemzése alapján. Olvasott beszédben a prozódia megbízhatóan jelezte a szintaktikai határokat: elkülönítette a tagmondathatárokat a szószerkezetek határaitól. Spontán beszéd esetén a prozódiai szerkezet alapján kevésbé volt megbízható a szintaktikai tagolás.

Váradi (2012) olyan egységeket elemzett a tagolhatóság szempontjából, amelyek jól elkülöníthetők a felolvasásban és a spontán beszédben is. Olaszy (2005) meghatározása alapján a szünettől szünetig terjedő beszédszakaszokat prozódiai frázisoknak (PF) nevezte el, melyek intonációs frázisokra (IF), azaz a hanglejtés és a hangsúlyozás szempontjából egy egységnek tekinthető frázisokra tagolhatók. Az eredményekből azt a következtetést vonta le, hogy a két prozódiai egység akusztikai fonetikai realizációja eltérő. Az intonációs frázisokra gyorsabb artikulációs tempó, alacsonyabb alapfrekvencia-maximum és szűkebb hangterjedelem jellemző, mint a prozódiai frázisokra. Az artikulációstempó-értékek a prozódiai frázisokban kisebb variabilitást mutattak, mint az intonációs frázisokban. Felolvasás esetén szignifikánsan gyorsabb volt az adatközlők artikulációs tempója a prozódiai frázisokban és az intonációs frázisokban is, mint spontán beszédben. Az artikulációs tempó variabilitása a kisebb prozódiai egységekben, a hangsúlyközi szakaszokban (IF) figyelhető meg, amely variabilitás a nagyobb egységekben kiegyenlítődni látszik. A felolvasás és a spontán beszéd eltérő akusztikai fonetikai megvalósulása a prozódiai egységek szintjén is igazolható volt.

Dankovičová (1999) angol és cseh nyelvű felolvasásokban elemezte a *rallentando*-t, vagyis a lassulási folyamatot a prozódiai egységeken (intonációs frázis és fonológiai szó) belül. A kutatás eredményei azt mutatták, hogy az artikulációstempó-értékek változása a prozódiai egységeken belül nem önkényesen, hanem rendszerszerűen szerveződött. A rallentando jelensége mind a cseh, mind az angol nyelvű anyagban igazolható volt. Az artikulációs tempót az adott egységen belüli pozíción túl az egység mérete (a szótagszám növekedésével gyorsul a tempó) és a szóosztály (tartalmas szavak tempója lassabb volt, mint a funkciószavaké) is befolyásolta.

A jelen kutatásunkban arra a kérdésre kerestük a választ, hogy magyar nyelvű felolvasásokban milyen mértékben jellemző a szavak artikulációs tempójának variabilitása az intonációs frázison belül, és milyen tényezők befolyásolják ezt a variabilitást. Kiinduló hipotézisünk, hogy az artikulációstempó-értékek változása nem véletlenszerű az intonációs frázison belül. Feltételezésünk szerint az artikulációs tempót befolyásoló tényezők a következők: a szó pozíciója (hányadik helyen áll) az intonációs frázison belül; az intonációs frázis pozíciója (hányadik helyen áll) a beszédszakaszon belül; a szót felépítő szótagok száma; a szó szófaja; szóosztály (tartalmas vagy funkciószó). Feltételeztük továbbá, hogy a tempóvariabilitás nem véletlenszerű, hanem meghatározott, ismétlődő mintázatokat követ.

Anyag, módszer, kísérleti személyek

A Beszélt nyelvi adatbázis (BEA, Gósy 2008) 4 férfi és 4 női adatközlőjének hangelvételeiből a felolvasásokat elemeztük. Az adatközlők átlagéletkora 28 év volt. Az adatközlőknek egy 13 mondatból álló tudománynépszerűsítő szöveget kellett felolvasniuk.

A lejegyzéshez a Praat 5.3 szoftvert használtuk (Boersma–Weenink 2012). A hanganyagot manuálisan, a spektogram és az oszcilogram alapján, folyamatos auditív ellenőrzéssel beszédszakasz-, intonációsfrázis- és szószinten annotáltuk. A beszédszakasz elkülönítésénél a szüneteket tekintettük határjelölőnek. Az intonációs frázisokat a hangsúlyozás és a hanglejtés egységessége alapján különítettük el egymástól. Abban az esetben tekintettük hangsúlyosnak az adott szótagot, ha a környezetéhez képest az alapfrekvencia és/vagy az intenzitás kiemelkedett az akusztikai lenyomatban a környező szótagok közül. Összesen 307 db beszédszakaszt, 609 db intonációs frázist és 1469 db szót adatoltunk a 8 adatközlő felolvasásában.

Az artikulációs tempó kinyeréséhez a Praat 5.1 szoftverben írt algoritmust alkalmaztuk: kinyertük a szavak és az intonációs frázisok artikulációs tempóját szótag/s-ben kifejezve (a szótagok számának megállapításához a szavakban lévő magánhangzók számát vettük alapul).

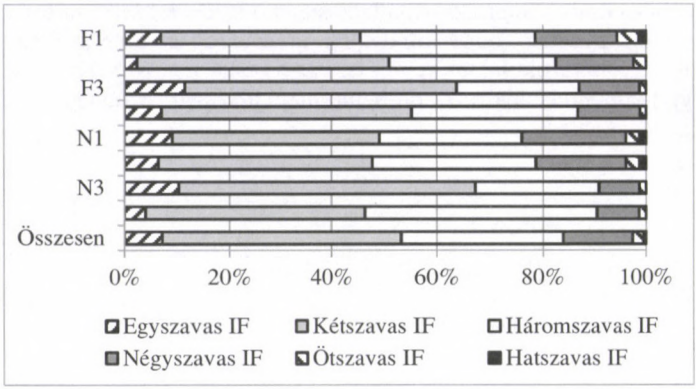
A statisztikai elemzések során az intonációs frázisok szavainak temporális variabilitását elemeztük. Az általánosított lineáris modellálás (Generalized Linear Model, GLM) módszerét választottuk Tukey-féle post hoc teszttel kombinálva, amely az átlagok közötti különbségek szignifikanciáját mutatja meg páronként. A lineáris modellek olyan magyarázó modellek, amelyek egy folytonos változó (függő változó) alakulását kívánják megragadni több független változó lineáris függvényben. Ennek általánosítása a GLM, amelynél nem magát a függő változót, hanem annak egy transzformált értékét becsüljük a magyarázó változók lineáris függvényével. Ebben az elemzésben a modell építésének jóságáról két fontos mérőszámot kell vizsgálnunk. A korrigált R^2 érték a felépített modell sikerességéről nyújt információt, az η^2 érték a kapcsolat erősségének mértéke, hozzávetőleges becslést ad a különböző tényezők relatív erejéről. Az egyének közötti hasonlóság mérésére korrelációelemzést végeztünk.

Az elemzéseket az SPSS 13.0 verziójával végeztük el 95%-os megbízhatósági szinten. A szófaji elemzéshez a HUNMORPH ingyenesen elérhető szoftvert használtuk, majd a kimenetet manuálisan egyértelműsítettük, illetve a szoftver által nem ismert szavakhoz szófaji kategóriát rendeltünk.

Eredmények

A vizsgált korpusz mintegy felét a két szóból felépülő intonációs frázisok teszik ki. A szószám növekedésével csökken az intonációs frázisok aránya, vagyis minél több szóból áll az intonációs frázis, annál kisebb arányban fordul elő. Ez a tendencia mindenkinél megfigyelhető, kivéve az N4 adatközlőt, akinél a háromszavas intonációs frázisokból eggyel többet adatoltunk, mint a kétszavasokból (1. ábra).

A további elemzésekbe csak a két, három, négy szóból álló intonációs frázisokat vontuk be, mivel az egy, illetve az öt vagy annál több szóból felépülő intonációs frázisok előfordulási aránya nagyon alacsony volt.



1. ábra
Az intonációs frázisok eloszlása az őket alkotó szavak száma alapján adatközlőnként és összesen (IF = intonációs frázis)

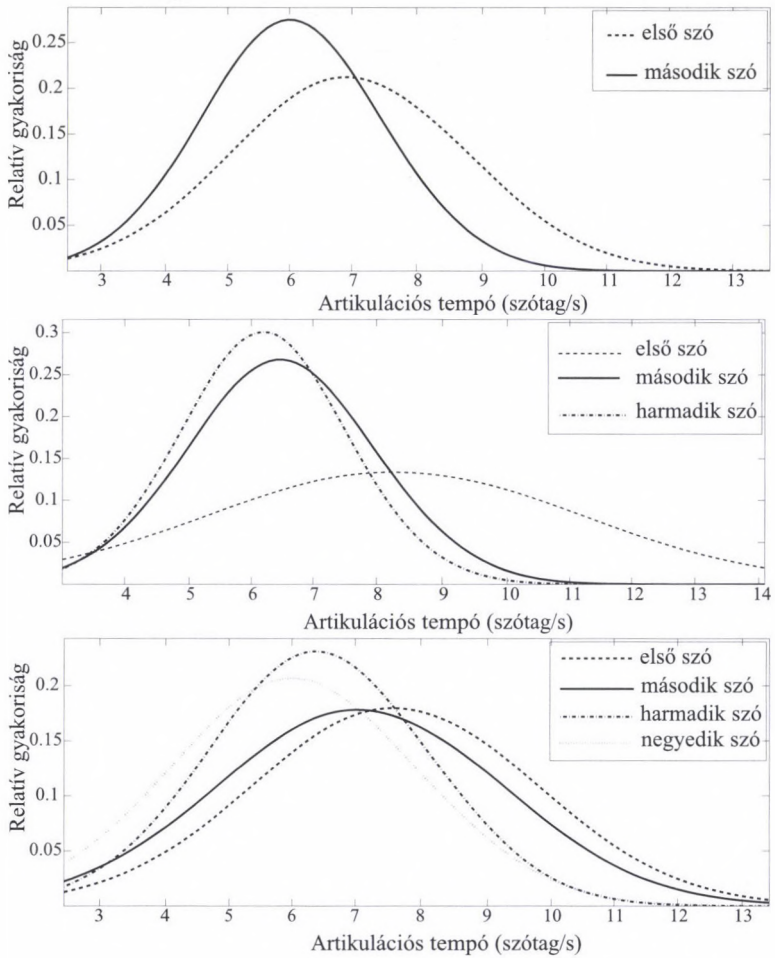
A felolvasott beszédben Pearson-féle korrelációelemzéssel vizsgálható, hogy az intonációs frázisok szavainak artikulációs tempója milyen mértékű variabilitást mutat a beszélők között (1. táblázat).

1. táblázat: A szavak artikulációs tempójának beszélők közötti korrelációja (R^2)

Beszélők	F1	F2	F3	F4	N1	N2	N3	N4
F1		0,838	0,793	0,900	0,840	0,787	0,786	0,896
F2	0,838		0,849	0,860	0,834	0,787	0,891	0,829
F3	0,793	0,849		0,854	0,714	0,752	0,857	0,785
F4	0,900	0,860	0,854		0,826	0,839	0,840	0,912
N1	0,840	0,834	0,714	0,826		0,747	0,753	0,825
N2	0,787	0,787	0,752	0,839	0,747		0,793	0,800
N3	0,786	0,891	0,857	0,840	0,753	0,793		0,787
N4	0,896	0,829	0,785	0,912	0,825	0,800	0,787	

Közel azonos artikulációs tempóval valósulnak meg az egyes szavak a felolvasásban, hiszen az R^2 értéke 99%-os konfidenciaintervallumon 0,714 és 0,912 között alakul. Az artikulációs tempó megválasztása közel sem önkényes a felolvasáskor, hanem valamilyen mintázatot követ. Ezen mintázat feltérképezésére megvizsgáltuk az egyes intonációs frázisokban realizálódott szavak artikulációs tempóját. Az első hipotézisünk tesztelésére összehasonlítottuk a két, három és négy szóból felépülő intonációs frázisokon belül a különböző pozíciókban előforduló szavak artikulációs tempójának relatív gyakoriságát (2. ábra). Az empirikus eloszlásra illesztett normál eloszlás terje-

delméből jól látszik, hogy az artikulációstempó-értékek széles tartományban szórnak. Megfigyelhető, hogy mindegyik intonációs frázisban az első szónak van a legnagyobb artikulációs tempója, vagyis ez a leggyorsabban megvalósuló egység az intonációs frázison belül, majd a szavak artikulációs sebessége csökken a vége felé.



2. ábra

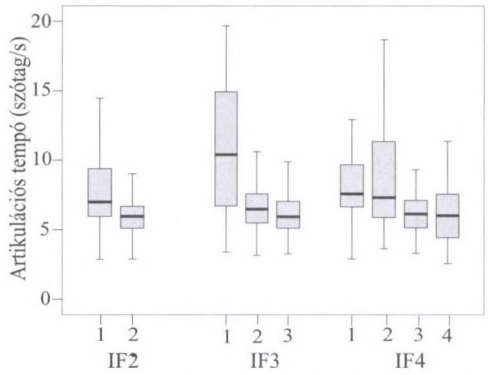
A szavak artikulációs tempójának eloszlása az intonációs frázisokon belül (felül: kétszavas intonációs frázisok, középen: háromszavas intonációs frázisok, alul: négyszavas intonációs frázisok)

A második hipotézisünk teszteléséhez a kétszavas, háromszavas és négyszavas intonációs frázisokban adatolt 1453 szót használtuk fel. A GLM analízis eredményeit a 2. táblázatban összegeztük. A statisztikai elemzésbe öt tényezőt vontunk be: a szót felépítő szótagok száma; a szót tartalmazó intonációs frázis pozíciója a beszédszakaszon belül; a szó pozíciója az intonációs frázison belül; a szó szófaja; illetve, hogy a szó tartalmas szó vagy funkciószó. Mindhárom modellünk sikeresnek mondható, mivel a kétszavas intonációs frázisok esetén 80,2%-ban, a háromszavasok esetén 77,9%-ban, míg a négyszavasok esetén 63,4%-ban az elemzésbe bevont öt tényezővel magyarázható az artikulációs tempó variabilitása. A szót tartalmazó intonációs frázis helye a beszédszakaszon belül egyik modell esetében sem szignifikáns, a háromszavas modellek esetén a szó helye sem szignifikáns tényező. A η^2 értékei azt mutatják, hogy mindhárom modell esetén a legerősebb tényező a szófaj. További elemzéseket igényel azonban, hogy ez az eredmény milyen mértékben az egyenlőtlen eloszlásnak, illetve a különböző szófajú szavak átlagos szótagszámának köszönhető.

2. táblázat: A GLM analízis eredményei a két, három és négy szóból felépülő intonációs frázisokra

Kétszavas IF	*df*	*F*	*p*	*Adj.* R^2	η^2
Modell	26	80,243	<,001	80,2%	
Szóhossz	1	13,628	<,001		,026
IF helye	1	5,001	=,026		,010
Szó helye	1	10,332	<,001		,020
Szófaj	11	73,032	<,001		,609
Tartalmas vs. funkciószó	1	72,904	<,001		,124
Háromszavas IF					
Modell	29	60,437	<,001	77,9%	
Szóhossz	1	11,373	<,001		,024
IF helye	1	4,403	=,036		,010
Szó helye	1	,030	=,863		,000
Szófaj	8	11,886	<,001		,172
Tartalmas vs. funkciószó	1	17,653	<,001		,037
Négyszavas IF					
Modell	28	19,971	<,001	63,4%	
Szóhossz	1	18,019	<,001		,061
IF helye	1	13,691	=,548		,047
Szó helye	1	,362	<,001		,001
Szófaj	9	8,610	<,001		,217
Tartalmas vs. funkciószó	1	11,625	=,806		,040

A 3. ábra dobozdiagramjai a két, három és négy szóból felépülő intonációs frázisok eltérő pozíciójú szavainak artikulációs tempóját ábrázolják. A statisztikai elemzések szerint a különböző pozíciójú szavak tempója között az eltérés szignifikáns [kétszavas intonációs frázis: $F(1,541)= 91,203$; $p < 0,001$; háromszavas intonációs frázis: $F(2,541)= 91,024$; $p <0,001$; négyszavas intonációs frázis $F(3,541)= 11,589$; $p <0,001$].



3. ábra
A szavak artikulációs tempója az intonációs frázisban elfoglalt pozíciójuk alapján (IF1 = egyszavas intonációs frázisok, IF2 = kétszavas intonációs frázisok, IF3 = háromszavas intonációs frázisok)

A kétszavas intonációs frázisokban a második szó (6,01 szótag/s) szignifikánsan lassabb, mint az első (8,59 szótag). A háromszavas intonációs frázisokban is megfigyelhető a lassulási tendencia, az első (10,91 szótag/s) és második (7,18 szótag/s), az első és harmadik (6,21 szótag/s), illetve a második és harmadik szó artikulációs tempója között is szignifikáns a különbség a Tukey-féle post hoc teszt alapján. A négy szóból felépülő intonációs frázisok esetén is a rallentando mintázat a jellemző, bár a második (8,87 szótag/s) és a negyedik szó (6,44 szótag/s) átlagosan valamivel gyorsabb, mint az őket megelőző (az első szó átlagos artikulációs tempója: 8,52 szótag/s, a harmadik szó átlagos artikulációs tempója: 6,41 szótag/s). Ezek a különbségek a Tukey-féle post hoc teszt alapján nem szignifikánsak, a többi pozíció közötti különbség azonban szignifikáns.

Harmadik hipotézisünk tesztelésére elemeztük az artikulációs tempó variabilitásának mintázatát. Az intonációs frázist felépítő szavak artikulációs tempóit hasonlítottuk össze, hogy leírhassuk a lassuló és gyorsuló tendenciákat az intonációs frázison belül. Az eredményeket külön tárgyaljuk a két, három és négy szóból felépülő intonációs frázisok esetén. Az artikulációstempó-

mintázatokat Dankovičová (1997) alapján ikonikusan is ábrázoltuk, de elneveztük őket betűsorokkal is, hogy a későbbiekben könnyebben hivatkozhassunk rájuk. A szavakat az intonációs frázisban elfoglalt helyük alapján jelöltük az ábécé betűivel (első szó jele A, második szó jele B stb.), és az artikulációs tempójuk szerint raktuk őket sorrendbe (az első a leggyorsabb szó betűjele stb.).

A 3. táblázat a két szót tartalmazó intonációs frázisok artikulációstempó-struktúráját mutatja be. Közel azonos arányban jellemző a két szóból álló intonációs frázisokra a lassulás (BA) és a gyorsulás (AB). A nyolc adatközlőből hatnál az AB, vagyis a lassuló tendenciát adatoltuk nagyobb arányban.

3. táblázat: A két szót tartalmazó intonációs frázisok artikulációstempó-mintázatainak aránya adatközlőnként

Beszélők	AB 	BA 
F1	46,15%	**53,85%**
F2	**51,28%**	48,72%
F3	47,50%	**52,50%**
F4	**55%**	45%
N1	**60%**	40%
N2	**51,61%**	48,39%
N3	**53,49%**	46,51%
N4	46,88%	**53,12%**
Összesen	**51,60%**	48,40%

A három szóból felépülő intonációs frázisoknál hatféle artikulációstempó-mintázatot különítettünk el (4. táblázat). Ezek közül az ACB mintázat (37,84%) volt a legjellemzőbb, de közel azonos arányban fordultak elő a CBA mintázatú (31,35%), vagyis a folyamatosan gyorsuló intonációs frázisok is. A hat mintázatból négynél (az intonációs frázisok 62,70%-a) – ABC, BAC, BCA, ACB – az első vagy a második a szó a leggyorsabb. Az adatközlők háromegyedénél az ACB mintázatot adatoltuk a legnagyobb arányban, vagyis az első szó a leggyorsabb, a második szó a leglassabb, majd az utolsó szónál gyorsulás figyelhető meg. Két adatközlőnél az egész frázison végigvonuló gyorsulási tendencia (CBA) volt a legjellemzőbb.

A négy szóból álló intonációs frázisoknál 19 típusba csoportosíthatók az artikulációstempó-mintázatok (5. táblázat). A leggyakrabban realizálódó mintázat a BADC (17,86%). Kilenc típusnál (ABCD, BACD, BCAD, ACBD,

BCDA, ABDC, ACDB, BADC, ADCB) az intonációs frázis első vagy második szava a leggyorsabb, vagyis az összes intonációs frázis 76,19%-ánál megfigyelhető ez a tendencia.

4. táblázat: A három szót tartalmazó intonációs frázisok artikulációstempó-mintázatainak aránya adatközlőnként

Beszélők	ABC 	BAC 	BCA 	ACB 	CAB 	CBA 
F1	13,64%	4,55%	9,09%	22,73%	9,09%	**40,91%**
F2	4,00%	12,00%	–	32,00%	16,00%	**36,00%**
F3	5,56%	11,11%	11,11%	**44,44%**	–	27,78%
F4	7,69%	19,23%	3,85%	**38,46%**	7,69%	23,08%
N1	10,00%	10,00%	5,00%	**40,00%**	5,00%	30,00%
N2	8,70%	8,70%	8,70%	**39,13%**	4,35%	30,43%
N3	5,88%	5,88%	–	**47,06%**	–	41,18%
N4	–	18,75%	9,34%	**43,75%**	3,13%	25,00%
Összesen	6,49%	12,43%	5,95%	**37,84%**	5,95%	31,35%

Az adatközlőkre jellemző mintázatokat is elemeztük. Mind a négy férfi adatközlőnél kilencféle, az N1 és az N2 adatközlőknél nyolcféle, az N3 adatközlőnél hatféle és az N4 adatközlőnél ötféle mintázatot adatoltunk. Az F2 és az N1 beszélőknél a BADC, az N2 beszélőnél az ABCD és az N4 beszélőnél a BACD mintázat jelent meg a legnagyobb arányban. Az F1 beszélőnél a BACD és a BADC, az F4 beszélőnél a BACD és az ACDB mintázatok voltak a leggyakoribbak. Az F3 és az N3 beszélőknél az előforduló kilencféle, illetve hatféle mintázat egyenlő arányban volt adatolható.

Dankovičová (1997) a rallentandot a következő kritériumokkal definiálja a négy szóból álló intonációs frázisok esetén: az intonációs frázis első vagy második szava a leggyorsabb; az egymást követő szavak között nincs egynél többször gyorsulás. Az accelerando (folyamatos gyorsulás) ennek az ellenkezője, vagyis az utolsó előtti vagy utolsó szó a leggyorsabb tempójú; az egymást követő szavak között nincs egynél többször lassulás. Ezeknek a kritériumoknak megfelelően csoportosítottuk a 19 artikulációstempó-mintázatot.

A rallentando első kritériumának megfelelő kilenc mintázat közül hét megfelel a második kritériumnak is (ABCD, BACD, BCAD, ACBD, BCDA, ABDC, ACBD), vagyis a négy szóból álló intonációs frázisok 53,57%-ára jellemző a lassulási tendencia. Öt típusra (CBAD, CBDA, CDBA, DCAB, DBCA), vagyis az intonációs frázisok 7,14%-ára az accelerando, vagyis az

artikulációs tempó folyamatos gyorsulása a jellemző. Hét típusnak (CADB, CABD, BADC, ADCB, DABC, CDAB, DACB) nincs olyan közös jellemzője, amely alapján egy csoportba sorolhatnánk őket.

5. táblázat: A négy szót tartalmazó intonációs frázisok artikulációstempó-mintázatainak aránya adatközlőnként

Beszélők	ABCD 	BACD 	BCAD 	ACBD 	BCDA 
F1	–	**18,18%**	9,09%	–	9,09%
F2	8,33%	–	–	–	8,33%
F3	**11,11%**	**11,11%**	**11,11%**	**11,11%**	**11,11%**
F4	9,09%	**18,18%**	9,09%	–	–
N1	20,00%	6,67%	–	13,33%	6,67%
N2	**28,57%**	–	21,43%	–	7,14%
N3	–	–	**16,67%**	–	**16,67%**
N4	–	**33,33%**	–	16,67%	16,67%
Összesen	11,90%	9,52%	8,33%	4,76%	8,33%

	ABDC 	ACDB 	CBAD 	CBDA 	CDBA 
F1	–	9,09%	9,09%	–	–
F2	8,33%	8,33%	–	–	8,33%
F3	–	–	–	–	–
F4	9,09%	**18,18%**	–	9,09%	–
N1	–	–	–	–	–
N2	–	–	–	7,14%	–
N3	**16,67%**	**16,67%**	–	–	–
N4	–	16,67%	–	–	–
Összesen	3,57%	7,14%	1,19%	2,38%	1,19%

	DCAB 	DBCA 	CADB 	CABD 	BADC 
F1	–	–	9,09%	–	**18,18%**
F2	8,33%	–	–	–	**25,00%**
F3	–	–	**11,11%**	–	**11,11%**
F4	–	–	–	9,09%	9,09%
N1	–	6,67%	–	–	**33,33%**
N2	–	–	–	7,14%	7,14%
N3	–	–	–	–	**16,67%**
N4	–	–	–	–	16,67%
Összesen	1,19%	1,19%	2,38%	2,38%	**17,86%**

	ADCB 	DABC 	CDAB 	DACB
F1	–	–	9,09%	9,09%
F2	16,67%	–	–	8,33%
F3	**11,11%**	**11,11%**	–	–
F4	9,09%	–	–	–
N1	6,67%	–	–	6,67%
N2	–	–	7,14%	14,29%
N3	–	–	–	**16,67%**
N4	–	–	–	–
Összesen	5,95%	1,19%	2,38%	7,14%

Következtetések

A jelen tanulmányunkban a szavak artikulációstempó-variabilitását elemeztük felolvasások intonációs frázisaiban. Magyar nyelvre is igazoltuk, hogy az artikulációs tempó bizonyos mintázatokat követ (vö. Dankovičová 1997, 1999).

Számos tényező befolyásolhatja ezt a variabilitást, ezek közül jelen kutatásunkban ötöt vizsgáltunk: a szó pozíciója az intonációs frázison belül; az intonációs frázis pozíciója a beszédszakaszon belül; a szót felépítő szótagok száma; a szó szófaja; szóosztály (tartalmas vs. funkciószó). A statisztikai elemzések szerint nagymértékben magyarázható ezzel az öt tényezővel a tempóvariabilitás a két, három és négy szóból felépülő intonációs frázisok esetén. Az intonációs frázis beszédszakaszbeli pozíciója, és a háromszavas frázisokban a szó helye az intonációs frázison belül nem befolyásolja szignifikánsan a variabilitást, a többi tényező azonban igen.

Dankovičová (1997) cseh nyelven végzett vizsgálatában kevesebb artikulációstempó-mintázat volt elkülöníthető, mint jelen kutatásunkban. Az eltérés hátterében a nyelvi különbségek mellett módszertani különbségek is meghúzódhatnak. Dankovičová ugyanis az intonációs frázisokat fonológiai szavakra tagolta, melyet a következőképpen definiált: szótagsorozat, melyet a hangsúly határoz meg, és figyelembe veszi a lexémahatárokat. Annak a kérdésnek a megválaszolása, hogy milyen mértékben köszönhető az eltérés nyelvi különbségeknek, további vizsgálatokat igényel.

Az artikulációstempó-mintázatokat, illetve a beszélők artikulációs tempójának összefüggéseit elemezve azt a következtetést vonhatjuk le, hogy a szavak artikulációs tempója az intonációs frázison belül nem önkényesen, hanem rendszerszerűen változik. Ez azzal magyarázható, hogy a beszéd temporális szerveződésének az intonációs frázis az egyik alapegysége. Mindhárom típusú intonációs frázis esetén a rallentando tendencia volt a legjellemzőbb. Az intonációs frázisok többségénél adatolható volt az egység vége felé a lassulás, amelyet a hallgatók felhasználhatnak a beszéd szegmentálása során.

Az intonációs frázisokon belül az első pozícióban lévő szavak magas szótag/s értékei feltételezhetően a funkciószavak redukált ejtésformáinak köszönhető, ami további kérdéseket vet fel a szószintű tagolás módszertanában.

További kutatásaink során tervezzük más faktorok bevonását az elemzésbe, mint például a szintaktikai jellemző; az igétől való távolság stb. Ezek elemzésével pontosíthatjuk modellünket. További megválaszolandó kérdés, hogy az artikulációs-tempó modellezhető-e ezekkel a módszerekkel a magyar spontán beszédben is.

A kutatás eredményei felhasználhatók a beszédszintézisben, illetve a beszédfelismerésben, a dadogók és a hadarók beszédterápiájában, és általában a beszédtechnikában.

Irodalom

Andó Éva 2002. *A történetmondás kommunikatív jellemzői.* PhD-értekezés. ELTE, Budapest.

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

Beke András 2008. A felolvasás és a spontán beszéd alaphangszerkezeteinek vizsgálata. *Beszédkutatás 2008.* 93–107.

Beke András – Gráczi Tekla Etelka 2010. A magánhangzók semlegesedése a spontán beszédben. In: Navracsics Judit (szerk.): *Nyelv, beszéd, írás 1. – Pszicholingvisztikai tanulmányok 1.* Tinta Könyvkiadó, Budapest. 57–64.

Boersma, Paul – Weenink, David 2012. *Praat: doing phonetics by computer.* (Version 5.3.32.). http://www.fon.hum.uva.nl/praat/download_win.html

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

Bóna Judit 2011. A különböző beszédstílusok az akusztikai-fonetikai és a percepciós vizsgálatok tükrében. *Alkalmazott Nyelvtudomány* XI/1–2. 39–48.

Dankovičová, Jana 1997. The domain of articulation rate variation in Czech. *Journal of Phonetics* 25. 287–312.

Dankovičová, Jana 1999. Articulation rate variation within the intonational phrase in Czech and English. In: *Proceedings of the 14th International Congress of Phonetic Sciences.* San Francisco. 269–272.

Fónagy, Iván – Magdics, Klára 1960. Speed of utterances in phrases of different length. *Language and Speech* 3. 179–193.

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

Gocsál Ákos 2001. Gyorsabban beszélnek-e a nők, mint a férfiak? *Beszédkutatás 2001.* 61–72.

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

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

Gósy Mária 2000. A beszédszünetek kettős funkciója. *Beszédkutatás 2000.* 1–14.

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

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

Gósy Mária – Beke András – Horváth Viktória 2011. Temporális variabilitás a spontán beszédben. *Beszédkutatás 2011*. 5–30.

Gósy Mária – Bóna Judit 2011. Beszédfolyamatok monitorozása. *Magyar Nyelvőr* 135. 393–414.

Gósy Mária – Horváth Viktória 2007. Fonetikai elemzések a spontán beszédben: alapok, kihívások. *Beszédkutatás 2007*. 7–18.

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

Gussenhoven, Carlos 2004. *The phonology of tone and intonation*. Cambridge University Press, Cambridge.

Hunyadi László 2002. *Hungarian sentence prosody and Universal Grammar*. Peter Lang, Frankfurt am Main–Berlin–Bern–Bruxelles–New York–Oxford–Wien.

Imre Angéla 2005. Különböző műfajú szövegek szupraszegmentális jellemzői. *Magyar Nyelvőr* 129. 510–520.

Jacewicz, Ewa – Fox, Robert Allen – O'Neill, Caitlin – Salmons, Joseph 2009. Articulation rate across dialect, age, and gender. *Language Variation and Change* 21. 233–256.

Jacewicz, Ewa – Fox, Robert Allen – Wei, Lai 2010. Between-speaker and within-speaker variation in speech tempo of American English. *Journal of the Acoustical Society of America* 128. 839–850.

Laan, Gitta P. M. 1997. The contribution of intonation, segmental durations, and spectral features to the perception of a spontaneous and a read speaking style. *Speech Communication* 22. 43–65.

Laczkó, Mária 1991. The interrelation of articulation rate and pauses in children's speech. In: Gósy, Mária (ed.): *Temporal factors in Speech. A collection of papers*. HAS Research Institute for Linguistics, Budapest. 139–151.

Laczkó, Mária 1993. A tempó és a szünet viszonya a hangos olvasásban. *Beszédkutatás '93*. 185–93.

Levelt, Willem J. M. 1999. Producing spoken language: A blueprint of the speaker. In Brown, Colin M. – Hagoort, Peter (eds.): *The neurocognition of language*. Oxford University Press, Oxford. 83–122.

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 2005. *A spontán beszéd néhány szupraszegmentális jellegzetessége. Monologikus és dialogikus szövegek összevetése, valamint a hümmögés vizsgálata*. Phd-disszertáció. ELTE, Budapest.

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

Markó Alexandra – Bóna Judit 2012. Eltérő beszédmódok intonációs sajátosságai fiatal és idős korban. In Balázs Géza – Veszelszki Ágnes (szerk.): *Nyelv és kultúra – kulturális nyelvészet*. Magyar Szemiotikai Társaság, Budapest. 253–258.

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

Menyhárt Krisztina 2010. A beszédsebesség objektív mérései és szubjektív észlelési eredményeinek összefüggései mai és 60 évvel ezelőtti beszélőknél. *Beszédkutatás 2010*. 110–124.

Olaszy Gábor 2000. Kísérlet a magyar beszédhangok specifikus időtartamainak meghatározására folyamatos beszédre. *Beszédkutatás 2000*. 26–38.

Olaszy Gábor 2006. Prozódiai szerkezetek jellemzése a hírfelolvasásban, a mesemondásban, a novella és a reklámok felolvasásában. *Beszédkutatás 2006*. 21–50.

Oyer, Herbert J. – Deal, Leo V. 1985. Temporal aspects of speech and the aging process. *Folia Phoniatrica* 37. 109–112.

Quené, Hugo 2004. What is the Just Noticeable Difference for tempo in speech? In Quené, Hugo – Heuven, V. J. (Eds.): *On speech and language: Studies for Sieb G. Nooteboom*. Netherlands Graduate School of Linguistics, Utrecht. 149–158.

Quené, Hugo 2007. On the just noticeable difference for tempo in speech. *Journal of Phonetics* 35. 353–362.

Quené, Hugo 2008. Multilevel modeling of between-speaker and within-speaker variation in spontaneous speech tempo. *Journal of the Acoustical Society of America* 123. 1104–1113.

Roca, Iggy – Johnson, Wyn 2005. *A course in phonology*. Blackwell, Oxford.

Savy, Renata – Voghera, Miriam 2010. A corpus-based study on syntactic and phonetic prosodic phrasing boundaries in spontaneous Italian speech. In: *Proceeding of Speech Prosody 2010*. Chicago. 1892–1896.

Subosits István 1990. Adatok az életkor és a beszédtempó összefüggéséhez egy eszközfonetikai vizsgálat alapján. *Egyetemi Fonetikai Füzetek* 3. 159–167.

Szaszák György – Beke András 2012. Statisztikai módszerek alkalmazása a szintaktikai szerkezet és a beszédjel prozódiai szerkezetének feltérképezéséhez olvasott és spontán beszédben. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest. 236–250.

Tsao, Y. Ch. – Weismer, G. 1997. Interspekaer variation in habitual speaking rate: Evidence for a neuromuscular component. *Journal of Speech, Language and Hearing Research* 40. 858–866.

Váradi Viola 2010. A felolvasás és a spontán beszéd temporális sajátosságainak összehasonlítása. *Beszédkutatás 2010*. 100–109.

Váradi Viola 2012. A felolvasás és a spontán beszéd tagolhatósága prozódiai egységek jellemzése alapján. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest. 108–120.

Varga László 1994. A hanglejtés. In Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 2. Fonológia*. Akadémiai Kiadó, Budapest. 468–549.

Verhoeven, Jo – De Pauw, Guy – Kloots, Hanne 2004. Speech rate in a pluricentric language: A comparison between Dutch in Belgium and the Netherlands. *Language and Speech* 47. 297–308.

Wacha Imre 1999. *Szöveg és hangzása*. Kodolányi János Főiskola, Székesfehérvár.

KÉRDŐ MEGNYILATKOZÁSOK A SPONTÁN BESZÉDBEN

Markó Alexandra

Bevezetés

A kérdő megnyilatkozások a szociális interakciók minden területén jelen vannak, ennek köszönhetően a nyelvészeti elemzések minden szintjén vizsgálják ezeket (Freed 1994). A prozódiai, szintaktikai, logikai, szemantikai, pragmatikai vagy pszicholingvisztikai sajátosságok leírása számos tudományos mű témája (például Austin 1962/1990; Searle 1969; Schegloff–Sacks 1973; Hymes 1974; Fishman 1978; Kiefer 2000; Vion–Colas 2006). Bár megegyezés van a szerzők között abban a tekintetben, hogy a kérdő megnyilatkozások azonosítható szintaktikai szerkezettel, sajátos prozódiával és szemantikai vagy propozicionális tartalommal rendelkeznek, amely lényegében független azoktól a változatos funkcióktól, amelyeket betöltenek a társas interakciókban, abban még sincs teljes egyetértés, hogy hogyan határozzuk meg, osztályozzuk és elemezzük a kérdéseket (a különböző megközelítések szakirodalmi áttekintését lásd pl. Freed 1994).

Az interakciós megközelítés szerint a kérdések elemzésében figyelemmel kell lennünk a kontextusra, amelyben megjelennek (Rossano 2010). Habár a kérdő megnyilatkozások felismerhetők pusztán formális nyelvi jegyek alapján is (pl. szintaxis, prozódia, lexikai elemek), ez nem jelenti azt, hogy a társalgásban ne tölthetnék be a kérdés funkcióját olyan megnyilatkozások, amelyek nem rendelkeznek ilyen formális jegyekkel (például az *Azt szeretném kérdezni, hogy...* kezdetű megnyilatkozás formailag nem kérdés, hanem állítás). Mindez ellenkezőleg is igaz: az, hogy egy megnyilatkozás a kérdés formai jegyeit hordozza, nem jelenti feltétlenül azt, hogy ez a közlés valóban kérdés (a választ váró vagy információt kérő megnyilatkozás értelmében). Például az *Ide tudod adni a sót?* esetében a beszélő nem arra kíváncsi, hogy a beszédpartnere képes-e megtenni az adott cselekedetet, hanem közvetett módon alkalmazza a kérés beszédaktusát (Searle 1969/1991). Ha a partner válaszol a kérdő formájú megnyilatkozásra, azzal megsérti az együttműködési alapelvet (Grice 1975/1991), hiszen az adott kérdésbe bújtatott kérésre a só átadása az adekvát reakció. Az interakciós megközelítés szerint tehát „valódi" kérdés az, amelyre választ vár és kap a beszélő.

Freed (1994) információs kontinuum mentén helyezi el a különböző kérdéstípusokat. Taxonómiájában 16 különböző funkcionális kategóriát különböztet meg a publikus információt kérő kérdéstől (amely a külvilágra vonat-

kozik, és új, tényszerű információt vár) a függő beszédig, azaz az idézetként elhangzott kérdésig. A két végpont között olyan kategóriák szerepelnek, mint az invitálás vagy ajánlat (pl. *Kérsz egy kávét?*), az információ tisztázására, illetve megismétlésére vonatkozó kér(d)és (pl. *Tessék?*), a fatikus kérdés (pl. *Tudod?*), a didaktikus, a retorikai funkció és így tovább. A Freed által megállapított 16 típus a korábbi szakirodalomban (pl. Kearsley 1976) meghatározott négy funkcionális osztályba sorolható be: tényekre (a társalgáson kívüli világ tényeire) vonatkozó kérdések, magára a konverzációra vonatkozó kérdések, a beszélő és a hallgató közötti verbális és társas viszonyra vonatkozó kérdések, valamint expresszív stílusú kérdések (inkább közvetítenek információt, mintsem kérnek). Freed (1994) és néhány más szerző (pl. Kearsley 1976) szerint a kérdések formája és funkciója között kimutatható kapcsolat.

Logikai szemantikai szempontból három fő kérdéstípust különít el a szakirodalom: 1. kiegészítendő kérdés, 2. eldöntendő kérdés, 3. választó kérdés. A kiegészítendő kérdések olyan kérdő mondatok, amelyek kérdő névmást tartalmaznak, és a rájuk adott válasz fókuszához tartozó fogalom ugyanabba az ontológiai kategóriába tartozik, mint a kérdőszóé (pl. a *ki?* kérdésre személyreferencia, a *hol?* kérdésre helyreferencia a válasz). A nyitott kérdéseket a kiegészítendő kérdés altípusának tekintik, amely szintaktikailag azonos vele, de szemantikai szempontból különbözik, hiszen bővebb kifejtést (nem csupán egy szót) vár válaszként (pl. *Hogyan oldjuk meg ezt a problémát?*) (Kiefer 2000; Kugler 2000). Prozódiai szempontból mind a kiegészítendő, mind a nyitott kérdés elöl eső karakterdallammal realizálódik (Deme 1962; Varga 2002a).

Az eldöntendő kérdés megerősítést vagy tagadást vár. A kérdést többféle nyelvi eszköz fejezheti ki, a magyarban kérdő intonáció (szökő-eső zárlat) vagy kérdő partikula (*-e, ugye*) jelzi, amelyek kizáró disztribúcióban vannak egymással Kugler (2000). A kérdő partikulát tartalmazó eldöntendő kérdések azonban – bár kérdő intonációval (szökő-eső zárlat) nem jólformáltak – szintén jelöltek prozódiailag (Fónagy–Magdics 1967; Gyuris 2009).

A választó kérdést sok szerző különálló kategóriának tartja (pl. Stivers 2010; Rossano 2010), míg mások az eldöntendő kérdés altípusaként tárgyalják (pl. Freed 1994; Kiefer 2000; Kugler 2000). A magyarban a választó kérdés szintaktikailag szinte azonos az eldöntendővel (a *vagy* kötőszó hozzáadásával; Kugler 2000), ugyanakkor prozódiai szempontból különbözőek (lásd Olaszy 2002; Varga 2002a).

2010-ben a *Journal of Pragmatics* 42. számában egy olyan átfogó kutatásról számolnak be, amely 10 nyelvben (Európából, az USA-ból, Délkelet Ázsiából, Mexikóból, Namíbiából és Pápua Új-Guineából) vizsgálta a kérdéseket és a rájuk adott válaszokat egy közös módszertan segítségével. Bár a kutatás alapvetően pragmatikai jellegű volt, természetszerűleg több más tényezőt is figyelembe vettek a szerzők, a logikai szemantikai struktúrától kezdve a morfoszintaktikai sajátosságokon keresztül a prozódiáig. A kutatás-

ban nyelvenként mintegy 350 darab, informális társalgásból származó kérdés-válasz párt elemeztek. Csak ún. funkcionális (értsd: választ vagy megerősítést váró) kérdéseket vizsgáltak (az erre vonatkozó módszertani kritériumokat lásd alább).

Magyarra eddig Schirm (2007) végzett pragmatikai megközelítésű elemzést a kérdésekre vonatkozóan. Tanulmányában áttekintette a szemantikai, logikai, retorikai kérdéstipológiákat, és részletesen ismertette a diskurzuselemzés szakirodalmának ilyetén vonatkozásait. Felhívta a figyelmet a kontextus vizsgálatának szükségességére a pragmatikai kérdésfunkciók meghatározásában, és új kérdéstipológiát állított fel. Vizsgálatának korpuszát többségében félintézményes vitaműsorok alkották, vagyis spontán-félspontán kérdések pragmatikai elemzését végezte el a szerző, amelyben szükség esetén prozódiai szempontokat is figyelembe vett.

A kérdő megnyilatkozások elemzése a magyar nyelvészeti szakirodalomban látszólag kevés nyitott kérdést tartogat, hiszen mind szintaktikai, szemantikai, fonológiai, mind pragmatikai, fonetikai szempontból születtek áttekintő munkák. Ugyanakkor a spontán beszéd kérdő megnyilatkozásaival, ennek a témakörnek a korpuszalapú feltérképezésével kevés kutatás foglalkozott eddig.

A 20. század fonetikusai nagy érdeklődéssel fordultak a beszéddallam kérdései felé. A tudományos igényű intonációkutatás Tolnai Vilmos (1915) nevéhez fűződik. A következő fontos lépést a szamosháti nyelvjárás hanglejtésformáinak leírásával Csűry Bálint (1925) tette meg. Hat kérdésformát különített el, beleértve az összetett mondatokat is. Eredményei azért is különösen fontosak, mert dialektológiai vizsgálatai a spontán beszédet állították a középpontba. Magyar anyagon Hegedűs Lajos (1930) végzett először műszeres beszéddallam-vizsgálatot (kimográffal).

A mai magyar nyelv rendszeré-ben Deme a mondattani fejezethez kapcsolódva mutatta be a mondatfajták jellemző hanglejtésformáit (1962). Fónagy és Magdics (1963 és 1967) a kérdések dallamát több szempontú elemzésnek vetette alá: vizsgáltak semleges és attitűd/érzelemjelző formákat, valamint kitértek a nemek és a generációk közötti nyelvhasználati különbségekre is. Megállapításaikat főként felolvasott és (színészek által) eljátszott közlések meghallgatására, lejegyzésére és lekottázására alapozták. Bár olvashatunk spontán beszédbeli megvalósulásokról is, a spontán kérdések rendszerszerű vizsgálatáról nem, illetőleg sok esetben nem világos, hogy eljátszott vagy spontán megnyilatkozásra vonatkoznak-e a leírtak.

Bartók (1978) a kérdő intonációt a magyar mint idegen nyelv oktatása szempontjából jellemezte.

A leírások többsége a kérdő mondatokból és azok funkcióiból indult ki, és az ezekhez tartozó lehetséges hanglejtésformákat vette számba. A *Nyelvművelő kézikönyv*-nek a hanglejtéssel foglalkozó címszava (Elekfi 1980) – érte-

lemszerűen – kiegészült az ún. hanglejtési hibák tárgyalásával is. Elekfi és Wacha (é. n. [2003]) munkája ugyancsak kitért a helytelen megvalósulásokra.

Varga László funkcionális vizsgálatainak (pl. 1993, 1994, 2002a) a célja a magyar dallamprozodémák teljes rendszerének leírása volt. Az eddig említett munkákkal ellentétben ezért (fonológiai szempontú) tanulmányaiban a formából indult ki, s az egyes karakterdallamok funkcióit – köztük a kérdőt – vette sorra.

Fónagy 1998-as áttekintésében összegezte többek között saját korábbi és Varga (1981 és 1983) eredményeit. Itt is olvashatunk megállapításokat a spontán beszédről, de ezek közül sok ugyancsak nem választható el a szótagrabok meghangosításával kapcsolatos eredményektől; illetőleg némelyik állítás túlzónak tűnik (pl. hogy a nők közötti társalgásban a kiegészítendő kérdések végi dallamemelés neutrális formának tartható).

A kutatás újabb irányvonalát jelentik a műszeres kísérleti fonetikai elemzések. Elsősorban percepciós nézőpontból közelítette meg a kiegészítendő kérdések dallamának sajátosságait Gósy (1993). Maga az elemzés nem spontán beszédanyagot vett alapul, hanem beszédszintézissel előállított változatokat tesztelt, de a kutatást a spontán beszéd motiválta. Ugyancsak szintézissel manipulált anyaggal végzett percepciós vizsgálat adta meg a választ a Gósy és Terken (1994) által megfogalmazott kérdésre: milyen prozódiai jegyek teszik lehetővé a két szótagú szóra végződő eldöntendő kérdések és (az utolsó szavukon hangsúlyos) állítások megkülönböztetését – tehát az ugyanazon szótagon megjelenő F_0-csúcs ellenére milyen járulékos paraméterek segítik az azonosítást. Gósy 2000-ben megjelent analízise produkcióközpontú volt: kérdő, felkiáltó és felszólító mondatok dallamát elemezte 12 pedagógus bemondásában.

Olaszy (1995) először végzett komplex akusztikai vizsgálatot, azaz a beszéddallam mellett az intenzitásról és a temporális szerkezetről, illetőleg az ezekben végbemenő változásokról is számot adott különböző beszédaktusok, s köztük a kérdések tekintetében. 2002-es tanulmánya részletezőbb volt, a kérdés több altípusát tárgyalta – az alaphangmagasság és az intenzitás tekintetében. Mindkét tanulmány felolvasott közlések adatait mutatta be.

Az úgynevezett laboratóriumi fonológia irányzatába illeszkedik Varga (2002b) elemzése, amelyben az egy szótagú eldöntendő kérdő mondat „allokontúrjait" vizsgálta – minidialógusokban elicitált kérdések alapján. Mycock (2007, 2010 stb.) szintén (olvasott) minidialógusokban elemezte a magyar kiegészítendő kérdéseket, lexikális funkcionális nyelvtani keretben végzett vizsgálata prozódiai vonatkozásokra is kitért. A kutatás kontrasztív szempontokat is figyelembe vett.

Egy korábbi kutatásban (Markó 2007) mintegy kétórányi társalgást rögzítettünk négy fiatal beszélő (2 nő, 2 férfi) részvételével. Az ebben adatolható (zajmentes) közel kétszáz kérdő megnyilatkozás elemzésének eredményeit összegezve azt állapíthatjuk meg, hogy – ahogy vártuk – a spontán megvaló-

suló kérdések sokszínűbbek, mint amilyennek a felolvasott, elicitált vagy eljátszott mintapéldáik alapján a szakirodalom lefesti őket. A változatosság elsősorban az F_0 modulációjának lehetőségeiben, másrészt pedig a hangköz értékeiben mutatkozott meg. Kérdésként merült fel, hogy bizonyos jellegű vagy mértékű eltérések indokolják-e további fonetikai altípusok felvételét.

Keszler Borbála (2009) a kérdés- és felelettípusok spontán beszédbeli előfordulásait elemezte Kiefer (2000) szemantikai kategória-rendszere alapján. 205 kérdés-felelet párt vizsgált 13 beszélt nyelvi szövegben, amelyeket 1976-ban rögzítettek rejtett mikrofonnal, kollégiumban (Beszélt nyelvi gyűjtemény). Tanulságként azt fogalmazta meg, hogy „a spontán beszéd vizsgálata nagyon sok átmeneti esettel és grammatikai érdekességgel szolgál, s ezek ismerete sok esetben nem megkönnyíti, hanem megnehezíti a nyelvtani kategorizálást" (19).

A jelen kutatásban nagyobb mennyiségű spontán társalgásban elemezzük a kérdő megnyilatkozásokat. Az alkalmazott megközelítés a prozódiai megvalósulást állítja a középpontba, de figyelembe vesz más (funkcionális és formai) szempontokat is. Hipotézisünk szerint a kérdések logikai szemantikai típusa, pragmatikai funkciója és prozódiai megvalósulása között összefüggés mutatható ki: a valódi kérdések esetében jellemzőbbek a prototipikus megvalósulások. Feltételezzük, hogy a kérdő megnyilatkozások prozódiai megvalósulása változatosságot mutat, és az egyes, a prototipikustól eltérő formák megjelenése nem függ a beszélő nemétől és/vagy életkorától.

Kísérleti személyek, anyag és módszer

Az elemzett hanganyag 30 társalgás volt, amelyeket a BEA adatbázisból (Gósy et al. 2012) válogattunk össze. A BEA protokollja szerint a társalgásfelvételeken az adatközlőn és a felvételvezetőn kívül egy további személy vesz részt. A téma változó, az élet mindennapjaihoz kapcsolódik (pl. ünnepek, gyermeknevelés, társadalmi kérdések, aktualitások). Feltételeztük, hogy a társalgás helyzetéből adódóan az adatközlőktől nagyszámú kérdő megnyilatkozást adatolhatunk. Annál is inkább, mivel az adatbázis fejlesztői szerint „a társalgás az a része a protokollnak, amelynek során az adatközlők a legjobban feloldódnak" (Gósy et al. 2012: 12).

A BEA adatközlői közül három életkori csoportba tartozó, korosztályonként azonos számú (5-5) nőt és férfit, összesen tehát 30 beszélőt választottunk ki. Mivel a korosztályi megoszlás természetszerűen nem teljesen egyenletes a BEA-ban, ezért az életkori határok nemek szerint eltérőek lehettek. Az 1. táblázat mutatja a beszélők életkori megoszlását nemenként.

A 30 társalgás összes időtartama meghaladja a 10 órát. Ebből pusztán az adatközlők beszéde 290,5 perc, azaz megközelítőleg 5 óra. Az egyes társalgásfelvételek időtartama 6 perc 42 másodperc és 73 perc 26 másodperc között változik, átlagosan 20 perc 50 másodperc. Ebből az adatközlők beszéd-

ideje 2 perc 38 másodperctől 45 perc 12 másodpercig terjed, átlagosan 9 perc 42 másodperc.

1. táblázat: A beszélők életkori megoszlása nemenként

	Nők		Férfiak	
	Szórás (év)	Átlag (év)	Szórás (év)	Átlag (év)
Fiatalok	21–24	22,2	21–24	22,0
Középkorúak	39–46	43,6	39–41	40,0
Idősebbek	54–60	57,4	57–66	61,0

A felvételeken felcímkéztük és lejegyeztük a kérdő megnyilatkozásokat. A BEA sajátosságai miatt (a felvételvezető és a társalgás harmadik résztvevője a BEA fejlesztői közül kerül ki, akiknek a beszéde emiatt erősen felülreprezentált az anyagban) csak az adatközlő által megfogalmazott kérdéseket vettük figyelembe, illetve ezek közül is csak azokat a példányokat, amelyek alkalmasak voltak a prozódiai elemzés elvégzésére (nem fedte el őket egyszerre beszélés, nevetés vagy más zaj). A felcímkézett kérdő megnyilatkozásokat a következő szempontok szerint csoportosítottuk: 1. logikai szemantikai szerkezet (eldöntendő, kiegészítendő vagy választó kérdés); 2. funkció (a kérdésre választ vár a beszélő, vagy ettől eltérő szerepben alkalmazza); 3. beszéddallam (a szakirodalomban leírt intonációs sajátosságokkal rendelkezike, vagy eltér ezektől, és ha eltér, miben).

A funkcionális csoportosításhoz természetesen az egyes kérdő megnyilatkozások kontextusát vizsgáltuk meg. A logikai szemantikai szerkezet tekintetében először az összes kérdéspéldányt elemeztük, majd különválasztva csak az információkérő (választ váró), illetve a nem információkérő (választ nem váró) kérdések szerkezeti megoszlását. A funkcionális elemzésben a Stivers és Enfield (2010) által bemutatott módszert követtük. Így például a *Tényleg? Igen?* típusú visszakérdezéseket valódi (információkérő) kérdésként interpretáltuk, mivel megerősítést várnak, nem tekintettük azonban valódi kérdésnek a történet részeként elhangzott, idézett kérdéseket, mert az adott kommunikációs helyzetben nem igényeltek sem választ, sem megerősítést.

A kérdések csoportosítását a logikai szemantikai és a funkcionális szempontokat együttesen alkalmazva is elvégeztük, és vizsgáltuk a kérdő megnyilatkozások gyakoriságát e tényezők szerint. Az intonációelemzés adatait ugyancsak e csoportosítás függvényében ismertetjük.

A kérdő funkciójú megnyilatkozások hangtani jellemzőit tekintve rögzítettük a zönge minőségét (kváziperiodikus, irreguláris vagy leheletes zönge, a módszerre lásd Dilley et al. 1996), valamint a hanglejtésformát. A kváziperiodikustól eltérő zöngeminőséggel realizált formákat kizártuk a további elemzésből, ahogyan az emocionális töltettel rendelkezőket is. A fennmaradó kérdéseket a szerint osztályoztuk, hogy prototipikus (a fonológiai leírásnak megfelelő, vö. Varga 2002a) dallamformával realizálódtak-e. Ennek megfele-

lően azokban az esetekben, amikor az intonáció azért tér el a szintaktikai szerkezet alapján elvárttól, mert a kérdés meghatározott kommunikációs funkciót tölt be, prototipikusnak minősítettük a példányt. Például az ekhókérdések szintaktikai szerkezetüket tekintve a kiegészítendő kérdések közé tartoznak, de intonációs kontúrjukat szökő-eső dallamzárlat jellemzi.

A magyar kiegészítendő kérdések egyik (viszonylag gyakori) realizációs típusáról máig vita folyik a nyelvészetben. Ez a közlésvégi dallamemeléssel záródó típus, amelynek a „szabályos" voltával kapcsolatban eltérő vélemények fogalmazódnak meg (a szakirodalombeli állásfoglalásokat illetően lásd Gósy 1993 és Olaszy 2002). Mivel a magyar beszéddallam fonológiai leírásában Varga László (2002a) variánsként tartja számon (konfliktus kifejezését társítva hozzá funkcióként), a jelen elemzésben prototipikusnak minősítjük ezeket a megvalósulásokat.

Az annotálást és az elemzést az 5.3 verziójú Praat szoftverrel végeztük (Boersma–Weenink 2012).

Eredmények

A 30 beszélő összesen mintegy 5 órányi elemzett beszédanyagában 134 (zajmentes) kérdő megnyilatkozást annotáltunk. Ez átlagban azt jelenti, hogy beszélőnként 4,5 (elemezhető) kérdés hangzott el, azaz átlagosan minden 2,2. percben egy. Az átlag azonban elfedi azt a tényt, hogy természetesen voltak olyan adatközlők (összesen kilencen), akik a társalgásfelvételen egyáltalán nem fogalmaztak meg kérdést, akár több mint 30 percnyi összes beszédidő alatt sem. Ugyanakkor egy fiatal férfi hanganyagában 31 kérdő megnyilatkozást adatoltunk (90 percnyi beszédben), egy idősebb férfi anyagában pedig 20-at (mintegy 30 perc alatt). Mindebből adódóan nem mutatható ki összefüggés a beszédidő hossza és az elhangzott kérdések száma között.

A legtöbb kérdést a fiatal beszélők anyagában adatoltuk (56-ot, a középkorúaknál 41-et, az idősebbeknél 35-öt), és összességében több mint kétszer annyi kérdést találtunk a férfi adatközlők beszédanyagában, mint a nőkében, időarányosan is (2. táblázat). Átlagosan időarányosan a fiatalok és a középkorúak hasonló gyakorisággal fogalmaztak meg kérdéseket (nagyjából 4 percenként), az idősebbek ennél jóval ritkábban (mintegy 6 percenként). Az idősebb adatközlők felénél (2 férfi, 3 nő) egyáltalán nem adatoltunk kérdő megnyilatkozást, a fiatalok és a középkorúak esetében egy-egy férfi és egy-egy nő hanganyagában nem volt elemezhető kérdés.

A kérdések logikai szemantikai és funkcionális sajátosságai

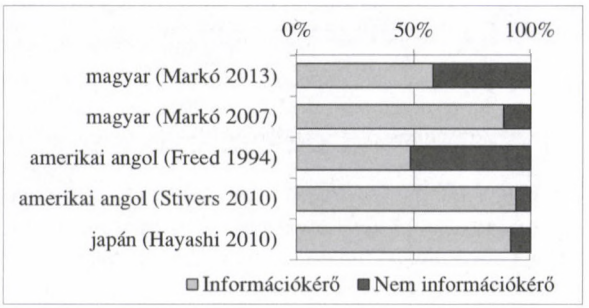
A logikai szemantikai szerkezetet illetően a két fő típus közel egyenlő arányban volt adatolható. A kiegészítendő kérdések aránya 45,52% volt (61 előfordulás), az eldöntendőké 52,24% (70 példány), és összesen 3 (2,24%) választó kérdést találtunk az anyagban.

Interakciós szempontból az információkérő kérdések aránya a teljes anyagban 58,21% (78 példány), azaz alig több mint a kérdések fele esetén

várt a beszélő választ/megerősítést a partnerétől. A nem információkérő kérdések meglepően nagy aránya (56 darab, 41,79%) miatt érdemesnek látszott ezt az eredményt összevetni más korpuszokra kapott adatokkal (1. ábra).

2. táblázat: Az elemezhető kérdések száma és gyakorisága a beszélők életkorának és nemének függvényében

	Nők		Férfiak	
	A kérdések száma (db)	A kérdések gyakorisága (db/perc)	A kérdések száma (db)	A kérdések gyakorisága (db/perc)
Fiatalok	12	0,16	44	0,34
Középkorúak	26	0,22	15	0,33
Idősebbek	4	0,04	31	0,29
Összesen	42	0,14	90	0,32



1. ábra
Az információkérő és a nem információkérő kérdések aránya néhány beszédkorpuszban

A korábban vizsgált, négybeszélős, kétórás magyar nyelvű korpuszban (Markó 2007) 199 kérdés volt elemezhető minőségű (vagyis mintegy másfélszer több, mint a most elemzett, 5 órányi beszédanyagban), ezeknek 88,44%-a volt információkérő kérdés. Freed (1994) amerikai angol korpuszában, amely kétbeszélős, célzott interjúhelyzetben rögzített társalgásokból állt, az 1275 kérdésnek mintegy a fele (51,69%) tartozott a választ nem váró (pl. fatikus, retorikai) kategóriába. Ugyanakkor egy másik amerikai angol kutatásban, amelyben 3–5 beszélős, familiáris helyzetben videóra rögzített interakciókban 350 kérdés vizsgálatát végezték el, mindössze a kérdő megnyilatkozások 6,29%-a volt „nem funkcionális" (Stivers 2010). Ugyanebben a projektumban japán nyelvű diskurzusok (ugyancsak informális, spontán társalgások 3-4 barát vagy családtag részvételével) 350 kérdése közül szintén mindössze

8,57% volt a választ nem várók aránya (Hayashi 2010). Ezek az adatok azt sugallják, hogy a nem információkérő kérdések arányára hatással lehet az interakció stílusa (az a tény, hogy a formális-informális skála mentén hol helyezkedik el), illetőleg olyan tényezők is, mint például a társalgás témája. Az informálisabb helyzetben rögzített társalgásokban, ahol a résztvevők közötti szociális távolság (Brown–Levinson 1987) kicsi, arányaiban jóval több valódi kérdés hangzik el, mint azokban a helyzetekben, amelyekben a szociális távolság nagy.

Megvizsgáltuk azt is, hogy a beszélők nemének és életkorának függvényében hogyan alakul az információkérő kérdések aránya (3. táblázat). A fiatal beszélők esetében ez az arány 70% feletti, és a beszélő életkorának előrehaladtával fokozatosan csökken mindkét nem esetében. A nemek szerint nincs számottevő eltérés.

3. táblázat: Az információkérő kérdések aránya a beszélők életkorának és nemének függvényében

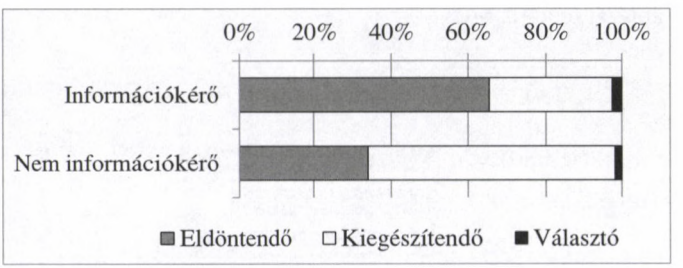
	Nők	**Férfiak**
Fiatalok	75,00%	70,45%
Középkorúak	53,85%	60,00%
Idősebbek	50,00%	41,94%

Az arányok alakulását feltehetően befolyásolja az a tény, hogy a társalgás állandó résztvevői (az adatközlő két beszédpartnere) fiatalok, így a fiatalabb adatközlőkkel könnyebben alakul informálissá a helyzet még viszonylag nagyobb szociális távolság esetén is. Az informálisabb szituáció pedig – mint láttuk – nagyobb arányban hív elő információkérő kérdéseket.

A jelen vizsgálati anyagban a nem információkérő kérdések mintegy fele (27 előfordulás, az összes nem valódi kérdés 48,21%-a) idézet volt, azaz egy felidézett (múltbeli vagy fiktív) beszédeseményből származott, pl. *Mér, maga soha nem járt itt az oktatójával?* (egy KRESZ-vizsgáról szóló történetből). 19 kérdés (a nem valódi kérdések 33,93%-a) retorikai funkcióval bírt (pl. *Mér az iskolára kell passzolni, hogy megnevelje a gyereket?*), 7 további (12,50%) pedig fatikus közbevetésként realizálódott (pl. *...tudod?*). 3 kérdéssel (5,36%) a beszélő a bizonytalanságát fejezte ki: bár tudatában volt annak, hogy a partnere nem tudhatja a választ, így jelezte, hogy nem egészen biztos a közlése tartalmában. Példaként idézethetjük a következő szomszédsági párt, amelyben B közlése az eldöntendő kérdő mondatra jellemző intonációval hangzott el:

A: *Hány óra volt az út?*
B: *Hát olyan tizenkettő?*

A 134-ből tehát 78 esetben követte a kérdést verbális vagy nonverbális válasz, illetve megerősítés. Ezen információkérő kérdések logikai szemantikai osztályozását külön is elvégeztük, és a teljes korpusz arányaitól, valamint a nem információkérő kérdések arányaitól is merőben eltérő eredményeket kaptunk (2. ábra).



2. ábra
Az egyes logikai szemantikai kategóriába tartozó kérdések aránya az információkérő és a nem információkérő kérdő megnyilatkozások között

Az információkérő kérdések között az eldöntendők aránya kimagasló (51 előfordulás, 65,38%) a kiegészítendőkéhez (25 példány, 32,05%) képest, és várhatóan igen kevés (2 darab, 2,56%) választó kérdés tartozott ebbe a csoportba is. Az 56 nem információkérő kérdés körében a két nagy kategória aránya épp fordított: a kiegészítendők adják a csoport kétharmadát (36 előfordulás, 64,29%), egyharmadot tesznek ki az eldöntendők (19 példány, 33,93%), és csak egy választó kérdés tartozott ide (1,79%).

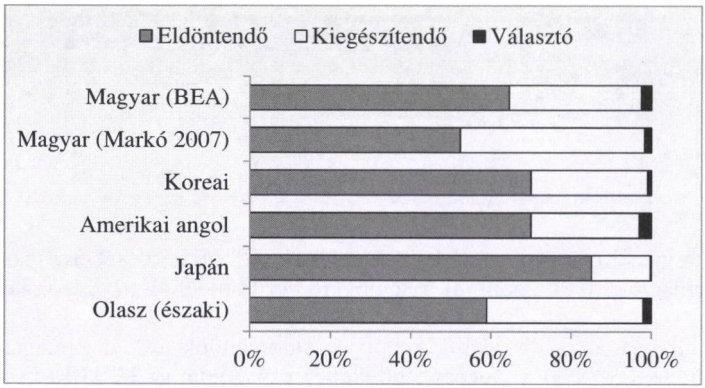
Az információkérő kérdések logikai szemantikai eloszlását ugyancsak összevetettük saját korábbi (magyar nyelvű) korpuszunk elemzési eredményeivel (Markó 2007), valamint a *Journal of Pragmatics* (2010) különszámában bemutatott, más nyelvre kapott eredményekkel is (3. ábra). (A többi korpuszban a nem információkérő kérdések kis száma miatt arra a csoportra vonatkozóan nem végeztük el az összehasonlítást.) Az információkérő kérdésekre vonatkozó korpuszadatok azt mutatják, hogy a választó kérdések aránya nyelvtől függetlenül elhanyagolható, illetve az eldöntendő kérdések aránya többé-kevésbé meghaladja a kiegészítendőkét. A kiegészítendő és az eldöntendő kérdések számbeli különbsége a 2007-es magyar vizsgálatban volt a legkisebb (mintegy 6 százalékpontnyi eltérés: 52,27 vs. 46,02%), és a japánban a legnagyobb (ahol majdnem 6-szor több eldöntendő kérdést adatoltak a vizsgált anyagban, mint kiegészítendőt).

A korpuszbeli kérdések prozódiai sajátosságai

A prozódiai elemzést 125 kérdő megnyilatkozáson végeztük el, mivel 5 kiegészítendő kérdés glottalizáltan (4), illetve leheletes zöngével (1) realizáló-

dott, valamint 4 további kérdést (3 eldöntendőt és 1 kiegészítendőt) emocionális töltetük (csalódottságot fejeztek ki) miatt zártunk ki a vizsgálatból.

Mindenekelőtt elemeztük a fonológiai leírásnak megfelelő, prototipikus megvalósulások gyakoriságát mind a teljes anyagra, mind külön az információkérő és a nem információkérő kérdésekre, valamint az egyes logikai szemantikai típusokra vonatkozóan. A csoportonkénti darabszámokat és arányokat a 4. táblázat mutatja be.



3. ábra

A szemantikai-logikai típusok megoszlása a valódi kérdések között magyar (a jelen kutatás és Markó 2007), koreai (Yoon 2010), amerikai angol (Stivers 2010), japán (Hayashi 2010) és északi olasz (Rossano 2010) társalgásokban

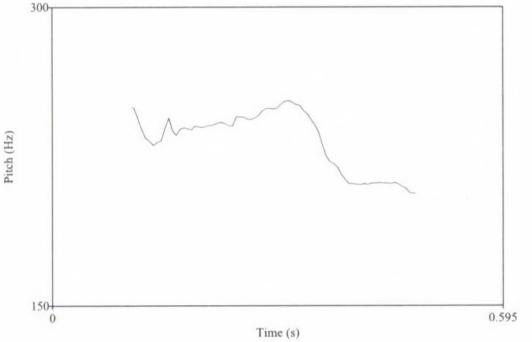
Az adatok alapján nincs számottevő különbség az információkérő és a nem információkérő kérdések között az intonáció prototipikus voltát tekintve, ugyanakkor az információkérő kérdések valamivel nagyobb arányban realizálódtak prototipikus kontúrral, mint a választ nem várók. A nem információkérő kérdések körében az elvárttól eltérő változatok az eldöntendő kérdések körében gyakoribbak.

A táblázatbeli számadatok általánosságban jól jellemzik a vizsgált korpuszt, ugyanakkor elfednek néhány, kisebb elemszámú adatra érvényes, ám funkcionális szempontból annál érdekesebb eredményt. Ahogy korábban már említettük, 3 olyan kérdő megnyilatkozás volt az anyagban, amely a beszélő bizonytalanságát fejezte ki. Formai szempontból ezek közül egy volt eldöntendő kérdés, amely (prototipikus) szökő-eső zárlattal realizálódott, a másik kettő azonban szintaktikailag kiegészítendő kérdés volt, amelyek (az ekhókérdéshez hasonlóan) ugyancsak az eldöntendő kérdés hanglejtésével valósultak meg (pl. *Mennyi ez?*, vö. 4. ábra). Emellett 3 retorikai (szintaktikailag kiegészítendő) kérdést adatoltunk szintén szökő-eső dallamzárral (pl. *Miről beszélünk?*). Ezek az eredmények arra utalnak, hogy a beszélő specifikus szándéka

kifejezésének céljából a kiegészítendő kérdések nemcsak abban az esetben „kölcsönözhetik" az eldöntendő kérdés intonációját, ha a kérdés funkcióját tekintve ekhó- vagy ismételtető kérdés.

4. táblázat: A prototipikus prozódiai megvalósulások aránya a különböző funkcionális és formális kérdéstípusokban

		Kiegészítendő	Eldöntendő	Választó	Összesen
Információkérő kérdések	A realizációk száma	24	49	2	75
	A prototipikus realizációk száma és aránya	19 (79,17%)	39 (79,59%)	1 (50,00%)	59 (78,67%)
Nem információkérő kérdések	A realizációk száma	31	18	1	50
	A prototipikus realizációk száma és aránya	23 (74,19%)	12 (66,67%)	1 (100,00%)	36 (72,00%)
Minden kérdés	A realizációk száma	55	67	3	125
	A prototipikus realizációk száma és aránya	42 (76,36%)	51 (76,12%)	2 (66,67%)	96 (76,80%)

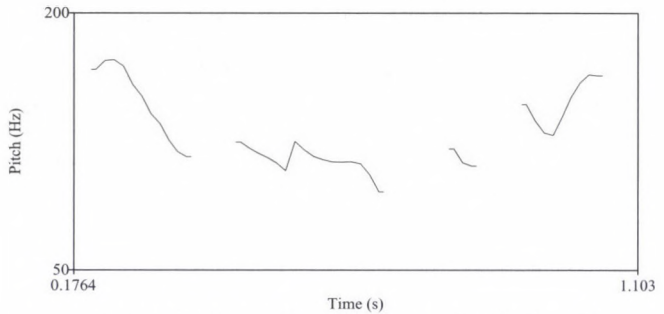


m e nny i e z

4. ábra
Bizonytalanságot kifejező kiegészítendő kérdés szökő-eső intonációval (középkorú női beszélő)

Habár az emelkedő dallammal záruló kiegészítendő kérdéseket gyakorinak tartják, mindössze 8 ilyen jelent meg a korpuszban. Annak ellenére, hogy a

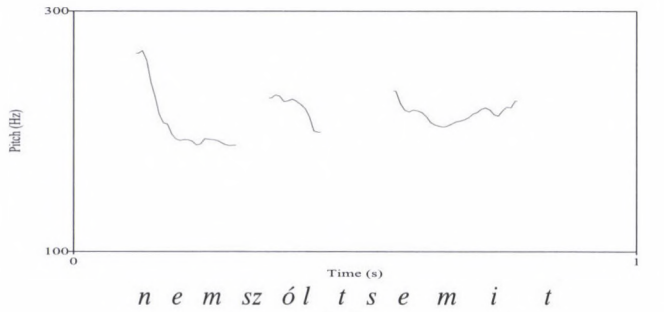
közvélekedés (példákkal szolgál Markó 2012) és néhány szakirodalmi forrás (pl. Fónagy 1998) szerint ez a forma a nők beszédére jellemző, a jelen korpuszban egyetlen előfordulást (fiatal nő) leszámítva férfiak (mindhárom korosztályból) hangfelvételein dokumentáltuk (5. ábra).



5. ábra

A *Melyik a Borszéki út?* kiegészítendő kérdés emelkedő dallamzárral (középkorú férfi beszélő, nem információkérő – idézett – kérdés)

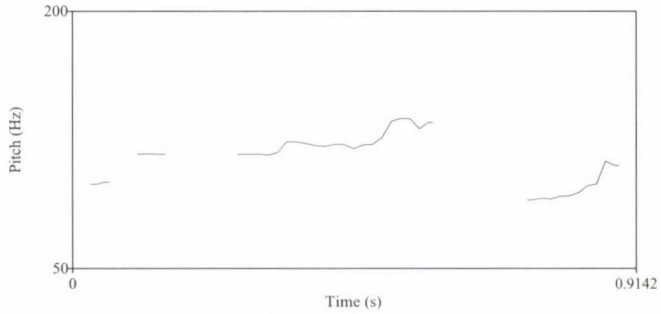
Az eldöntendő kérdések között is voltak olyanok – mintegy az esetek negyedében –, amelyek a prototipikus kérdő intonációtól eltérő dallammal valósultak meg. Ezek nagyobbrészt ugyanazokat a változatokat reprezentálják, mint amelyeket a korábbi korpuszalapú vizsgálatban (Markó 2007) találtunk. A prototipikustól eltérő intonációjú 16 eset mintegy harmadában (2 fiatal férfi és 3 fiatal nő hangfelvételén) a szökő-eső dallamívet újabb F_0-emelkedés követte (6. ábra), vagy nem is valósult meg a dallamesés, az alaphang magasan lebegő maradt.



6. ábra

A *Nem szólt semmit?* eldöntendő kérdés a kérdő mag után újból emelkedő dallamzárral (fiatal női beszélő, információkérő kérdés)

Ugyanilyen arányban fordult elő az, hogy az alaphangmagasság hirtelen megemelkedése helyett magasan lebegő volt a dallam az utolsó előtti szótagig, és ezt követte az esés az utolsó szótagon (és ezt a formát megítélésünk szerint nem motiválta valamely sajátos beszélői attitűd). Más esetekben (2 fiatal, 1 középkorú és 2 idősebb férfi felvételén) az F_0 nem az utolsóelőtti, hanem az azt megelőző szótagon érte el a csúcsát, és alkalmanként ezt is követte újabb dallamemelés a közlés végén (7. ábra).



7. ábra

Az *Ez is már rögzítve van?* eldöntendő kérdés két F_0-csúccsal (fiatal férfi beszélő, információkérő kérdés)

Következtetések

A kérdések spontán beszédbeli vizsgálata jelentős módszertani nehézségekbe ütközik, hiszen kérdő mondatformákat csak társalgási helyzetekben lehet vizsgálni – mégpedig olyanokban, amelyekben a beszélők egyenrangú partnerek, és ez a szimmetria az adatközlőt is „feljogosítja" kérdések megfogalmazására. Hiszen a korábban rögzített, akusztikai fonetikai eszközökkel is elemezhető, jó minőségű spontánbeszéd-felvételek túlnyomó része interjúhelyzetben készült, amelyben a beszélőtől többnyire kvázimonologikus narratívát vártak el a kutatók. Ezekben legfeljebb a feltett kérdésre vagy az elhangzott instrukcióra vonatkozó visszakérdezések (ekhókérdés) jelentek meg. Egy-egy kiegyenlített társalgási helyzetben készült hangfelvételt leszámítva (pl. Markó 2007) tehát nincsenek adatok a kérdések spontán beszédbeli megvalósulásáról. Ezen a területen a társalgást is tartalmazó magyar nyelvű beszédadatbázisok létrejötte jelenthet áttörést. Ugyanakkor természetesen ezeknek az esetében is felmerül a kérdés, hogy egy adott szituációban az egyes résztvevők között kialakul-e olyan kontaktus, amely előhív kérdő megnyilatkozásokat az adatközlő részéről is. Ezt befolyásolhatja a beszédpartnerek életkori közelsége, neme, a beszédtéma, az ahhoz való viszonyulásuk, érzelmi állapotuk. Mindemellett a megalapozott fonetikai vizsgálatokhoz zaj-

mentes felvételek szükségesek, a kifejezés információelméleti értelmében. Azaz csak azok a kérdő megnyilatkozások alkalmasak akusztikai elemzésre, amelyeknek az elhangzása közben a többi beszélő nem szólal meg (nincs egyszerre beszélés, hümmögés stb.), valamint amelyeket a beszélő nonverbális kommunikációja sem torzít (pl. nevetéssel, megváltozott zöngeminőséggel), sem külső zajok (pl. a gesztikulációból adódó kopogás) nem rontják a hangminőséget.

A spontán beszéd vizsgálata – a módszertani nehézségek ellenére – a kérdések elemzésében is gyümölcsözőnek mondható, hiszen ha kis számban is, olyan változatokat adatolhattunk (a 2007-es kutatásban és a BEA korpuszán is), amelyek pusztán introspektív módszerrel vagy hallási tapasztalat alapján tudomásunk szerint nem kerültek fel a kérdések korábbi intonációs térképére.

Tanulmányunkban a BEA adatbázis 30 társalgásából elemeztük az adatközlők kérdő megnyilatkozásait funkcionális és formális szempontok alapján, és ahol lehetett, összevetettük az eredményeket más korpuszok adataival. Külön is vizsgáltuk az információkérő és az egyéb kérdéseket, elemeztük a kérdéstípusok lehetséges összefüggéseit a beszélő nemével és életkorával, valamint a társalgási helyezet egyéb sajátosságaival.

Más nyelvekre kapott adatokkal is összevetve eredményeinket azt találtuk, hogy a kérdéstípusok gyakorisága sem funkcionális, sem logikai szemantikai tekintetben nem nyelvfüggő, ugyanakkor feltételezhető kapcsolat van az interakció formalitásának foka és a kérdések típusgyakorisága között (vö. Kearsley 1976).

Azt feltételeztük, hogy van összefüggés a kérdések logikai szemantikai típusa, pragmatikai funkciója és prozódiai megvalósulása között. Ez a hipotézis részben igazolódott: az információkérő és a nem információkérő kérdések között éppen fordított eloszlásban találunk eldöntendő és kiegészítendő kérdéseket. Ugyanakkor csak az eldöntendő kérdések esetében bizonyult igaznak az a feltételezés, hogy az információkérő kérdések nagyobb arányban realizálódnak prototipikus intonációs kontúrral. Ez vélhetőleg azzal függ össze, hogy az információkérő eldöntendő kérdések esetében a jól azonosítható intonációs kontúr (más jelölő hiányában) fontos észlelési kulcs.

Elemzéseink alapján a magyar spontán beszédbeli kérdő megnyilatkozások mintegy 20-30%-ban eltérnek a szakirodalom által leírt prototipikus intonációs formáktól. Néhány alternatív realizációs típus feltehetően pragmatikai tényezőkre vezethető vissza, amelyeket a korábbi (főleg felolvasott és eljátszott mintákat feldolgozó) kutatásokban kevéssé vettek tekintetbe. A megszokottól eltérő realizációkat jelentősen nagyobb számban adatoltuk férfi adatközlők beszédében, ami nyilvánvalóan összefügg azzal a ténnyel, hogy a vizsgált kérdések mintegy kétharmada férfiak társalgásfelvételérő származott, az arányok eltolódását önmagában ez a tény azonban mégsem magyarázza. Más, felolvasott és spontán beszédanyagon végzett elemzések (vö. Markó 2012) adatai szerint például a kiegészítendő kérdés emelkedő dallamzárlata egyálta-

lán nem jellemzőbb a nők beszédére. Feltehető, hogy a nemek között feltételezett különbség valójában „érzékcsalódás", hallási észlelésünk frekvenciafüggő voltából adódik. Az életkori csoportok között sem találtunk eltérést abban, hogy milyen gyakran jelennek meg a prototipikustól eltérő dallamkontúrral megvalósított kérdések. A második hipotézisünk tehát igazoltnak tekinthető.

Az eredmények felvetik azt a kérdést, hogy a prototipikustól eltérő prozódiai realizációk típusai mennyiben tekinthetők „szabálytalannak", és mennyiben a fonológia által meghatározott alapváltozat variánsainak. Feltehető, hogy ahogyan a kiegészítendő kérdő szókőzáras megvalósulása a kanonizált formák közé került, a spontán beszédbeli kérdő megnyilatkozások további elemzése más, hasonlóan gyakori formák funkcióiról is számot adhat, és ezáltal árnyalja a kérdések formai és funkcionális sajátosságairól és csoportosításáról alkotott képünket.

Irodalom

Austin, John L. 1962/1990. *Tetten ért szavak.* Akadémiai Kiadó, Budapest.

Bartók János 1978. A négyféle kérdő hanglejtés. *Magyar Fonetikai Füzetek 1.* 97–103.

Boersma, Paul – Weenink, David 2012. *Praat: doing phonetics by computer (Version 5.3).* http://www.fon.hum.uva.nl/praat/download_win.html

Brown, Penelope – Levinson, Stephen 1987. *Politeness. Some universals in language use.* Cambridge University Press, Cambridge.

Csűry Bálint 1925. *A szamosháti nyelvjárás hanglejtésformái.* A Magyar Nyelvtudományi Társaság Közleményei 22. Magyar Nyelvtudományi Társaság, Budapest.

Deme László 1962. Hangsúly, szórend, hanglejtés, szünet. In Tompa József (szerk.): *A mai magyar nyelv rendszere. Leíró nyelvtan II.* Akadémiai Kiadó, Budapest, 457–522.

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] 1980. Hanglejtés. In Grétsy László – Kovalovszky Miklós (főszerk.): *Nyelvművelő kézikönyv. Első kötet. A–K.* Akadémiai Kiadó, Budapest, 755–784.

Elekfi László – Wacha Imre é. n. [2003]. Az értelmes beszéd hangzása. Mondatfonetika – kitekintéssel a szövegfonetikára. Szemimpex Kiadó, Budapest.

Fishman, Pamela 1978. Interaction: The work women do. *Social Problems* 25. 397–406.

Fónagy Iván 1998. Intonation in Hungarian. In Hirst, Daniel – di Cristo, Albert (eds.): *Intonation systems. A survey of twenty languages.* Cambridge University Press, Cambridge, 328–344.

Fónagy Iván – Magdics Klára 1963. A kérdő mondatok dallamáról. *Nyelvtudományi Értekezések* 40. 89–106.

Fónagy Iván – Magdics Klára 1967. *A magyar beszéd dallama.* Akadémiai Kiadó, Budapest.

Freed, Alice F. 1994. The form and function of questions in informal dyadic conversation. *Journal of Pragmatics* 21. 621–644.

Gósy Mária 1993. A kiegészítendő kérdés dallamváltozása. *Magyar Nyelvőr* 117. 443–447.

Gósy Mária 2000. A kérdő, a felkiáltó és a felszólító mondat dallamának azonossága és különbözősége. In T. Molnár István – Klaudy Kinga (szerk.): *Papp Ferenc akadémikus 70. születésnapjára.* Debreceni Egyetem Kossuth Egyetemi Kiadó, Debrecen, 123–129.

Gósy, Mária – Terken, Jacques 1994. Question marking in Hungarian: timing and height of pitch peaks. *Journal of Phonetics* 22. 269–281.

Gósy Mária – Gyarmathy Dorottya – Horváth Viktória – Gráczi Tekla Etelka – Beke András – Neuberger Tilda – Nikléczy Péter 2012. BEA: beszélt nyelvi adatbázis. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások.* Akadémiai Kiadó, Budapest, 9–24.

Grice, Paul H. 1975/1991. A társalgás logikája. In Pléh Csaba – Síklaki István – Terestyéni Tamás (szerk.): *Nyelv – kommunikáció – cselekvés.* Osiris, Budapest, 213–28.

Gyuris, Beáta 2009. Sentence-types, discourse particles and intonation in Hungarian. In Solstad, Torgrim – Riester, Arndt (eds.): *SinSpeC Volume 5. Proceedings of Sinn und Bedeutung 13.* Stuttgart University, Stuttgart. 157–170. http://elib.uni-stuttgart.de/opus/volltexte/2009/4138/pdf/sub13_band1_190609.pdf

Hayashi, Makoto 2010. An overview of the question–response system in Japanese. *Journal of Pragmatics* 42. 2685–2702.

Hegedűs Lajos 1930. Magyar hanglejtésformák grafikus ábrázolása. Kísérletfonetikai tanulmány. A bécsi Collegium Hungaricum füzetei V. Budapest.

Hymes, Dell 1974. *Foundations in Sociolinguistics.* University of Pennsylvania Press, Philadelphia, PA.

Kearsley, Greg 1976. Questions and question-asking in verbal discourse: A cross-disciplinary review. *Journal of Psycholinguistic Research* 5/4: 355–375.

Keszler Borbála 2009. Kérdés- és felelettípusok a spontán beszédben. In Keszler Borbála – Tátrai Szilárd (szerk.): *Diskurzus a grammatikában – grammatika a diskurzusban.* Segédkönyvek a nyelvészet tanulmányozásához 88. Tinta Könyvkiadó, Budapest. 13–20.

Kiefer Ferenc 2000. *Jelentéselmélet.* Corvina Kiadó, Budapest.

Kugler Nóra 2000. A mondattan általános kérdései. In Keszler Borbála (szerk.): *Magyar grammatika.* Nemzeti Tankönyvkiadó, Budapest, 369–393.

Markó Alexandra 2007. Kérdő funkciójú hanglejtésformák a spontán beszédben. *Beszédkutatás 2007.* 59–74.

Markó Alexandra 2012. Vélemények és tények egy magyar hanglejtésformáról. *Anyanyelv-pedagógia* 2012/3. http://www.anyanyelv-pedagogia.hu/cikkek.php?id=400

Mycock, Louise 2007. The role of prosody in constituent question formation: a comparison of Hungarian and Japanese. *The Phonetician* 95. 7–18.

Mycock, Louise 2010. Multiple-clause constituent questions: Intonation and variation in Hungarian. *Acta Linguistica Hungarica* 57/2–3. 268–287.

Olaszy Gábor 1995. A kérés, a figyelmeztetés és a kérdés prozódiája a kijelentő mondatok tükrében. *Beszédkutatás '95.* 46–61.

Olaszy Gábor 2002. A magyar kérdés dallamformáinak és intenzitásszerkezetének fonetikai vizsgálata. *Beszédkutatás 2002.* 83–99.

Rossano, Federico 2010. Questioning and responding in Italian. *Journal of Pragmatics* 42. 2756–2771.
Schegloff, Emmanuel – Sacks, Harvey 1973. Opening up closings. *Semiotica* 7/4. 289–327.
Schirm Anita 2007. A kérdések pragmatikája. In Váradi Tamás (szerk.): *Alknyelvdok 2007. I. Alkalmazott Nyelvészeti Doktorandusz Konferencia*. MTA Nyelvtudományi Intézet, Budapest. 161–171. http://www.nytud.hu/alknyelvdok07/proceedings07/Schirm.pdf
Searle, John 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press, Cambridge.
Stivers, Tanya 2010. An overview of the question–response system in American English conversation. *Journal of Pragmatics* 42. 2772–2781.
Tolnai Vilmos 1915. Adatok a magyar hanglejtéshez. *Magyar Nyelv* 11. 51–59, 108–116, 152–156.
Varga László 1981. A magyar intonáció – funkcionális szempontból. *Nyelvtudományi Közlemények* 83. 313–339.
Varga László 1983. Hungarian sentence prosody: An outline. *Folia Linguistica* 17. 117–151.
Varga László 1993. *A magyar beszéddallamok fonológiai, szemantikai és szintaktikai vonatkozásai*. Nyelvtudományi Értekezések 135. Akadémiai Kiadó, Budapest.
Varga László 1994. A hanglejtés. In Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 2. Fonológia*. Akadémiai Kiadó, Budapest, 468–549.
Varga László 2002a. *Intonation and Stress. Evidence from Hungarian*. Palgrave Macmillan, Houndmills, Basingstoke
Varga László 2002b. Az egy szótagú magyar eldöntendő kérdések intonációja. In Hunyadi László (szerk.): *Kísérleti fonetika – laboratóriumi fonológia 2002*. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen, 127–140.
Vion, Monique and Annie Colas 2006. Pitch cues for the recognition of Yes-No questions in French. *Journal of Psycholinguistic Research* 35: 427–445.
Yoon, Kyung-Eun 2010. Questions and responses in Korean conversation. *Journal of Pragmatics* 42. 2782–2798.

A tanulmány a Bolyai János kutatási ösztöndíj támogatásával készült.

A BESZÉDSZÜNETEK FONETIKAI SAJÁTOSSÁGAI A BESZÉDTÍPUS FÜGGVÉNYÉBEN

Bóna Judit

Bevezetés

A beszéd folyamatosságát gyakran szünetek és megakadásjelenségek szakítják meg. A beszédszünet egyrészt fiziológiai szükséglet (a levegővételhez biztosít időt), másrészt lehet akaratlagos (a beszélő tagolni akarja a mondanivalóját, illetve hatásszünetet tart), illetve jelezheti a beszélő beszédtervezési nehézségeit is (Gósy 2000). A beszédszünet meghatározása nem teljesen egyértelmű a szakirodalomban (vö. Hieke et al. 1983; Gósy 2000), és funkciója, akusztikai-fonetikai és percepciós paraméterei is vitás kérdéseket vetnek fel. A szünetmeghatározások korábban ugyanis vagy a formát, vagy a funkciót tekintették kiindulásnak, és ez számos ellentmondáshoz és osztályozási nehézséghez vezetett. Sallai és Szende (1995) a szünetet a nyelvi szerkezetekkel összefüggésben határozza meg, és a szekvencia (szeriális szerkezet) egyfajta információt képző, hordozó megszakításaként definiálja.

Számos kérdést vet fel azonban, hogy a szekvenciákat megszakító elemek közül egyáltalán mit tekintünk szünetnek, hiszen vannak a beszédben például olyan jelkimaradások, amelyek egy-egy beszédhang részei. Ilyen jelkimaradások, de nem beszédszünetek a zöngétlen explozívák és a zöngétlen affrikáták zárszakaszai (Gósy 2004). Egyes kutatók szerint a junktúra is tulajdonképpen 0 ms-os szünet, és egyfajta szünetnek tekintendők az olyan „szünetkompenzációs" esetek is, amikor a néma szünet előtt megnyúlik a magánhangzó (Sallai–Szende 1995). Mások – az angol nyelvre vonatkozóan – a szünetek közé sorolják a különböző megakadásjelenségeket is, mint például az ismétlést, téves kezdéseket stb. (Mahl 1956; idézi Gósy 2004). Ezeket a mai magyar szakirodalomban a szünetektől elkülönítve – bizonytalanságból fakadó megakadásokként vagy hibákként – szokás elemezni (Gósy 2005).

A szünet nem feltétlenül néma jelkimaradás, a jellel kitöltött szüneteket hezitálásnak hívjuk. Ezek leggyakoribb formája a magyarban az *öö* és a *mm* (Gósy 2000), de megjelenhetnek például *öm, öh, ömh, höh, eh, eö* stb. formában is (Horváth 2010; Gósy et al. 2013). A néma szünet és a hezitálás kombinálódhat is (Gósy 2003; Markó 2005a; Horváth 2010), vannak olyan szünetek, amelyekben néma szünet + hezitálás, hezitálás + néma szünet vagy néma szünet + hezitálás + néma szünet váltja egymást (ezeket kombinált szünetnek hívjuk). A szünettartás egyénre jellemző sajátosság, ha egy beszélőre hosszabb néma szünetek a jellemzők, akkor valószínűleg hosszabb kitöltött szü-

neteket tart (Gósy–Beke–Horváth 2011). A jelen tanulmányban a *beszédszünet/szünet* kifejezés alatt értjük mind a néma, mind a kitöltött és kombinált szüneteket. Amikor nem általánosságban a beszédszünetet vizsgáljuk, ott konkrétan megnevezzük az elemzett szünettípust.

A beszédszünet minimum tartamának megállapítása a kezdetektől vita tárgyát képezte a szakirodalomban. Eleinte 250 ms-ban határozták meg a szünet minimum időtartamát, mivel az akkori elemző berendezésekkel nem lehetett könnyen megkülönböztetni a néma szünetet a beszédhanghoz tartozó jelkimaradásoktól, illetve az elemzés nem adott lehetőséget az igen rövid tartamok biztos szegmentálására. Azóta a kezdetben 250 ms-os alsó határt lecsökkentették egészen a 30 ms-os értékig, de a különböző kutatók mindmáig eltérő értékeket adnak meg a szünet legrövidebb tartamaként (vö. Gósy 2000).

A jelkimaradás nem mindig tekinthető funkcionális szempontból beszédszünetnek: a „csend" (silence) megjelenhet 1. egy folyamatban levő egységen belül, 2. társalgási egységek között, illetve 3. jelezheti a társalgás végét (Levelt 1989). Markó (2005a) a szünet mellé a hallgatás terminusát vezeti be, amelyet úgy definiál, mint a beszédjelnek két társalgási egység közötti kimaradása. Mintegy 5 másodpercben határozza meg azt az időtartamot, amely még a társalgás résztvevői számára elfogadható; ha ennél hosszabbra nyúlna a hallgatás, valakinek meg kell szólalnia. A hallgatás elfogadható maximum időtartama a beszédpartnerek kapcsolatának függvényében, illetőleg a társalgás különböző szakaszaiban változhat (Bata 2009).

Fónagy (1967) a legfontosabb szünetfunkciónak a beszéd tagolását tartja. A tagolásban látja a szünetek fő szerepét Szende (1976) is, aki a beszédszünet megjelenési formájának csak két típusát írja le: a lélegzetvételt és a hezitációs szünetet. A szakirodalomban számos más szünetfunkcióról is olvashatunk (Zellner 1994; Rose 1998), például gondolkodási szünetről, a beszédtervezési bizonytalanságokat „feloldó" szünetről és hatásszünetről, de a szünet kaphat például nyomatékosító funkciót is (Strangert 2003). A hezitálásoknak, azaz a kitöltött szüneteknek Horváth (2010) szerint három fő funkciójuk lehet: a beszédszándék kifejezése, a bizonytalanság jelzése, illetve a hibajavításhoz biztosítanak időt. A társalgásban további funkciói lehetnek: például a beszélő kitöltött szünettel jelezheti a beszédszakasz végén a beszédpartnernek, hogy át kívánja adni a szót, vagy épp ellenkezőleg, magánál akarja tartani (Horváth 2010).

A beszédszünetnek a percepcióban betöltött szerepéről már a '60-as évekből találunk utalást (Goldman-Eisler 1968), Hegedűs Lajos pedig az 1957-ben végzett beszédtempó elemzések kapcsán jelenti ki, hogy a szünetek fontos szereppel bírnak a tempóélmény kialakításában. A szünetek percepcióját azóta sokféle szempontból vizsgálták magyar nyelven is (Kassai 1988; Kassai–Fagyal 1996; Menyhárt 1998; Gósy 2000; Bóna 2006). Gósy (2000) eredményei szerint az emberek a beszédszünetek mintegy kétharmadát észlelik; és a szünet időtartama és észlelése szoros összefüggést mutat. Saját kuta-

tásunkban (Bóna 2006) azt találtuk, hogy a hallgatók a szünetek mintegy 54%-át észlelték helyesen, az összes észlelt szünet 10%-át hibásan. Nagy különbség volt a tesztben szereplő adatközlők között: a legjobban teljesítő személy a szünetek 96%-át, a legrosszabbul teljesítő 31%-ot észlelt helyesen. A kutatások eredményeiből az is látszik, hogy a szünetészlelést meghatározzák az egyéni sajátosságok, a szöveg típusa, tartalma, illetve a szünetidőtartamok is. A kísérleti személyek általában a hezitálásokat és az 500 ms-nál hosszabb néma szüneteket észlelték jól. A szünet szükséges a megfelelő beszédmegértéshez is. Azokban a vizsgálatokban, amelyekben olyan szövegek megértését vizsgálták, amelyekből előzőleg kivágták a szüneteket, a kísérleti személyek szövegértése nagymértékben csökkent a szünetes szövegnél kapott teljesítményhez képest (Gósy 2000; Bóna 2009).

A beszédprodukciós és a beszédpercepciós szempontokat is figyelembe veszi Gósy Mária (2000: 2) definíciója, amely szerint „a beszédfolyamatban jelentkező szünet olyan kismértékben akaratlagos beszédkimaradás, amely néma vagy jellel kitöltött, de független a beszédhang képzésétől. Funkcióját tekintve a beszédprodukcióban 1. biztosítja az artikulációt lehetővé tevő légáramot, 2. elősegíti a közlés értelmi tagolását, 3. a beszédtervezés során az ún. ellentmondások, téves utak stb. feloldására szolgál, 4. a mentális lexikonban történő keresési idő kitöltését biztosítja, illetőleg lehetőséget nyújt a nyelvi kódolás módosítására. Funkciói a beszédmegértésben: 1. az elhangzottak könnyebb feldolgozása, 2. az entrópia csökkentése és 3. a megértés és az értelmezés működési folyamatainak biztosítása". Kutatásunkban ezt a szünetdefiníciót fogadtuk el, elemzéseinket ennek alapján végeztük.

A beszédszünetek aránya a teljes beszédidőhöz viszonyítva egyén- és szituációfüggő, általában 20–30% (Gósy 2005). A beszédszünetek gyakorisága és időtartama számos tényezőtől függ: a mondat hosszától és összetettségétől (Volskaya 2003; Krivokapic 2007); a beszélő személy adottságaitól, a beszédben való jártasságától; a témától; a beszédhelyzettől (Markó 2005b); a műfajtól (Imre 2005; Olaszy 2005), a beszédstílustól (Duez 1982) stb.

A beszédszüneteket további grammatikai és szociolingvisztikai változók mentén is vizsgálták: elemezték például a beszédszünetek helyét (Sallai–Szende 1995; Gósy 2004; Vallent 2005; Menyhárt 2010); a 60 évvel ezelőtti és a mai beszélők szünethasználatát (Vallent 2005; Menyhárt 2010; 2012); a falusi és a városi beszélők szünetezését (Menyhárt 2010); illetve a különböző életkorú beszélők beszédszüneteit (Vallent 2010; Laczkó 2009; Bóna 2010; 2012a). Ezekben a vizsgálatokban a néma, a kitöltött és a kombinált szüneteket is tekintetbe vették. A beszéd szünetezését nagymértékben befolyásolja a beszédtípus; azaz nem mindegy, hogy spontán beszédet vagy felolvasást elemzünk; az adatközlő a saját életéről mond egy monológot, egy képsor alapján alkot egy történetet, egy hallott szöveget próbál visszaidézni, vagy éppen egy többszereplős társalgásban vesz részt. A különböző beszédhelyze-

tek ugyanis eltérő beszédtervezési stratégiákat igényelnek, amelyek elsősorban a beszéd temporális sajátosságaiban jelennek meg.

A spontán beszéd produkciója során a beszédtervezés és -kivitelezés egyszerre zajlik, a beszélő maga tervezi meg az üzenet tartalmát, szabadon válogathat a szavakból és a nyelvtani szerkezetekből. Mivel a cél ilyenkor gyakran az, hogy a beszélő önállóan, megszakítás nélkül minél hosszabb beszédproduktumot hozzon létre, van elég ideje a mondanivalója megfogalmazására (Gósy 2005). A felolvasás során ezzel szemben nincs szükség a magasabb tervezési folyamatok működtetésére, hiszen a felolvasandó nyelvi anyag adott. Ez utóbbi beszédstílusban tehát nem játszik szerepet a makrotervezés (az üzenet megfogalmazása) és a mikrotervezés néhány folyamata (a nyelvi átalakítás, vö. Levelt 1989). Az irányított spontán beszéd létrehozásakor, például a tartalom-visszamondáskor egy hallott szöveg tartalmi elemeit, szerkezetét, szókincsét kell felidéznie az adatközlőnek (Gósy 2010). A tartalomvisszamondást számos tényező befolyásolja, ezek közül a legfontosabbak a beszédfeldolgozás sikeressége, a figyelmi folyamatok, a rövid és a hosszú távú memória működése, illetve a narratív kompetencia (vö. Juncos-Rabadán–Pereiro 2000). Ugyanakkor a társalgás is különböző beszédstratégiákat kíván meg, mint egy monologikus beszédtípus, hiszen ebben az esetben egyrészt egyfajta „verseny" alakul ki a beszélők között, másrészt a többiek megszólalásai időt adnak a beszélőnek a beszédtervezésre (Markó 2005a). Wacha Imre (1974) az elhangzó beszédet négy fő stíluskategóriába sorolta (1. spontán, kötetlen beszéd; 2. felolvasás; 3. reproduktív vagy interpretatív beszéd; 4. félreproduktív beszéd), az akusztikai szerkezet szerint a két legszélsőségesebb kategória a spontán beszéd és a felolvasás. Ezeknek mind sajátos hangzásuk van, és mindegyiket eltérő szünettartás jellemzi.

Különböző beszédhelyzetben rögzített beszédtípusok szupraszegmentális fonetikai összevetését, így a szünettartás sajátosságainak elemzését a magyar szakirodalomban elsőként Markó Alexandra (2005a) végezte el. Kutatásában négy adatközlő többféle monologikus szövegét vetette össze ugyanazon beszélők négyszereplős társalgásával. A monologikus beszédprodukciók között volt egy „kontrollált" narratíva, amikor az adatközlőnek egy a kísérletvezetővel közösen átélt élményt kellett elmesélnie; egy „szabad" spontán beszédprodukció, amikor a nyári élményeiket fogalmazták meg a beszélők; a harmadik feladat pedig irányított spontán beszéd volt, amelyben egy képsorozat alapján meséltek el egy történetet. Az eredmények szerint a különböző monológokban a (néma, kitöltött és kombinált) szünetek aránya beszélőfüggő volt, és nem a beszédtípusra volt jellemző. A társalgásban azonban nagyon hasonló eredmények születtek mindegyik adatközlőnél (azaz ebben az esetben kevésbé volt egyéni a szünethasználat), a társalgásban adatolt szünetarány mintegy harmada-negyede volt a monológokban mértnek. A szünetek időtartamai is szignifikáns különbséget mutattak a társalgás és a monologikus szövegek között (és egyes monologikus szövegtípusok között is), továbbá a tár-

salgásban 50%-kal csökkent a kitöltött szünetek előfordulási aránya minden beszélőnél.

Váradi Viola (2010) a felolvasás és a spontán beszéd temporális sajátosságait vetette össze hat fő, 25–40 éves kor közötti beszélő felvételei alapján. Az eredményei szerint a beszédtípus csak kismértékben befolyásolta a beszédszünetek időtartamát, amely inkább egyénfüggő sajátosság. A szünetek mindegyik beszélőnél nagyobb arányban és gyakorisággal jelentek meg a spontán beszédben, mint a felolvasásban; és nagy különbség volt a kitöltött szünetek előfordulásában is a két beszédtípus között: a felolvasásban egyáltalán nem jelent meg kitöltött szünet. Ebben a beszédtípusban ugyanis nincs nyelvi tervezés, így a felolvasásban a beszédszünet a spontán beszédtől eltérően csak három funkcióval bír: 1. a levegővételhez szükséges idő biztosítása; 2. értelmi tagolás; 3. a hallgató észlelésének és megértésének megkönnyítése (Váradi 2011).

Saját kutatásunkban azonos nyelvi tartalmú beszédanyagon vizsgáltuk a spontán beszéd és a felolvasás szupraszegmentális sajátosságait, így a szünettartást is (Bóna 2011b). Ebben a vizsgálatban is elemeztük mindegyik szünettípust. A kutatáshoz négy adatközlő spontán beszédét rögzítettük, majd lejegyeztük, és egy második ülésben ezt a lejegyzett hangfelvételt olvastattuk fel az adatközlőkkel. Mindegyikük a saját spontán beszédének a lejegyzett változatát hangosította meg a szöveg alapos átolvasása és néma értelmezése után. Az eredmények szerint mindegyik beszélőre jellemző volt, hogy a spontán beszédben gyakrabban tartott szünetet (3–4,2 szavanként), mint a felolvasásban (5,9–9 szavanként). Az adatközlők a spontán beszédben átlagosan 24,3%, a felolvasásban átlagosan 13,2% szünetet tartottak. A szünetek időtartamában két beszélőnél nem, két beszélőnél pedig szignifikáns különbség volt a kétféle beszédtípus között. A felolvasás során egyetlen beszélőnél adatoltunk kitöltött szünetet.

További vizsgálatainkban (Bóna 2010; 2012b) fiatal és idős adatközlők beszédprodukcióit vetettük össze spontán narratívában és hallott szövegek tartalmának visszamondásakor több más paraméter között a beszédszünetek szempontjából is. Az eredmények azt mutatták, hogy a tartalom-visszamondásban a szünetek előfordulásának gyakorisága mindkét életkori csoportban nőtt. Ebben a feladatban a szünetarány is szignifikánsan megnőtt mindkét életkori csoportnál a narratívához képest, a növekedés a fiataloknál nagyobb mértékű volt. A néma szüneteket elemezve az idősek hasonló időtartamú szüneteket produkáltak a narratívákban, mint a tartalom-visszamondáskor; a fiataloknál azonban szignifikánsan hosszabb néma szüneteket adatoltunk ez utóbbi beszédtípusban. A kitöltött szünetek aránya a teljes beszédidőhöz képest a fiatalok tartalom-visszamondásában matematikailag igazolható módon megnőtt (Bóna 2012b). Ennek az is lehetett az oka, hogy a fiatalok pontosabb interpretációra törekedtek, mint az idősek (Bóna 2011a).

Különleges beszédtípusok prozódiáját vetette össze Menyhárt Krisztina (2011). A Hegedűs-archívum felvételei közül ugyanazon mesemondó spontán beszédét és különböző műfajú meséit és elemezte: egy tündérmesét, egy reális mesét és egy állatmesét. Megállapította, hogy több más beszédparaméter mellett, a szünettartásban is jellegzetes eltérések vannak a különböző meseműfajok között.

A különböző beszédtípusok szünetjellemzőinek elemzése nemcsak fonetikai és pszicholingvisztikai szempontból érdekes, hanem a gyakorlat szempontjából is fontos lehet (Campione–Véronis 2002). Jelentősége van egyes beszédpatológiák (pl. dadogás, hadarás) megismerésében és terápiájának megtervezésében (vö. Bóna 2012c; Gósy–Bóna 2012) vagy a gépi beszéd spontán hangzásának megvalósítását célzó kutatásokban (vö. Zainkó et al. 2010).

A szünetekről tehát magyarra vonatkozóan is számos kutatás született, azok különböző beszédhelyzetekben való előfordulásáról, sajátosságairól. Ezek a kutatások azonban vagy kevés adatközlővel készültek (Markó 2005a; Váradi 2010; Bóna 2011b; Menyhárt 2011), és ezért bizonyos paraméterek (pl. a gyakoriság) nem voltak statisztikailag vizsgálhatók, illetve az egyéni sajátosságok miatt nem mutattak a beszédtípusra jellemző mintázatot; vagy pedig mindössze két beszédtípust vetettek össze (pl. Bóna 2012b). A jelen kutatás célja az, hogy relatíve nagyszámú adatközlő beszédfelvételeiben elemezze a szünettartás sajátosságait négy eltérő beszédtervezési működést igénylő beszédtípusban: 1. spontán narratívában; 2. hallott szövegek visszamondásakor; 3. társalgásban; 4. szövegfelolvasásban.

Hipotéziseink szerint az egyes beszédtípusok között szignifikáns különbséget találunk a szünetek 1. gyakoriságában, 2. a teljes beszédidőhöz viszonyított arányában, 3. az időtartamában, 4. a típusaiban, illetve 5. a hezitálásokat jellemző paraméterekben (gyakoriság, arány, időtartam) is.

Kísérleti személyek, anyag, módszer

A kutatáshoz a BEA magyar beszélt nyelvi adatbázisból választottuk ki 20 fő hanganyagát (Gósy 2012). Az adatközlők között 10 férfi és 10 nő volt, az életkoruk 20–31 év közé esett. Mindnyájan magyar anyanyelvűek, ép hallók voltak, a magyar köznyelvet beszélték.

A hangfelvételekből négy szövegtípust választottunk ki, az első a spontán narratíva volt. Ennek során az adatközlők a tanulmányaikról, munkájukról, illetve a hobbijukról beszéltek. A felvételt készítő személy csak nagyon ritkán tett fel nekik segítő kérdést, lehetőleg úgy, hogy a kérdésre az adatközlő minél hosszabban tudjon válaszolni.

A második, tartalom-visszamondási feladatban az adatközlők két rövid szöveget hallgattak meg: egy 174 szóból álló, 97 s időtartamú tudományos ismeretterjesztő szöveget; illetve egy 270 szóból álló 125 s időtartamú törté-

nelmi anekdotát. Az egyes szövegek meghallgatása után az volt az adatközlők feladata, hogy lehetőleg minél pontosabban ismételjék el a hallottakat.

A harmadik beszédtípus rögzítésekor a felvételkészítő mellett egy kollégája is jelen volt, és így egy háromszereplős társalgás alakult ki egy, a felvételkészítés idején aktuális témáról. Mivel a jelen kutatásban szereplő adatközlők és a felvételkészítők közel azonos életkorúak voltak, az elemzett társalgások oldott légkörben zajlottak, a résztvevők gyakran egymás szavába vágtak, közbebeszéltek.

A negyedik beszédtípusban az adatközlők egy 291 szóból álló tudománynépszerűsítő szöveget olvastak fel.

A narratívákból és a társalgásokból adatközlőnként átlagosan 4,5-5 percnyi hanganyagot dolgoztunk fel (a társalgásban sok rövidebb részletből állt össze az 5 perc, amely csak az adatközlő megszólalásait tartalmazta), a tartalomvisszamondás és a szövegfelolvasás időtartama adatközlőnként változott (1. táblázat). Összesen 280 percnyi, azaz 4,7 órányi hanganyagot elemeztünk.

1. táblázat: A feldolgozott hanganyagok időtartama beszédtípusonként

	Átlag (s)	Összes idő (s)
Narratíva	291	5820
Visszamondás	172	3440
Társalgás	256	5120
Szövegfelolvasás	120	2400

A felvételeket a Praat 5.0-val (Boersma–Weenink 2008) annotáltuk a szünetek és a beszédszakaszok szerint. Szünetnek tekintettünk minden olyan jelkimaradást, amelynél jól hallható volt a nyelvi jel folyamatosságának „megtorpanása", illetve valamilyen hezitációs jelenséget észleltünk; minimum időtartamot nem határoztunk meg (a legrövidebb időtartamú szünet 33 ms volt). A funkciótól függetlenül minden szünetet elemeztünk, így a légzési szünetet és a beszédtervezési nehézség miatt létrejövőt egyaránt. Meghatároztuk adatközlőnként és beszédtípusonként a szünetek számát, a szószámhoz és a jelidőhöz viszonyított gyakoriságát, a szünetek arányát a teljes beszédidőhöz képest, a szünetek típusait (néma, kitöltött, kombinált), a szünetek hosszát, a kitöltött szünetek arányát, gyakoriságát és időtartamát. Az eredményeket összevetettük a beszédtípusok szerint.

Az adatokon statisztikai elemzést végeztünk az SPSS 13.0 szoftverrel (egytényezős ANOVA, egyváltozós GLM, Tukey post hoc teszt, Kruskal–Wallis-próba, Mann–Whitney-próba) 95%-os konfidenciaintervallumban.

Eredmények

Összesen 6115 darab szünetet elemeztünk. Az egyes beszédtípusokban különböző gyakorisággal fordultak elő a beszédszünetek. A szószámra vetített gyakoriság, azaz a tagoltsági paraméterek szerint (vö. Olaszy 2005) két szü-

net között a legtöbb szót a társalgásban és a felolvasáskor ejtették az adatközlők, míg a legkevesebbet a tartalom-visszamondáskor (2. táblázat). Ez utóbbi beszédtípusban volt a legkisebb a beszélők közötti különbség a tagoltsági paraméterben. Az egytényezős ANOVA szerint szignifikáns különbség volt az egyes beszédtípusok között: $F(3,76) = 27,855$; $p \leq 0,001$; a Tukey post hoc teszt azonban megerősítette az átlagok által mutatott hasonlóságot: a társalgás és a felolvasás között nem volt szignifikáns a különbség. Nem különbözött statisztikailag a tagoltságban a narratíva és a társalgás sem (az összes többi esetben $p \leq 0,001$).

2. táblázat: A tagoltsági paraméter az egyes beszédtípusokban

Beszédtípus	Átlag	Átlagos eltérés	Minimum–maximum
Narratíva	5,2	1,05	3,3–7,6
Visszamondás	3,8	0,95	2,0–5,5
Társalgás	6,1	1,46	4,2–10,3
Felolvasás	6,9	1,07	4,2–9,8

A gyakoriságot úgy is megvizsgáltuk, hogy elemeztük az átlagos jelidőt, amely két szünet közé esett az egyes beszédtípusokban (ez tulajdonképpen a beszédszakaszok átlagos időtartamát jelenti). Eszerint két szünet között a leghosszabb jelidő a felolvasásban volt mérhető, a legrövidebb pedig a tartalom-visszamondásban (3. táblázat). A legkisebb egyéni különbségek ebben az esetben is a tartalom-visszamondásra voltak jellemzőek, míg a legnagyobb egyének közötti variabilitást – a tagoltsághoz hasonlóan – a társalgásban mértük. A Kruskal–Wallis-próba szerint szignifikáns különbség volt a beszédtípusok között: $\chi^2 = 49,584$; $p \leq 0,001$. A Mann–Whitney U-próba szignifikáns különbséget mutatott a narratíva és a tartalom-visszamondás ($Z = -3,138$; $p = 0,001$); a narratíva és a felolvasás ($Z = -5,221$; $p \leq 0,001$); a tartalom-visszamondás és a társalgás ($Z = -3,557$; $p \leq 0,001$), illetve felolvasás ($Z = -5,410$; $p \leq 0,001$) között; illetőleg a társalgás és a felolvasás között is ($Z = -4,734$; $p \leq 0,001$).

3. táblázat: A szünetek gyakorisága a jelidőhöz viszonyítva az egyes beszédtípusokban (ms)

Beszédtípus	Átlag	Átlagos eltérés	Minimum–maximum
Narratíva	1861	322	1432–2783
Visszamondás	1487	314	973–2050
Társalgás	2016	475	1319–3076
Felolvasás	3033	444	2077–4095

Az egyes beszédtípusokban adatolt tagoltsági paraméterek és a két szünettartás közötti időtartamok nem azonos arányban különböznek az egyes be-

szédtípusok között. Mivel erre elvileg kétféle magyarázat adható (1. a különböző beszédtípusokban különböző artikulációs tempóval beszéltek az adatközlők; 2. az egyes beszédtípusokban eltérő hosszúságú szavak fordultak elő), megvizsgáltuk a különböző beszédtípusokban mérhető artikulációs tempókat is (4. táblázat). Az egytényezős ANOVA ugyan szignifikáns különbséget mutatott a beszédtípusok között ($F(3,76) = 6,517$; $p = 0,001$); a post hoc teszt szerint csak a tartalom-visszamondás és a többi beszédtípus között volt szignifikáns különbség az artikulációs tempóban ($p \leq 0,033$). Ráadásul a társalgásban mért átlagos artikulációs tempó lassabb volt a felolvasásban adatoltnál. Ez azt jelenti, hogy nem az artikulációs tempó különbsége a magyarázat a kétféle gyakorisági érték eltérésére, hanem a szóhosszúság.

4. táblázat: Az artikulációs tempók az egyes beszédtípusokban (hang/s)

Beszédtípus	Átlag	Átlagos eltérés	Minimum–maximum
Narratíva	13,7	1,4	11,3–16,2
Visszamondás	12,5	1,4	9,8–15,0
Társalgás	13,8	1,4	10,9–16,7
Felolvasás	14,3	1,0	12,0–15,7

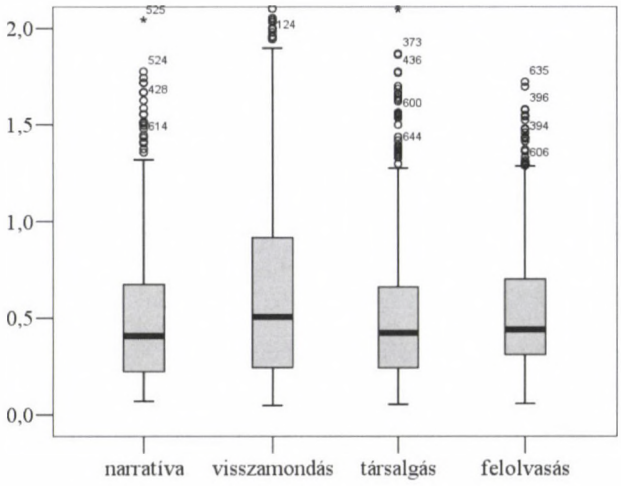
Eltérően alakult az összes szünetidőtartam teljes beszédidőhöz viszonyított aránya is az egyes beszédtípusokban (5. táblázat). A legnagyobb arányú szünettartást a tartalom-visszamondásokban adatoltuk, és itt volt a legnagyobb különbség az egyes beszélők között is. A legkisebb arányban a felolvasásban fordultak elő szünetek, ebben a beszédtípusban volt adatolható a legkisebb egyének közötti variabilitás is. A Kruskal–Wallis-próba szerint szignifikáns különbség volt a beszédtípusok között: $\chi^2 = 52,686$; $p \leq 0,001$; ezt a szignifikáns különbséget két-két beszédtípust páronként összehasonlítva is igazolta a Mann–Whitney U-teszt ($p \leq 0,002$).

5. táblázat: Az összes szünettartam aránya a teljes beszédidőhöz viszonyítva az egyes beszédtípusokban (%)

Beszédtípus	Átlag	Átlagos eltérés	Minimum–maximum
Narratíva	26,5	6,7	18,0–42,9
Visszamondás	34,1	7,8	21,5–46,3
Társalgás	19,6	5,4	11,1–29,9
Felolvasás	14,9	2,3	9,4–18,7

Az egyes beszédtípusokban előforduló szünetek időtartamának eloszlása az 1. ábrán olvasható (a kiugróan magas értékeket nem szerepeltettük). A leghosszabb szünetet a tartalom-visszamondásban mértük: ez 7993 ms volt. A leghosszabb átlagos szünetidőt is a tartalom-visszamondásban adatoltuk (781 ms); a legrövidebbet a felolvasásban (536 ms). A felolvasáséhoz hasonló

átlagos szünettartamot mértünk a társalgásban (589 ms), míg a narratívában adatolt átlag a társalgás és a tartalomösszegzés közé esett (673 ms). A Kruskal–Wallis-próba szerint a beszédtípusok között szignifikáns volt a különbség: $\chi^2 = 65{,}008$; $p \leq 0{,}001$. Két-két beszédtípust (a Mann–Whitney-teszttel) összehasonlítva azt találtuk, hogy a társalgásban és a felolvasásokban mért szünetek időtartamai között nincs szignifikáns különbség, míg minden más beszédtípus között van ($p \leq 0{,}001$).

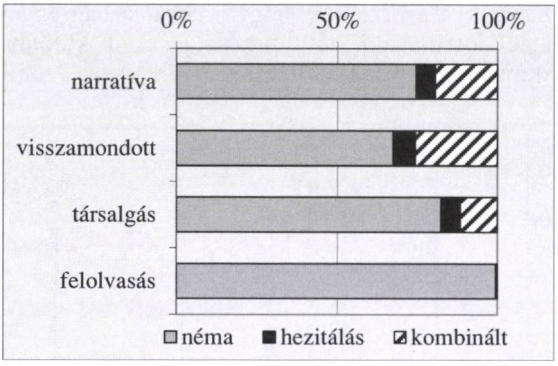


1. ábra
A szünetidőtartamok eloszlása (s)

A szünetek között a beszédtípustól függően különböző arányban fordultak elő a csak néma, a csak kitöltött és a kombinált szünetek (2. ábra). A felolvasásban mindössze egyetlen kitöltött (0,15%) és három kombinált (0,45%) szünetet produkált a húsz adatközlő. A legnagyobb arányban a tartalom-visszamondáskor fordultak elő kitöltött és kombinált szünetek (összesen a két szünettípus: 32,65%). Olyan kombinált szüneteket is adatoltunk, amelyekben több néma szünet és több hezitálás fordult elő. Például (a számok a néma szünetek és a hezitálások időtartamát jelzik): *és a vár két ööm (424) 482 ööm (503) 1302 támadást is vissza tudott verni.*

Összevetettük a szünetidőket a szünettípus és a beszédtípus szerint is (6. táblázat). Mindhárom szünettípusból a tartalom-visszamondásban adatoltuk a leghosszabb átlagos időtartamot, míg a legrövidebb átlag és a legkisebb átlagos eltérés a felolvasásra volt jellemző. Az egyváltozós GLM elemzés szerint ugyan szignifikáns hatása van a beszédtípusnak is a szünethosszra, de ez csak mindössze 3%-ban magyarázza az eredményeket (parciális η^2 =

0,003; $p \leq 0{,}001$). Erősebben befolyásoló tényező a szünet típusa (parciális $\eta^2 = 0{,}141$; $p \leq 0{,}001$); míg a két tényező együttes hatása igen kicsi (parciális $\eta^2 = 0{,}002$; $p = 0{,}012$).



2. ábra
A szünettípusok előfordulásának aránya a különböző beszédtípusokban

6. táblázat: Szünetidőtartamok a szünettípus és a beszédtípus szerint

Néma szünet (ms)			
	Átlag	Átlagos eltérés	Minimum–maximum
Narratíva	580	503	33–4497
Visszamondás	619	558	47–5763
Társalgás	520	417	54–4200
Felolvasás	537	374	57–4109
Kitöltött szünet (ms)			
	Átlag	Átlagos eltérés	Minimum–maximum
Narratíva	374	254	94–1658
Visszamondás	384	242	80–1322
Társalgás	334	156	81–952
Felolvasás	109	–	–
Kombinált szünet (ms)			
	Átlag	Átlagos eltérés	Minimum–maximum
Narratíva	1142	794	89–5566
Visszamondás	1312	960	144–7993
Társalgás	1062	652	208–3693
Felolvasás	402	171	206–516

Külön elemeztük a hezitálások (kitöltött szünetek; 1468 darab) gyakoriságát, az összes szünetidőhöz viszonyított arányát, illetve időtartamait – függetlenül attól, hogy önálló kitöltött szünetként vagy kombinált szünetként jelentek-e meg. A felolvasásokban mindössze két nőnél és két férfinál fordult elő egy-egy hezitálás. Az egyik női adatközlő a társalgás során egyszer sem hezitált. A hezitálások gyakoriságában a narratíva, a tartalom-visszamondás és a társalgás között szignifikáns különbség volt: $\chi^2 = 27{,}247$; $p \leq 0{,}001$. Két-két beszédtípust összevetve, a Mann–Whitney U-próba is szignifikáns különbséget mutatott ($p \leq 0{,}001$). A narratívákban átlagosan 25 szavanként (szórás: 13,6), a tartalom-visszamondásban 14,1 szavanként (szórás: 11,4), a társalgásban (nem számítva azt az adatközlőt, akinél egyáltalán nem fordult elő kitöltött szünet) 51,8 szavanként (szórás: 35,1) adatoltunk hezitálást.

A hezitálásoknak az összes szünetidőhöz viszonyított aránya is változott a beszédtípus függvényében: a narratívákban átlagosan 14,7% (szórás: 7,4%), a tartalom-visszamondásban átlagosan 17,9% (szórás: 8,8%), a társalgásban 11,7% (szórás: 5,8%) volt. Az egytényezős ANOVA szerint e három csoport között szignifikáns volt a különbség: $F(2,57) = 3{,}524$; $p = 0{,}036$. A post hoc teszt szerint azonban csak a tartalom-visszamondásban és a társalgásban mért időarány különbözött szignifikánsan ($p = 0{,}027$). A hezitálások időtartamában nem volt szignifikáns különbség a beszédtípusok között. A narratívában a hezitálások időtartama átlagosan 378 ms (szórás: 209 ms), a tartalom-visszamondásban 391 ms (szórás: 204 ms), a társalgásban 373 ms (szórás: 189 ms), a felolvasásban 190 ms (összesen négy előfordulás; a szórás: 116 ms) volt.

Következtetések

Kiinduló hipotéziseink, amelyek szerint az egyes beszédtípusok között szignifikáns különbséget találunk minden vizsgált paraméterben, részben igazolódtak. A beszédtervezési folyamatok különbsége a szünettartásban is jelentkezik, de vannak olyan beszédtípusok, amelyek a szünetek tekintetében hasonlítanak egymásra, míg mások nagyobb mértékben különböznek. Másrészt egyes beszédtípusokban nagyobb mértékben jelentkeznek az egyéni sajátosságok, mások kötöttebb műfajnak tekinthetők, s így kisebb az egyéni variabilitás a szünettartásban.

A beszédtervezés szempontjából a legnehezebb feladat a tartalom-visszamondás (Bóna 2010; Gósy 2010), ami az elemzett paraméterekben is megmutatkozott. Ebben a beszédtípusban adatoltuk a leggyakoribb, illetve a leghosszabb szüneteket, és a kitöltött és a kombinált szünetek aránya is itt volt a legnagyobb. Mivel a hallott szövegek felidézését számos, egyéntől függő sajátosság (mint pl. a beszédfeldolgozás sikeressége, a figyelem, a memóriaműködés, illetve a narratív kompetencia; Juncos-Rabadán–Pereiro 2000) befolyásolja, ebben a beszédtípusban adatoltuk a legnagyobb egyéni különbségeket is. Azaz attól függően, hogy a beszélő milyen pontosan akarta visszaadni a hallott szövegek tartalmát, mit értett meg belőlük, illetve mire

emlékezett, változtak a szünettartás sajátosságai is. Ezt megerősítik egy korábbi kutatásunk eredményei, amelyben fiatalok, idősödők és idősek tartalom-visszamondásait elemeztük. Abban a vizsgálatban csak egyetlen beszédtípust elemeztünk, és azt az eredményt kaptuk, hogy a fiatalok, akik szignifikánsan több tartalmi egységet tudtak felidézni pontosan a hallott szövegekből, nagyobb arányban produkáltak néma és kitöltött szünetet, mint a másik két életkori csoport (Bóna 2011a).

Az artikulációsan és akusztikailag is leginkább meghatározott beszédtípus a felolvasás (Wacha 1974; Markó 2012; Váradi 2011, 2012), ami a szünetezésben is megmutatkozott. Mivel a felolvasás során nem egyszerre történik a fogalmi és a nyelvi tervezés az artikulációs tervezéssel, mert a meghangosítandó szöveg adott, ezért a felolvasónak több ideje és mentális energiája marad a szünettartás megtervezésére is (Gósy 2005). Ezt a beszédtípust a többinél rövidebb, ritkábban tartott szünetek jellemezték, a hezitálások pedig csak minimális arányban fordultak elő benne. Az egyéni különbségek is itt voltak a legkisebbek. Nem várt eredmény, hogy több paraméterben is hasonlított egymásra a társalgás és a felolvasás, így a szünetek szószámra vetített gyakoriságában, illetve az időtartamukban. E két beszédtípusban a beszélők bizonyos szempontból hasonló szünettartási stratégiákat alkalmaztak, ennek hátterében azonban eltérő okok állhatnak. Amíg a felolvasásban nincs szükség a magasabb tervezési szintek működtetésére, addig a társalgásban a megszólalás előtt már gyakran lehetőség van a mondanivaló megtervezésére. Ráadásul a (vizsgált) társalgásokban relatíve rövidek voltak a beszédlépések, vagyis a beszélő átadhatta a szót, ha nem volt több mondandója, vagy az hosszabb tervezési időt (és így több szünetet) igényelt volna.

Adataink egyrészt megerősítik a korábbi, kisszámú (4-6) adatközlővel végzett vizsgálatok (Markó 2005; Váradi 2010; Bóna 2011b) eredményeit, másrészt a nagyobb számú (20 fő) adatközlő miatt lehetőség nyílt olyan paraméterek statisztikai elemzésére, amelyekről korábban tendenciaszerű megállapításokat tettek a kutatók (pl. gyakoriság, időarány). Vizsgálatunk felhívja a figyelmet arra is, hogy egyes alkalmazott fonetikai kutatásoknál (például az életkori sajátosságok vizsgálatánál, a női-férfi beszédjellemzők közötti különbségek elemzésénél, a beszédpatológiás kutatásoknál) szükség lehet az eltérő beszédtípusok elemzésére is. (A hadarás vizsgálatakor például teljesen más eredményeket kaphatunk, és másként ítélhetjük meg a hadarás súlyosságát, ha a beszélőtől spontán narratívát vagy történetmesélést, netán szónoki beszédet rögzítünk, Bóna 2012c). A mindennapi kommunikáció során a különböző (általunk is elemzett) beszédtípusok mind megjelenhetnek, és együttesen határozzák meg az egyén beszédkészségéről, kommunikációjáról alkotott képünket.

Irodalom

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

Boersma, Paul – Weenink, David 2008. Praat: doing phonetics by computer (Version 5.0.1). http://www.fon.hum.uva.nl/praat/download_win.html

Bóna Judit 2006. A megakadásjelenségek akusztikai és percepciós sajátosságai. *Beszédkutatás 2006*. 101–113.

Bóna Judit 2009. *A gyors beszéd. Produkciós és percepciós sajátosságok*. MTA Könyvtára–Lexica Kiadó, Budapest.

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

Bóna Judit 2011a. A beszédpercepció és a beszédprodukció összefüggései fiatal, idősödő és idős korban. *Gyógypedagógiai Szemle* XXXIX/3–4. 221–232.

Bóna Judit 2011b. A különböző beszédstílusok az akusztikai-fonetikai és percepciós vizsgálatok tükrében. *Alkalmazott Nyelvtudomány* 11. 39–48.

Bóna Judit 2012a. A spontán beszéd sajátosságai idősödő, idős és matuzsálemi korban. In Markó Alexandra (szerk.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngekezdési időig*. ELTE BTK–MTA Nyelvtudományi Intézet, Budapest, 100–115.

Bóna Judit 2012b. *Az időskori beszéd fonetikai sajátosságai*. Kézirat. Budapest.

Bóna Judit 2012c. Linguistic-phonetic characteristics of cluttering across different speaking styles: A pilot study from Hungarian. *Poznań Studies in Contemporary Linguistics* 48/2. 203–222.

Campione, Estelle – Véronis, Jean 2002. A large-scale multilingual study of silent pause duration. *Speech Prosody 2002*. Aix-en-Provance, France. 199–202.

Duez, Danielle 1982. Silent and non-silent pauses in three speech styles. *Language and Speech* 25. 11–25.

Fónagy Iván 1967. Áthajlás, szünet, szerkezet. *Nyelvtudományi Közlemények* LXIX. 313–343.

Goldman-Eisler, Frieda 1968. *Psycholinguistics: Experiments in spontaneous speech*. Academic Press, London.

Gósy Mária 2000. A beszédszünetek kettős funkciója. *Beszédkutatás 2000*. 1–15.

Gósy Mária 2003. 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 2010. Szövegértés alapú narratívák. In Bárdosi Vilmos (szerk.): *Világkép a nyelvben és a nyelvhasználatban*. Tinta Kiadó, Budapest. 113–124.

Gósy Mária 2012. Multifunkcionális beszélt nyelvi adatbázis – BEA. *Általános Nyelvészeti Tanulmányok* XXIV. 329–349.

Gósy Mária – Beke András – Horváth Viktória 2011. Temporális variabilitás a spontán beszédben. *Beszédkutatás 2011*. 5–30.

Gósy, Mária – Bóna, Judit 2012. A case study on the effect of discourse type on fluency level in stuttering. *The Phonetician* 103–104/1–2. 57–76.

Gósy Mária – Bóna Judit – Beke András – Horváth Viktória 2013. A kitöltött szünetek fonetikai sajátosságai az életkor függvényében. *Beszédkutatás 2013*. 121–143.

Hegedűs Lajos 1957. Beszédtempó-elemzések. *Magyar Nyelvőr* 81. 223–227.

Hieke, Adolf E. – Kowal, Sabine – O'Connel, Daniel C. 1983. The trouble with „articulatory" pauses. *Language and Speech* 26. 203–219.
Horváth, Viktória 2010. Filled pauses in Hungarian: Their phonetic form and function. *Acta Linguistica Hungarica* 57/2–3. 288–306.
Imre Angéla 2005. Különböző műfajú szövegek szupraszegmentális jellemzői. *Magyar Nyelvőr* 129. 510–520.
Juncos-Rabadán, Onésimo – Pereiro, Arturo X. 2000. Telling stories in the elderly. Influence of attentional and working memory proceses (preliminary study). In *Psycholinguistics on the threshold of the year 2000*. 155–159.
Kassai Ilona 1988. A szünet kérdésköre a szöveglejegyzésben. In Kontra Miklós (szerk.): *Beszélt nyelvi tanulmányok. Linguistica, Series A, Studia et Dissertationes* 1. MTA Nyelvtudományi Intézet, Budapest. 22–43.
Kassai Ilona – Fagyal Zsuzsanna 1996. Hogyan észlelik a magyar beszéd szüneteit magyar és francia anyanyelvű hallgatók? *Magyar Nyelvőr* 120. 209–220.
Krivokapic, Jelena 2007. Prosodic planning: Effects of phrasal lenght and complexity on pause duration. *Journal of Phonetics* 35/2. 162–179.
Laczkó Mária 2009. Tizenévesek beszédének fonetikai és stilisztikai elemzése. *Anyanyelv-pedagógia* 2009/1. http://www.anyanyelv-pedagogia.hu/cikkek.php?id=151
Levelt, Willem J. M. 1989. *Speaking: From intention to articulation*. A Bradford Book. The MIT Press, Cambridge (Massachusetts)–London (England).
Mahl, George F. 1956. Disturbances and silences in the patient's speech in psychotherapy. *Journal of Abnormal and Social Psychology* 53. 1–15.
Markó Alexandra 2005a. *A spontán beszéd néhány szupraszegmentális jellegzetessége*. PhD-értekezés. ELTE, Budapest.
Markó Alexandra 2005b. A temporális szerkezet jellegzetességei eltérő kommunikációs helyzetekben. *Beszédkutatás 2005*. 63–77.
Markó Alexandra 2012. Az irreguláris zönge funkciói és gyakorisága olvasott és spontán beszédben. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest. 25–42.
Menyhárt Krisztina 1998. Nyelvi meghatározottság a beszédszünetek észlelésében. *Beszédkutatás '98*. 47–57.
Menyhárt Krisztina 2010. A beszédsebesség objektív mérési és szubjektív észlelési eredményeinek összefüggései mai és 60 évvel ezelőtti beszélőknél. *Beszédkutatás 2010*. 110–124.
Menyhárt Krisztina 2011. Régi mesék prozódiája: Palkó Józsefné meséi. *Beszédkutatás 2011*. 96–108.
Menyhárt Krisztina 2012. A beszéd temporális jellemzői 60 évvel ezelőtti gyermek beszélőknél. *Beszédkutatás 2012*. 246–259.
Olaszy Gábor 2005. Prozódiai szerkezetek jellemzése a hírfelolvasásban, a mesemondásban, a novella és a reklámok felolvasásában. *Beszédkutatás 2005*. 21–50.
Rose, Ralph Leon 1998. *The communicative value of filled pauses in spontaneous speech*. PhD thesis. University of Birmingham.
Sallai János – Szende Tamás 1995. Spontán közlések beszédszüneteinek pszicholingvisztikai értelmezése (egészséges és skizofrén közlők szövegeiben). *Általános Nyelvészeti Tanulmányok* XVIII. 209–222.
Strangert, Eva 2003. Emphasis by pausing. In: *Proceedings of the 15h International Congress of Phonetic Sciences*. Barcelona. 2477–2480.

Szende Tamás 1976. *A beszédfolyamat alaptényezői*. Akadémiai Kiadó, Budapest.
Vallent Brigitta 2005. A spontán beszéd ötven éve és ma. Esettanulmány. *Beszédkutatás 2005*. 99–111.
Vallent Brigitta 2010. Középiskolások spontán narratíváinak jellemzői. *Beszédkutatás 2010*. 199–210.
Váradi Viola 2010. A felolvasás és a spontán beszéd temporális sajátosságainak összehasonlítása. *Beszédkutatás 2010*. 100–109.
Váradi Viola 2011. A felolvasás és a spontán beszéd összevetésének pedagógiai vonatkozásai. *Anyanyelv-pedagógia* 2011/4. http://www.anyp.hu/cikkek.php?id=343
Váradi Viola 2012. A felolvasás és a spontán beszéd tagolhatósága prozódiai egységek jellemzése alapján. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest. 108–120.
Volkskaya, Nina B. 2003. Virtual and real pauses at clause and sentence boundaries. In: *Proceedings of the 15h International Congress of Phonetic Sciences*. Barcelona. 499–502.
Wacha Imre 1974. Az elhangzó beszéd főbb akusztikus stíluskategóriáiról. *Általános Nyelvészeti Tanulmányok* X. 203–216.
Zainkó, Csaba – Csapó, Tamás Gábor – Németh, Géza 2010. Special speech synthesis for social network websites. *Text, Speech and Dialogue, Lecture Notes in Computer Science* 6231. 455–463.
Zellner, Brigitte 1994. Pauses and the temporal structure of speech. In Keller, Eric (ed.): *Fundamentals of speech synthesis and speech recognition*. John Wiley, Chichester. 41–62.

THE INTONATION OF SPANISH YES-NO QUESTIONS IN SPONTANEOUS SPEECH AND IN A DIDACTIC MATERIAL

Baditzné Pálvölgyi Kata – Dolors Font-Rotchés

Introduction

This work presents a comparison between a prestigious didactic material and real-life communication from an intonational point of view. We compare 70 contours that appeared in the Spanish textbook, *El ventilador* (Chamorro et al. 2006) with 80 contours of spontaneous speech, collected between 1999 and 2011 from Spanish television programmes.

The theoretical framework of the data analysis follows the *Melodic Analysis of Speech (MAS)* theory, proposed by Cantero (2002), whose protocol is explained in detail in Font-Rotchés and Cantero (2009). Their method is characterized by analysing spontaneous speech (with software such as Praat). The first step is to get rid of insignificant micromelodic variations. Every tonally stable vowel is measured at its mid F_0 value, except for the last vowel in the utterance, which is measured at its last F_0 value. If there is a tonal difference of at least 10% within one vowel, it is called an inflection and is given two F_0 values in the representation. Then the melody is standardized, so that the intonation of different speakers could be compared objectively, by converting the first F_0 value to the arbitrary number 100, and all the other values are represented in % (with respect to the previous value). The standardized melody is submitted to perceptive tests, in order to verify whether the contour is melodically identical to the original one.

Every contour has its significant melodic traits, for instance its First Peak, or the movement that characterizes its Final Inflection (FI, the tonal movement starting from the last stressed syllable).

A melodic pattern has its characteristic binary features (/±interrogative/, /±suspended/ and /±emphatic/) and can be realized within certain dispersion margins. The patterns are defined with perceptive tests again. Listeners are given various contours and they decide after hearing them whether they sound /+interrogative/, /+suspended/ and /+emphatic/. Contours that share the same binary features are examined by analysing their melodic traits, and thus, the dispersion margins of the melodic patterns are settled.

Earlier works that presented the intonational patterns found in yes-no questions insisted on one existing typical melody: a pattern with a high rise at the end. On the universality (or at least the characteristic majority) of rising FIs

for Spanish interrogative intonation, see Alcoba and Murillo (1998: 160), Congosto et al. (2008: 9), Face (2007: 194), Fernández et al. (2003: 191, 197), Henriksen (2010), Jiménez Gómez (2010: 296–297), Martínez Celdrán (2003: 72, 74), Quilis (1995: 429–431), Sosa (1999: 30) and tacitly in Zamora et al. (2005: 127). Martínez Celdrán (2003: 78) agrees with Cantero (2002: 168–169) that in peninsular Spanish, absolute interrogatives must have a rising inflection starting from a low Syntagmatic Accent (= Nucleus, the last stressed syllable). Sosa (1999), with an autosegmental treatment, found only rising yes-no questions in Spanish. A possible reason for this is that he made the informants read out yes-no questions. In read yes-no questions, presumably, speakers are cautious to read out the "expected, normative" version. Canellada and Madsen (1987) also found only rising FIs in yes-no interrogatives. Still, as Sosa (1999: 211) remarks, Cunningham (1983) found only high falling FIs in Oviedo Spanish. Díaz Tejera (1973: 106) also permits a circumflex intonation for yes-no interrogatives. The only significant work which permits various intonational patterns that occur in Spanish yes-no questions is Navarro Tomás (1966).

The descriptions we use here as model also reveal that Spanish questions do not exclusively have a rising FI. The presentation of the Spanish intonational patterns found in the corpora is based on Cantero–Font-Rotchés (2007) and on Font-Rotchés–Mateo (2011, and in press). There are four Spanish interrogative intonational patterns found until nowadays since the application of the *MAS* method, according to the research realized by the *Grup de Recerca en Entonació i Parla* (Intonation and Speech Research Group, 2009 SGR 00233) from the Applied Phonetics Laboratory of the University of Barcelona (http://www.ub.edu/lfa/).

The scope of our research is to examine whether a widely-taught method reflects the linguistic reality that characterizes Spanish yes-no question intonation, or it only offers a reduced panorama of intonational phenomena related to yes-no interrogatives. This kind of interrogatives often cause misunderstanding between native speakers of Spanish and foreigners, as the Spanish do not understand when they are asked a question, due to its inadequate intonation. It can be seen what sort of problems can be connected to inappropriate intonation by Hungarian (Baditzné in press), Italian (Devís 2011), Brazilian (Fonseca–Cantero 2011), Swedish (Martorell 2011) or Chinese (Liu 2005) learners of Spanish.

The methodology of data analysis

This section will concentrate on two basic constituents of the applied methodology: the structure of the intonational unit and the representation of the melodies. An empirical and experimental-based methodology has been used, focusing on the assumptions of the *Melodic Analysis of Speech (MAS)* theory, expounded in detail in Cantero (2002), revised in Font-Rotchés

(2007), and established as a protocol in Cantero–Font-Rotchés (2009) and Font-Rotchés–Cantero (2009).

Cantero and Font-Rotchés take their intonational data from spontaneous speech and accept nothing but the phonetic reality as the basis of their analysis. They analyse the intonation of each utterance with a voice analysis-and-synthesis programme, and standardize the received results – the F_0 value of each syllable – so that the intonation of the sentences by different speakers could be compared. The following points will detail the process of data analysis and representation.

The structure of the melodic unit

The elements of the intonational unit for Cantero (2002) are the following: *Anacrusis, Body* and *Final Inflection (FI)*. By *Anacrusis* he means the unaccented syllables preceding the *First Peak,* which is normally on the first accented vowel in the contour but can also be displaced to the left or to the right from the first lexical stress. This implies that the Anacrusis can contain lexically stressed syllables when the First Peak is shifted to the right of the first lexical stress. He defines as *Body* the syllables between the First Peak and the last accented vowel in the contour (the latter also known as the *Nucleus*), from which the *Final Inflection* begins, see Figure 1. Thus, there are two important accents that separate these three parts, the First Peak and the Nucleus.

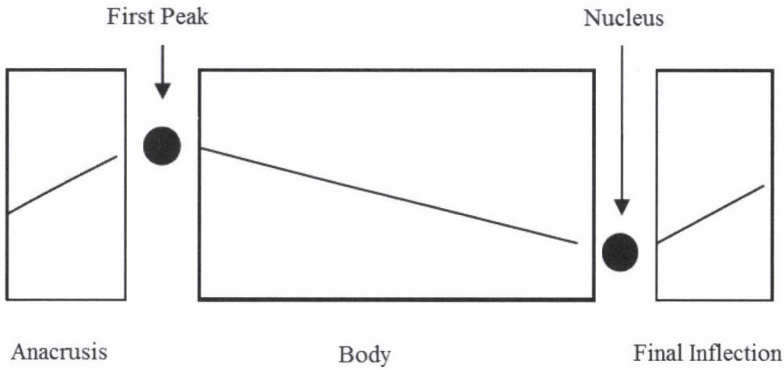


Figure 1.
The structure of the intonational unit according to Cantero (2002)

The process of obtaining F_0 values

Cantero defines the tonal segment as "every more or less stable tonal stage which is clearly perceptible and the duration of which coincides with one *mora*" (2002: 89). Thus, each vowel has minimally one tonal segment, that is, it has the duration of one mora, and it may have two or three if it constitutes

an inflection. If three, the inflection is circumflex (for Cantero *circumflex* is two-directional, i.e. both rising-falling and falling-rising), which means that it has two tonal movements, with three tonal segments altogether. This is why an inflection is the succession of at least two distinct and contiguous tonal segments.

Figure 2 below shows how the distinct tonal values can be perceived in the utterance *¿Hace deporte?* 'Do you do any sport?'. The diagrams are produced by the voice analysis software Praat, version 4.0 (the text is our addition). We can see five syllables, each of which has a tonally stable vowel, and there is a bigger pitch jump between the penultimate and the last syllable.

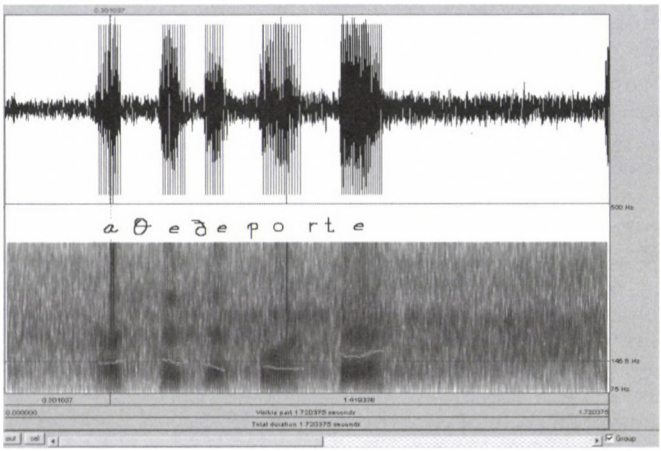


Figure 2.
Pitch values of the vowels in *¿Hace deporte?* 'Do you do any sport?'

In Figure 3 we can see a schematic graphic display showing the whole curve with the relevant pitch heights measured at each vowel. While in the case of last vowels it is the endpoint of the vowel, in the case of other vowels it is the mid-point of the vowel's duration where the pitch-height is measured.

As can be seen in Figures 2 and 3, there is tonal stability (i.e. no or hardly any pitch change) in the first three vowels, but there is a jump, that is, tonal instability (i.e. radical pitch change) between the fourth and the fifth vowels, which characterizes the Final Inflection.

Although the reduction of every syllable to one pitch point can solve the problem of redundant micromelodic variations in the curve, but we still have to face the problem of the significance of each movement. This is a question of considerable importance, as tonal movements of a certain size can convey different meanings to the listener: a certain amount of inflection in Spanish suggests that the utterance is a question, for example. But it would be impossible

to generalize this amount if we did not standardize the melodies. Let us imagine, for example, the same utterance produced by a male adult and a child. Obviously, at first the two versions seem to be very different, as the child pronounces the melody with much higher pitch; even the same movements may seem steeper. We cannot compare these melodies unless we have a reliable system of data standardization. Our next point will explain this step.

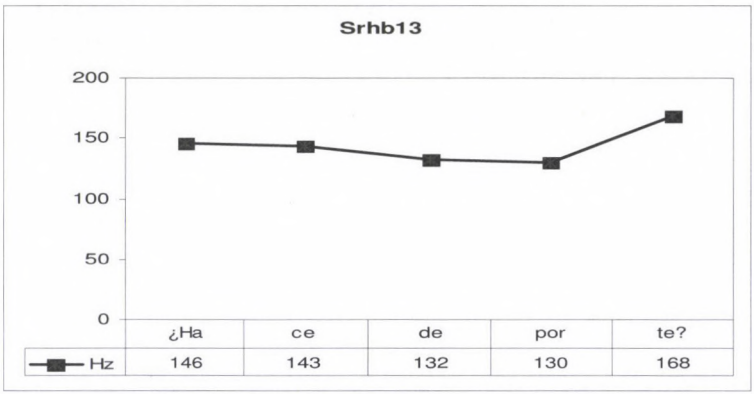
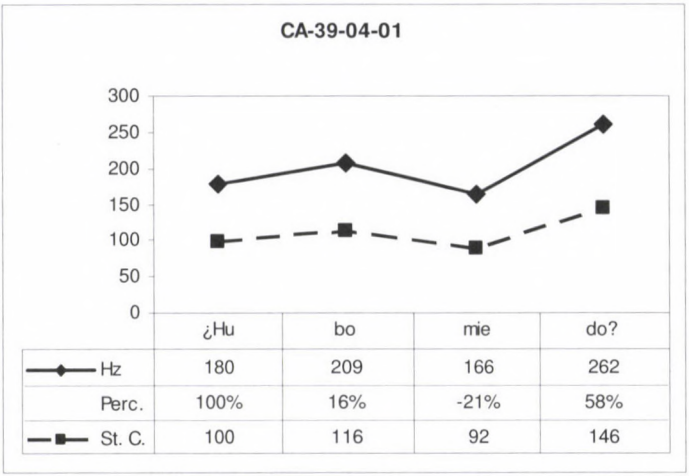


Figure 3.
The representation of the pitch values of all five vowels in the utterance in *¿Hace deporte?* 'Do you do any sport?'

The standardization of contours

The standardization is a following but also necessary step in the intonational analysis. With the help of standardized contours, we will have melodies independent of the speaker's sex or age, as well as of the idiosyncratic factors of the speaker's voice.

The standardization procedure consists in representing the F_0 value of each syllable as a percentage relative to the F_0 value of the preceding syllable, while giving the first syllable the arbitrary percentage of 100%. For example, if in the sentence of Figure 4, *¿Hubo miedo?* 'Was there fear?', the measured F_0 value in Hertz (Hz) for the first syllable is 180 Hz, and for the second syllable, 209 Hz, then it means a rise of 16.1% between the first and the second syllables (in Perc. = percentages); if the first syllable is given a value of 100 as a relative value, the second syllable must be given 116. In Figure 4 the original curve with F_0 values is shown with dotted line and the standardized curve (St. C.), with a solid line. The standardized curve is melodically identical to the original one, but with a different tessitura.



	¿Hu	bo	mie	do?
Hz	180	209	166	262
Perc.	100%	16%	-21%	58%
St. C.	100	116	92	146

Figure 4.
Diagram of the original and the standardized contour of *¿Hubo miedo?* 'Was there fear?'

The perceptive phase: the validation of standardized contours and data interpretation

After the acoustic phase, the validity of the results is checked with a series of perceptive tests submitted to the listener's judgment from an exact copy of the analyzed sentences. This synthesized copy is achieved by the programme *Praat* (with PSOLA method), in which all original relative pitch values are erased and then replaced by the standardized values. This way it can be checked if the melodic analysis has been correct and if it reflects the original melody, without micromelodic variations and with the normalized values.

The following step is the interpretation of the obtained melody. In this theoretical pattern, there are three phonological features that are studied: /± interrogative/ /± emphatic/ and /± suspended/, whose combination allows us to characterize the language *tonemes*. In these perceptive tests, the listeners are asked whether the utterances they hear (with the standardized F_0 values already) sound to them as questions, that is, /± interrogative/; as finished, i. e. /± suspended/, and as exclamations, that is to say, /± emphatic/.

The melodic features are, in turn, the characteristic of the functional elements of the contour: *Anacrusis,* the *First Peak,* the *Body* and the *Final Inflection (FI).* The description of these elements (especially the description of the FI) allows them to define the contour melody and to establish the typical melodic patterns in a corpus (as "typical contours" of the tonemes) and their dispersion margins.

The /+interrogative/ intonational patterns in Spanish

There are four /+interrogative/ patterns and other non-interrogatives ones which can be used for making yes-no questions. The difference between these two groups is that while in the case of /+interrogative/ patterns the melody is decoded as interrogative by listeners independently of the context, non-interrogative intonational patterns can only be used for asking yes-no questions if the context it permits.

In this work we will give predominance to the /+interrogative/ intonational patterns, because from a didactic point of view, we consider it essential to teach students the /+interrogative/ patterns first, as they will be recognized easily as interrogatives by the listeners independently of the context.

Now let us see which are the /+interrogative/ patterns in Spanish. The following inventory (Table 1) first gives a schematic figure about the characteristics of the pattern in question, and that in each case we resume the melodic traits in the following column.

The corpus

The spontaneous corpus was collected by Cantero et al. (2002), Mateo (unpublished PhD dissertation) and Ballesteros (2011). It contains more than 2500 spontaneous sentences produced by anonymous speakers between 1999 and 2011 in various Spanish television programmes. According to Cantero (2002), in television programmes we can get good-quality spontaneous utterances where the speakers are unaware of being recorded for the purposes of a linguistic investigation.

For this article, we selected 80 yes-no questions of this spontaneous corpus produced by 72 speakers (27 men and 45 women, whose age is between 16 and 83 years). In Table 2 we reproduce some examples from these sentences with their English translation, classified according to their melodic pattern.

The textbook corpus was collected exclusively from *El ventilador* (Chamorro et al. 2006). The choice of this particular textbook is explained by the following facts: first, it is a widely known and used material throughout the world. Second, it aims at preparing for a C1 level (CEFR) command of language, which induces us to believe that students are familiarized with a wider palette of intonational solutions than, for example, in the case of an A2 or a B1 level textbook they would. Given that there are few advanced (C1 or higher) level Spanish complex textbooks (the purpose of which is not the preparation for a special aptitude test, but to reach that level in all competences), *El ventilador* is a reasonable choice.

Altogether there are 70 yes-no questions analysed, all of which represent the standard peninsular Spanish dialect. Thus, only Latin-american realisations of yes-no questions were excluded from the investigation, but otherwise the corpus includes all the yes-no questions which, according to the transcription of the textbook, are meant to be yes-no questions.

Table 1: Spanish /+interrogative/ patterns

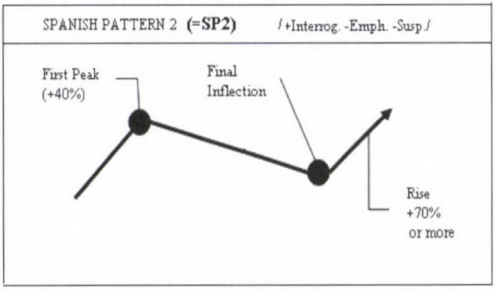
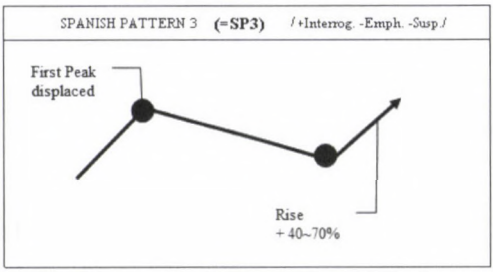
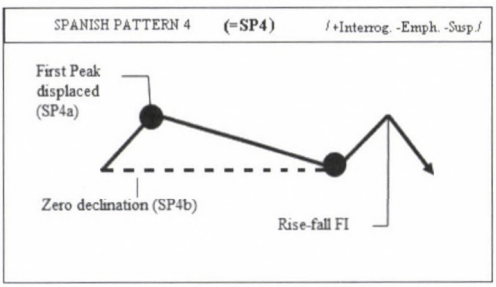
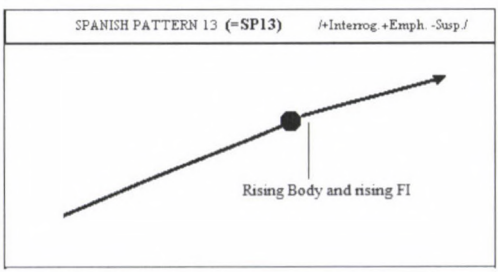
Patterns	Description
	Anacrusis: Rise till the First Peak of up to 40%. First Peak: The first stressed syllable, at the first highest point of the contour. Body: Smooth descending. Final Inflection: A rise of more than 70%.
	Anacrusis: Rise till the First Peak of up to 40%. First Peak: Displaced to the unstressed syllable following the first stress. Body: Smooth descending. Final Inflection: A rise of 40-70% (originally with the maximum of 60%, but according to recent studies up to 70%, cf. Font-Rotchés–Mateo in press).
	Anacrusis: Rise till the First Peak of up to 40% (SP4a) or level (SP4b). First Peak: Displaced to the unstressed syllable following the first stress (SP4a) or none (SP4b). Body: Smooth descending (4a) or zero declination (4b). Final Inflection: Rise-Fall.
	First Peak: None. Anacrusis: None. Body: Rising. Final Inflection: Rising.

Table 2: Examples from the spontaneous corpus

Code	Spanish sentences and translation
PATTERN 2 /+INTERROGATIVE/	
CM-10-05-01	¿Me lo puedes repetir? 'Can you repeat it to me?'
PV-01-23-02	¿Te lo has probado? 'Have you tried it?'
PATTERN 3 /+INTERROGATIVE/	
AI-15	¿Tú le has visto su cuerpo? 'Have you seen his body?
EX-46-09-05	¿Tienes permiso de conducir? 'Do you have driving licence?'
PATTERN 4 /+INTERROGATIVE/	
EX-44-03-09	¿Eso se está haciendo? 'Is it being done now?'
MU-01-11-05	¿Te gusta el queso? 'Do you like cheese?'
PATTERN 13 /+INTERROGATIVE/	
AN-23-04-06	¿Tienes hijos? 'Do you have children?'
AN-24-02-06	¿Estás grueso? 'Are you stout?'
/-INTERROGATIVE/ PATTERNS	
CA-59-02-01	¿Un día de trabajo nada más? 'A working day, nothing more?'
PV-03-05-04	¿Puedes decir algo? 'Can you say something?'

In Table 3 we reproduce the questions in the textbook corpus, with their English counterparts. All of them form part of dialogues of the sound material of the textbook (with the exception of the last one, which is an interrogative in a monologue), which have been read by 15 speakers – 9 men and 6 women. The speakers knew the aim of these dialogues: they were read to be listened by learners of Spanish. There isn't in the textbook any explanation about the intonation of the interrogatives, nor some specialised exercises in order to practise them. The students only can listen the yes-no questions in these dialogues.

The results of our investigation

Now let us compare the patterns found in each corpus. The following tables contain the number of the patterns found and also their proportion. First there are the four interrogative patterns examined (in the case of variants of SP4, the subtypes are counted in the same category), then they are summed up, and finally the proportion of non-interrogative patterns is exposed.

If we compare the results in Tables 4 and 5, the proportion of /+interrogative/ and non-interrogative intonational patterns is roughly identical in the two corpora, but the proportion of the /+interrogative/ patterns in the textbook corpus is higher (74% with respect to 62%). It is positive as we consider that didactic materials should concentrate more on familiarizing stu-

dents with the /+interrogative/ patterns first, thus, they should appear in higher proportion.

Table 3: Examples of the textbook corpus

Code	Spanish sentences and translation
PATTERN 2 /+INTERROGATIVE/	
v9g	¿Verdad que coincide? 'It coincides, is that true?'
v25c	¿El tiempo te gusta? 'Do you like the weather?'
PATTERN 3 /+INTERROGATIVE/	
v29b	¿Estás ahí? 'Are you here?'
v39e	¿Puede ser? 'Can it be?'
PATTERN 4 /+INTERROGATIVE/	
v9k	¿Verdad? 'True?'
v27e	¿Y te gusta? 'And do you like it?'
PATTERN 13 /+INTERROGATIVE/	
v15d	¿Que no me altere? 'That I shouldn't lose my temper?'
/-INTERROGATIVE/ PATTERNS	
v15b	¿Que diga algo? 'That I should say something?'
v15g	¿No le parece? 'Don't you think so?'

Table 4: Patterns in the spontaneous corpus

Pattern	2	3	4	13	Contour /+Interrog./	Contour /-Interrog./
Number (out of 80)	10	17	6	17	50	30
Frequency (%)	12.5%	21.2%	8%	21.2%	62%	38%

Table 5: Patterns in the textbook corpus

Pattern	2	3	4	13	Contour /+Interrog./	Contour /-Interrog./
Number (out of 70)	41	5	5	1	52	18
Frequency (%)	59%	7%	7%	1%	74%	26%

Now we will deal with the two intonational patterns used for asking neutral yes-no questions, SP2 and SP3. From a pragmatic point of view, they are used to obtain or confirm information, to ask for something or to make a rhetoric question (Font-Rotchés–Mateo, in press). These are the basic patterns that should be learnt first.

Spanish Pattern 2, /+interrogative –emphatic/, in the textbook corpus is predominant, with 41 questions (59%). That is, even if it is obvious that in

spontaneous speech the presence of this pattern is much lower (only 10 questions, a12.5%), the authors of the textbook found it necessary to follow the traditional view that the "normative" yes-no question intonation is characterized by a considerably high final inflection. The contours of this pattern have a rise of minimally 70% (100% is equivalent to an octave of the musical scale). The different methods and authors described the intonation of Spanish absolute yes-no questions with a final rise, but without giving the exact quantity of the final rise. However, Cantero et al. (2002) have demonstrated based on perceptive tests (following the MAS method) that in this pattern the final rise should be at least 70% so that the contour could be recognized as /+interrogative/ by listeners.

If we have a look at Table 6, we can observe that the textbook does not only offer more examples of SP2 than the spontaneous corpus, but it presents, in three thirds of the cases (75.6%) a Final Inflection with a final rise much higher – between 91% and 201% – than in spontaneous speech, in which the maximum amount of rise is 91% (as examples, cf. Figures 5 and 6). The average value of the rise is much higher in the case of the textbook corpus: 114.8% as opposed to the spontaneous average 81.1%.

In the diagrams, syllables between parentheses mean that they are not realized audibly, whereas syllables marked with an asterisk (*) mean that they are lengthened.

In Figure 5 we have an authentic spontaneous speech contour with a final rise of 70.6%, which characterizes SP2, whereas in Figure 6 we have a contour from the textbook corpus which ends in a very high rise, of 177%.

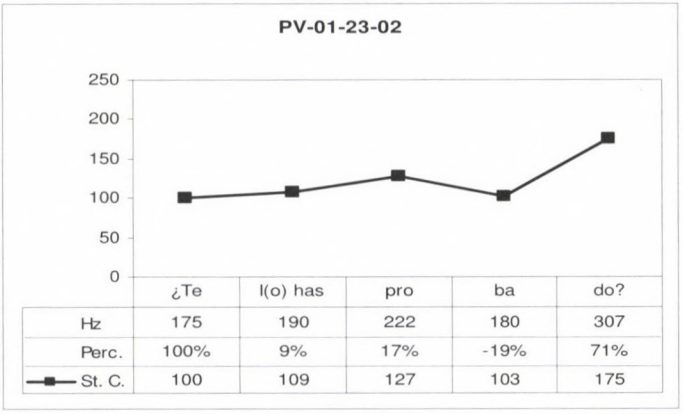
Table 6: Final rises in SP2 contours

Pattern	Rise	Frequency in spontaneous speech		Frequency in the textbook	
2	70%-91%	10	100%	10	24.3%
	91%-201%	–	–	31	75.6%

	Spontaneous speech corpus	Textbook corpus
Rise (%, minimum value)	70	70
Rise (%, maximum value)	91	201
Average rise (%)	81.1	114.8
Standard deviation (%)	6.8	35.2

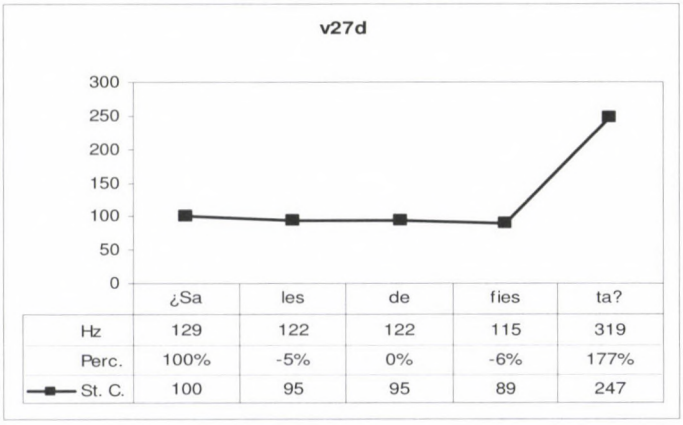
The reason for so high final rises can lie in the fact that the text is read and is thought to be directed to learners of Spanish. This is why different speakers make special efforts to utter interrogatives, which is highly beneficial to learn-

ers who listen: the higher the final rises, the easier students identify and reproduce them. According to Markó (2012), however, this phenomenon in itself is explained by the read nature of the corpus as opposed to spontaneous speech.



	¿Te	l(o) has	pro	ba	do?
Hz	175	190	222	180	307
Perc.	100%	9%	17%	-19%	71%
St. C.	100	109	127	103	175

Figure 5.
A contour of SP2 of spontaneous speech: *¿Te lo has probado?* 'Have you tried it?'



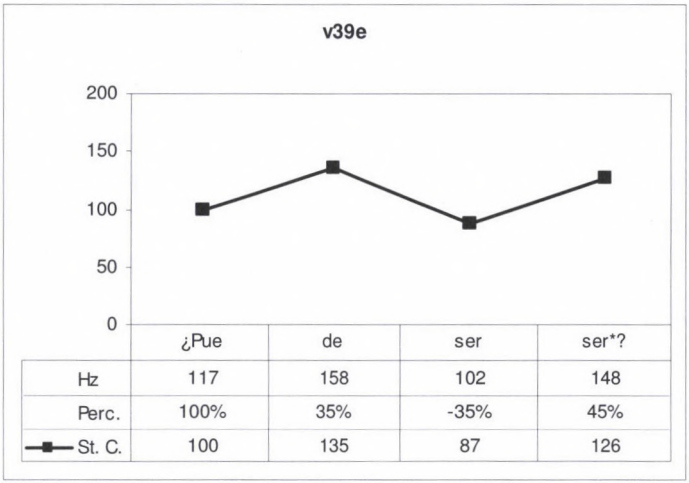
	¿Sa	les	de	fies	ta?
Hz	129	122	122	115	319
Perc.	100%	-5%	0%	-6%	177%
St. C.	100	95	95	89	247

Figure 6.
Contour of SP2 of the textbook: *¿Sales de fiesta?* 'Do you go out to have a party?'

Spanish Pattern 3 is more often used in spontaneous speech than in the textbook, in 21.2% (17 questions) with respect to 7% (5 questions). In spon-

taneous speech it is more abundant as native speakers know that in certain contexts in which the questions can be anticipated without having to make so huge efforts, by a pattern with a smaller final rise (between 40% and 70%) they can be interpreted when they make questions. In the textbook, the objective is just the opposite: it should be absolutely clear that we are facing a question. This is why we have a lot more SP2 contours (59%) as compared to the spontaneous corpus, with 7%.

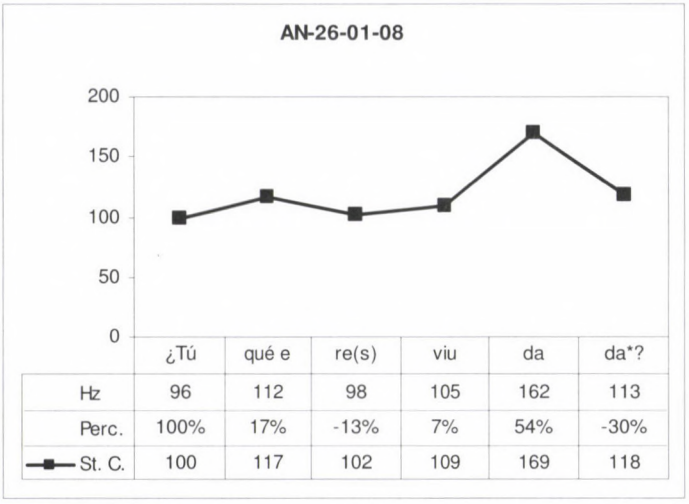
SP3 contours must have, apart from a rising FI (with 40%-70%), another important melodic trait: the First Peak must be shifted to one of the following unstressed syllables, as can be seen in Figure 7. The example in Figure 7 comes from the textbook, and it is characterized by the First Peak shifted to the unstressed syllable *–de,* after the first stressed syllable *pue-,* and a final rise of 45%. In our view, as this pattern is frequently used by native speakers, there should be more examples in the textbook so that language learners could get familiarized with it.



	¿Pue	de	ser	ser*?
Hz	117	158	102	148
Perc.	100%	35%	-35%	45%
St. C.	100	135	87	126

Figure 7.
Contour of SP3 of the textbook *¿Puede ser?* 'Can it be?'

Besides the neutral /+interrogative/ patterns, there exist at least two more /+interrogative/ patterns in Spanish, which are also /+emphatic/. These are SP4 (with a rising-falling FI), which only appears in polite questions, and SP13, with rising Body and FI, which appears in exclamatory questions and offers (Font-Rotchés–Mateo, in press).



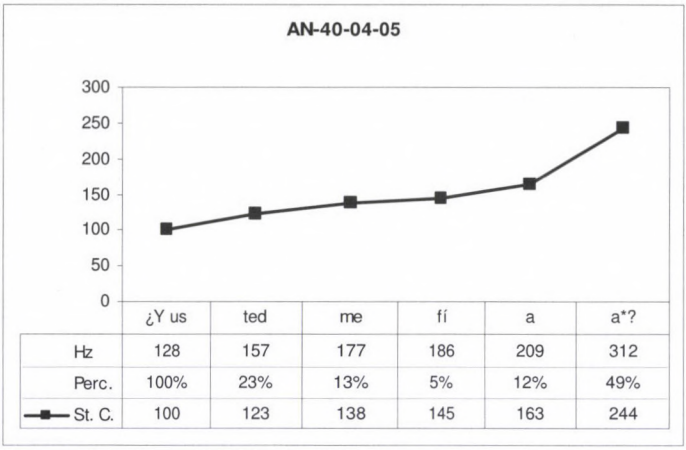
	¿Tú	qué e	re(s)	viu	da	da*?
Hz	96	112	98	105	162	113
Perc.	100%	17%	-13%	7%	54%	-30%
St. C.	100	117	102	109	169	118

Figure 8.
Contour of SP4 of spontaneous speech. *¿Tú qué eres, viuda?* 'What are you, a widow?'

SP4 is rare in both corpora (8% and 7%, respectively). The contours which are characterized by this pattern need special contexts. It is important to dominate this melody because it permits to mitigate the content of the question by the help of intonation, as in the case of the question in Figure 8 *¿Tú qué eres, viuda?* 'What are you, a widow?. Another possibility of mitigating the content could be reached by syntactic mechanisms, such as in *¿Me podría decir si es viuda?* 'Could you tell me if you are widow?'. In our view, in higher levels of language teaching (C1 and C2), this pattern should be explained and taught. In this textbook there are only 5 contours like this which appear in dialogues.

Apart from this low number of occurrences, they are characterized by having the length of only one word *¿Ok?* (twice),*¿verdad?* 'True?', *¿Sí?* 'Yes?', and a longer one *¿Y te gusta?* 'And do you like it?'. As we see, C1-level learners would hardly learn this pattern based on so few and short examples. This is worrying because SP4 has important pragmatic functions that would help learners to reach a better communicative competence.

Finally, SP13, which is the second most frequent pattern in spontaneous speech with 21.2% – is scarcely represented in the textbook corpus (only one example, 1%). In Figure 9, we can see an example for SP13 with a Body and a FI rising, and the total rise is of 144%.



	¿Y us	ted	me	fí	a	a*?
Hz	128	157	177	186	209	312
Perc.	100%	23%	13%	5%	12%	49%
St. C.	100	123	138	145	163	244

Figure 9.
Contour of SP13 of spontaneous speech. *¿Y usted me fía?* 'And do you confide in me?'

In this case, examples of this pattern are practically inexistent, although it would be very useful to language learners, as it is used for expressing emotions in exclamatory questions by native speakers. The contours of this pattern are perceived as emphatic questions (Font-Rotchés–Mateo, in press) if the global rise is over 90% (the Body and the FI taken together).

Conclusions

We have compared two corpora – a spontaneous and a textbook one – from an intonational point of view. We have examined whether in the textbook corpus (taken from *El Ventilador*) yes-no questions present a variety of intonational patterns as wide as in the spontaneous corpus.

We have analysed the intonation of the utterances with the same method: the Melodic Analysis of Speech by Font-Rotchés and Cantero (2009). According to this theory, the utterances are analysed with the help of a sound analysis and synthesis software (Praat), in a way that every syllable is given one (or maximally three, if it constitutes a complex inflection) characteristic F_0 value, in order to get rid of insignificant micromelodic variations. After that, the melodies are standardized, so as to eliminate the individual factors of the speaker's voice.

Based on recent corpus investigations, we illustrated the /+interrogative/ Spanish intonational patterns, including the neutral and the emphatic ones. The presence of these patterns was examined in *El Ventilador*.

As a conclusion, we can see that *El ventilador* presents almost the same proportion between /+interrogative/ and /-interrogative/ patterns in yes-no questions as the spontaneous corpus does. As *El ventilador* is a didactic material, it gives more emphasis and thus more presence to the /+interrogative/ patterns, as these are the ones which should be taught first.

As for the dispersion of the /+interrogative/ patterns, *El ventilador* and the spontaneous corpus show considerable differences.

Neutral yes-no question intonation patterns (SP2 and SP3) are used to ask for information or for anything else, or to confirm data. It is important to note that the textbook presents a high occurrence of SP2 (characterized by a Final Inflection with a rise up to 70%) and the contours have a very high rise (most of them overcome 100%), whereas SP3, with a less considerable rise (40%~70%), has less presence if we compare with spontaneous speech. This fact could be explained by the effort speakers who read the dialogues made when producing the text in order to mark neutral questions with a considerable rise. As a consequence of that, other melodies with less rise were disregarded because they would have less guarantee to be decoded as questions, depending on the context.

As far as the emphatic interrogative patterns are concerned, which are used for expressing politeness (SP4) or for exclamatory questions (SP13), these patterns should be learnt at higher levels of language learning. Still, this textbook offers only a small number of examples, which are, moreover, not long enough. The presence of SP13, which is considerable in the spontaneous corpus, is almost inexistent in *El ventilador*.

To sum up, the textbook analysed, which is a widely-taught method, doesn't reflect the linguistic reality that characterizes Spanish yes-no question intonation and only offers a reduced panorama of intonational phenomena related to yes-no interrogatives. In future, textbooks should treat intonational aspects as well in order to help Spanish learners to improve their communicative competence.

References

Alcoba, Santiago – Murillo, Julio 1998. Intonation in Spanish. In: Hirst, Daniel – Di Cristo, Albert (eds.) 1998. *Intonation Systems. A survey of twenty languages*. Oxford University Press, Oxford. 152–166.

Baditzné Pálvölgyi, Kata in press. *Spanish intonation of Hungarian learners of Spanish: yes-or-no questions*. PhD thesis. 2012. ELTE, Budapest, Biblioteca Phonica. http://www.publicacions.ub.edu/revistes/ phonica-biblioteca/.

Ballesteros, Maria Pilar 2011. *La entonación del español del norte*. PhD Thesis. Laboratori de Fonètica Aplicada. Universitat de Barcelona.

Canellada, María Josefa – Madsen, John Kuhlmann 1987. *Pronunciación del español*. Castalia, Madrid.

Cantero Serena, Francisco José 2002. *Teoría y análisis de la entonación*. Universitat de Barcelona, Barcelona.

Cantero, Francisco José – Mª De Araújo, Amália – Yen-Hui Liu et al. 2002. Patrones melódicos de la entonación interrogativa del español en habla espontánea. In: *Actas del II Congreso de Fonética Experimental*. Universidad de Sevilla, Sevilla. 118–123.

Cantero Serena, Francisco José – Font-Rotchés, Dolors 2007. Entonación del español peninsular en habla espontánea: patrones melódicos y márgenes de dispersión. *Moenia* 13. 69–92.

Cantero Serena, Francisco José – Font-Rotchés, Dolors 2009. Protocolo para el análisis melódico del habla. *Estudios de Fonética Experimental* 18. 17–32. http://stel.ub.edu/labfon/ca/estudios-de-fonetica-experimental-xviii-2009 (last consulted 28/08/2012).

Chamorro, María Dolores – Lozano, Gracia – Ríos, Aurelio – Rosales, Francisco – Ruiz, José Plácido – Guadalupe, Ruiz 2006. *El ventilador*. Difusión, Barcelona. Transcripciones: http://www.difusion.com/ele/ideas-y-recursos/ (last consulted 20/08/2012)

Congosto, Yolanda – Fabián Martín-Arroyo, María Dolores – Fernández Barranco, Concepción 2008. Picos tonales, acentos y límites sintagmáticos en enunciados declarativos e interrogativos absolutos sin expansión vs. con expansión (en el sujeto y en el objeto). *Language Design. Journal of Theoretical and Experimental Linguistics. Special Issue 2: Experimental Prosody*. 203–212. http://elies.rediris.es/Language_Design/LD-SI-2/23-Congosto-Martin-Barranco-BREVE.pdf (last consulted 20/08/2012).

Cunningham, Una 1983. Aspects of the Intonation of Spanish. *Nottingham Linguistic Circular* 12. 21–54.

Devís, E. 2011. La entonación del español hablado por italianos. *Didáctica (Lengua y Literatura)* 23. 35–58.

Díaz Tejera, Alberto 1973. La frase interrogativa como modalidad. *Revista Española de Lingüística* 3, 95–116. http://dialnet.unirioja.es/servlet/articulo?codigo=409 26 (last consulted 20/08/2012).

Face, Timothy 2007. The role of intonational cues in the perception of declaratives and absolute interrogatives in Castilian Spanish. *Estudios de Fonética Experimental* 16. 185–226. http://stel.ub.edu/labfon/ca/estudios-de-fonetica-experimental-xvi-2007 (last consulted 20/08/2012).

Fernández Planas, Ana María – Martínez Celdrán, Eugenio 2003. El tono fundamental y la duración: dos aspectos de la taxonomía prosódica en dos modalidades del habla (enunciativa e interrogativa) del español. *Estudios de Fonética Experimental* 12. 165–200. http://stel.ub.edu/labfon/ca/estudios-de-fonetica-experimental-xii-2003 (last consulted 20/08/2012).

Fonseca, A. y Cantero, F. J. 2011. Características da entonação do espanhol falado por brasileiros. *Actas do VII congresso internacional da ABRALIN*. Curitiba (Brasil), 84–98.

Font-Rotchés, Dolors 2007. *L'entonació del català*. Publicacions de l'Abadia de Montserrat, Barcelona.

Font-Rotchés, Dolors – Francisco José Cantero 2009. Melodic analysis of speech method applied to Spanish and Catalan. *Phonica* 5. 33–47. http://www.publicacions.ub.edu/revistes/phonica5/Default.asp (last consulted 9/08/2012).

Font-Rotchés, Dolors – Mateo, Miguel 2011. Absolute interrogatives in Spanish: a new melodic pattern. *Actas do VII congresso internacional da ABRALIN*. Curitiba (Brasil). 1111–1125.

Font-Rotchés, Dolors – Mateo, Miguel (in press). Entonación de las interrogativas absolutas del español meridional en habla espontánea.

Henriksen, Nicolas 2010. Nuclear rises and final rises in Manchego Peninsular Spanish yes/no questions. In: *Speech prosody 2010. Fifth international conference on speech prosody*. Chicago, Illinois, USA. May 11-14, 2010. http://aune.lpl.univ-aix.fr/~sprosig/sp2010/papers/100212.pdf (last consulted 20/08/2012).

Jiménez Gómez, Juan Julián 2010. Claves acústicas en la distinción fonológica interrogativa-declarativa en español: La estructura formántica de las vocales. *Estudios de Fonética Experimental* 19. 293–322. http://stel.ub.edu/labfon/ca/estudios-de-fonetica-experimental-xix-2010 (last consulted 20/08/2012).

Liu, Yen Hui 2005. *La entonación del español hablado por taiwaneses*. Biblioteca Phonica, 2. www.ub.es/lfa. (last consulted 20/08/2012).

Markó, Alexandra 2012. A magyar hangsúly realizációinak és észlelésének összefüggése felolvasásban és spontán beszédben. In: Markó Alexandra (szerk): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngekezdési időig*. ELTE BTK–MTA Nyelvtudományi Intézet 277–303.

Martínez Celdrán, Eugenio 2003. Análisis por niveles: la Escuela Americana. In Prieto, Pilar (ed.): *Teorías de la entonación*. Ariel Lingüística, Barcelona. 64–95.

Martorell, Laura 2011. *Les interrogatives absolutes de l'espanyol parlat per suecs*. (unpublished document). Departament de Didàctica de la Llengua i la Literatura, Universitat de Barcelona.

Mateo, Miguel (in press): *La entonación del español meridional*. PhD thesis. Dep. Didàctica de la Llengua i la Literatura, Universitat de Barcelona.

Navarro Tomás, Tomás 1966. *Manual de entonación española*. Edición Revolucionaria, La Habana.

Quilis, Antonio 1995. *Tratado de fonología y fonética españolas*. Biblioteca Románica Hispánica. Editorial Gredos, Madrid.

Sosa, Juan Manuel 1999. Entonación española. Su estructura fónica, variabilidad y dialectología. Cátedra, Madrid.

Zamora, Francisco José – Carrera, Micaela – Meléndez, María Mar 2005. Variación geoprosódica en dos entonaciones de Castilla y León: Análisis de frases declarativas e interrogativas sin expansión. *Estudios de Fonética Experimental* 14. 125–140. http://stel.ub.edu/labfon/es/estudios-de-fonetica-experimental-xiv-2005 (last consulted 20/08/2012).

EXPLOZÍVÁK ÉS AFFRIKÁTÁK IDŐVISZONYAI

Gráczi Tekla Etelka

Bevezetés

A beszédhangok képzése során az egyes artikulációs gesztusok egymással összeköttetésben, egymásra hatva működnek. Így például nemcsak a hangszalagok felől érkező, azok működése által meghatározott légáramot módosítja a szupraglottális artikuláció, hanem a beszédképző szervek az összeköttetések révén hatnak egymás beállításaira, illetve a hangszalagok feletti beszédszervek működésének okozta aerodinamikai jellemzők visszahatnak a hangszalagok működésére is.

Modális fonáció során a hangszalagokat összezárjuk, a tüdőből kiáramló levegő pedig felgyülemlik alattuk. Amikor a felgyűlő levegő nyomása elég nagy, hogy legyőzze a hangszalagok ellenállását, felnyitja azokat. Ez azonban csak akkor következhet be, ha a szupraglottális nyomás alacsonyabb, mint a hangszalagok felnyitásához szükséges nyomás (van den Berg 1958 alapján az eltérésnek legalább 3 H_2O cm-nek kell lennie – idézi Stevens 1998). Az obstruensek képzésekor az akadály fenntartása alatt az intraorális nyomás emelkedésével a hangszalagok feletti és – nyitott glottisz esetében – közötti nyomás is növekszik. Ezáltal a rezgéshez szükséges nyomáskülönbséghez közelíthet, vagy az alá csökkenhet az eltérés, így a hangszalagok rezgése megnehezedik, a nyitott szakasz időtartama növekszik (pl. Bickley–Stevens 1986), illetve le is állíthatja a zöngeképzést.

Az explozívák és az affrikáták képzésekor a szájüregben zárat hozunk létre. A zár megképzésétől a szájüregben a glottisz felől áramló levegő felgyülemlik, így az intraorális, illetve szupraglottális nyomás növekedni kezd, amely a szájüreg falát kifelé, a gégét pedig lefelé nyomja; a hangszalagok között megnövekvő nyomás pedig szétnyomja a hangszalagokat. Az intraorális nyomás növekedésével a szájüreg és a pharynx fala passzívan nagyobb térfogat felé nyomódik nyugalmi pozíciójából (Ohala 1990). Ezt a folyamatot a szövetek, izmok lazításával segíthetjük elő. Az affrikáták zárképzése nem tökéletes (Szende 1974: ha a nyelvcsatorna mégis zárat képez, az „párnahatás" [ti. Bernoulli] eredménye), emiatt felpattanások, réses áramlás is megjelenhet.

Az explozívák zárfeloldásakor a szájüregben megemelkedett nyomás hat a „szándékolt" artikulációs gesztusra, azaz meggyorsítja a zár felnyitódását. Az explozívák esetében a zár felnyílásakor hirtelen felerősödik a légáram, így a nyomás ugyancsak hirtelen csökken. A (korábbi) zár területén annak felnyílása után lecsökken a nyomás, ami miatt visszazáródhat a zár, így többszöri

felpattanást eredményezhet (magyarra vö. Gráczi–Kohári 2012). Ez a hatás a veláris explozívák esetében jelentősebb az érintkező területek nagyobb mérete miatt (Stevens 1998; Keating et al. 1983). A zár felnyílásakor először egy tranziens szakasz jön létre, majd ahogy az intraorális nyomás csökkenésével a szájfal nyugalmi állapotába visszahelyezkedik, és megindul a glottisz felőli légáram, turbulens rezgés jön létre, amelyet hehezet, levegős zöngeképzés követhet, amennyiben a hangszalagrezgés leállt (Stevens 1992). Mindezen jellemzők időtartamának aránya függ a felnyílás sebességétől, amely a képzési helynek is függvénye a dorzum lassabb mozgása miatt.

Az affrikáták képzése során a zár elülső, kis része hirtelen nyílik fel – az explozívákhoz hasonlóan – azonban a hátsóbb, nagyobb felületű akadályterület csak ezután, résesen nyílik (Stevens 1998). Az elülső akadályrész felnyitását a zár mögött megnövekedett nyomás (ugyancsak az explozívák feloldásához hasonlóan) gyorsítja, a hirtelen nyomáscsökkenés pedig a hátsóbb felület átmeneti feljebb emelkedését okozza (Stevens 1998).

Az explozívák zárfeloldása során a nyomás fokozatosan csökken, míg az affrikáták esetében a résbe átnyíló képzés a spiránsokhoz hasonló hatást feltételez, azaz a résképzés, a spirantikus zörej keltésének és fenntartásának artikulációs szándéka tovább nehezíti a zönge fenntartását. Az intenzív spirantikus zörej eléréséhez nagyobb akadály mögötti nyomás szükséges, amelyet a zöngétlen képzés (nyitott glottisz) során érhetünk el. Ha a szájüregi rés vagy a glottisz területén megnő a keresztmetszet, csökken az adott gerjesztés keltette rezgés amplitúdója (spiránsokra: Stevens et al. 1992). A képzés végén, amikor a szájüregbeli akadály nyitódása megindul, ismét növekszik a légáram, aminek következtében csökken az intraorális, a hangszalagok közötti és a transzglottális nyomás. A réses akadály és a glottisz keresztmetszete azonban nem állandó, hiszen a növekedő intraorális és hangszalagok közötti nyomás mindkettőre erőt fejt ki (spiránsokra: Stevens 1997).

A szupraglottális üregrendszer aktív és passzív (nyomásnövekedés hatására történő) tágításával (kb. 20%-a a nyugalmi térfogatnak) csökkenthető a glottisz feletti nyomás növekedése. Ez elérhető a gége süllyesztésével, a pharynx tágításával, a lágyszájpad emelésével, a nyelvgyök és az epiglottisz előretolásával (Perkell 1969; Stevens 1998; Maddieson 1997; Ohala 1997; Shadle 1997). A hangszalagok lazítása és az aktív szupraglottálisüreg-tágítás során a nyelv felszínéhez közeli izmok is lazulnak, további passzív tágulást okozva (Svirsky et al. 1997). A zöngeképzés lehetőségét fenntartó artikulációs gesztus lehet még a nazális üreg megnyitása, vagyis a zárszakasz egy részében megjelenő nazalitás, amely a nyomás lassabb növekedését szolgálja; esetleg a légáram „szökésének" biztosítása gyenge frikció révén (explozívákra: pl. Ohala 1997; Solé 2011; Solé–Sprouse 2011). További kompenzáció eredménye a zöngés obstruensek képzése közbeni zöngefenntartás nehézségének következtében, hogy a zöngétlen párhoz viszonyítva jellemzően rövidebb időtartamban valósulnak meg (Maddieson 1997; Ohala 1997). Néhány ejtésben

pedig megváltozhat a képzési mód. Magyarra például Magdics 1969-ben a /z/ ejtésében adatolt approximáns jellegű realizációt.

A zöngeképzésre számos tényező van hatással. Intervokális helyzetben a hangszalagok beállítása és a nyomásviszonyok miatt könnyebb fenntartani a zöngét, mint szünet utáni vagy előtti helyzetben (Westburry–Keating 1986). A mássalhangzó képzési helyéhez közeli, felső nyelvállású magánhangzók esetében tágabb a pharynx, ami a lassabb nyomásnövekedést segíti elő (Ohala–Riordan 1979; Pape et al. 2006). Magasabb nyelvállásfokú magánhangzó környezetében ugyanakkor megnyúlik az explozívák zöngekezdési ideje, illetve a zörejes elem időtartama is, mivel a nyelv kisebb süllyesztésével csak lassabban távozik a szájüregből a levegő, így lassabban csökken az intraorális nyomás, és emelkedik a transzglottális nyomáskülönbség (pl. Stevens 1971; Ohala 1976, 1983). A mássalhangzó hátsóbb képzéshelye esetében kisebb térfogat, kisebb tágítható felület és nagyobb zárkiterjedés (nagyobb ellenállás, lomhább mozgás) nehezíti a zöngeképzés fenntartását (pl. Shadle 1997; Stevens 1998). A képzési hely hatásának következménye például a magyar palatális zárhangok képzésmódjának kérdése. Számos elemzés sorolta az explozívák, mások az affrikáták közé (vö. Kázmér 1961; Szende 1997: 111), azonban fonológiai szempontból, valamint a képzési hely sajátosságai alapján általánosan explozívának fogadható el (Ladefoged 1941; Cseresnyési 1992; Siptár 1995).

Mindezek mellett további hatások és egyéni beszélői sajátosságok (adott fizikai paraméterek, pl. a szájüreg térfogata; artikulációs jellemzők – van Alphen–Smith 2004) is befolyásol(hat)ják a megvalósuló beszédhang szerkezetét.

Az eddigi, magyar nyelvre készült vizsgálatok közül kevés érinti magát a zöngésség időviszonyait, a mássalhangzó időtartamával azonban számos kutatás foglalkozott. Gósy és Ringen (2009) olvasott szólistán alapuló elemzésében kimutatta, hogy a képzési hely hatással van a zöngekezdési időre szó belseji és végi helyzetben a zöngés és a zöngétlen felpattanó zárhangok esetében is. Ezt a tendenciát spontán beszédben is igazolták szó belseji helyzetben (Gráczi et al. 2009). A mássalhangzó teljes időtartama alatti zönge aránya, a zöngekezdési idő és az időtartam alapján a képzési hely függvényében változó gyakoriságú volt az oppozíció (részleges) neutralizációja (Gráczi 2011a, b). Igazolták magyar nyelvben is a magánhangzó-környezet és a képzési hely együttes hatását, azaz a „képzési helyek" távolságának szerepét a zöngekezdési idő alakulásában (Gósy 2000). A beszélők közötti variabilitást az idősek és a fiatalok ejtésében megjelenő VOT-értékek eltérésében igazolták (Bóna 2011). Nemcsak a két csoport azonos beszédmódjában (olvasás és spontán beszéd) találtak eltérést, hanem az idősek esetében a két beszédmód mutatta zöngekezdésiidő-mintázat nem tért el, míg a fiatalokéban igen. Azaz az életkorral nem csak az artikuláció, hanem a beszédstílus okozta hatás is változott.

A mássalhangzók időtartamának elemzése minden esetben kimutatta, hogy a zöngés fonémák megvalósulásai rövidebbek a zöngétleneknél. Az eltérés mértékére azonban számos tényező volt hatással: pl. a képzési mód (explozíva, affrikáta, spiráns), a képzési hely, a hangsorbeli helyzet és a hangsúlyosság, a hangsor hossza, a szomszédos magánhangzók képzési jegyei (pl. Gombocz–Meyer 1909; Magdics 1969; Kassai 1982; Olaszy 2006).

A jelen vizsgálatban azt a kérdést tettük fel, hogy az egyes zárhangpárok belső időtartamviszonyai hogyan alakulnak a zöngésség, a képzési hely és a kontextus függvényében, illetve milyen jellemzők hasonlóak és különbözőek a két képzési módban. Feltételeztük, hogy az explozívák zöngésségi oppozíciója csak ritkán neutralizálódik, amire a képzési hely, a kontextus és a beszélő ejtése is hatással lesz. Az affrikáták esetében a réses elem miatt gyakoribb részleges zöngétlenedést feltételeztünk, valamint azt vártuk, hogy a képzési helyek közelebbi mivolta miatt ez a jegy nem befolyásolja lényegesen az egyes vizsgált paraméterek alakulását. A fonetikai kontextus esetében feltételeztük, hogy nemcsak a magánhangzó saját, hanem a mássalhangzóéihoz viszonyított képzési jegyei is befolyásolják, hogy milyen mértékben hat a konszonánsok megvalósulásaira.

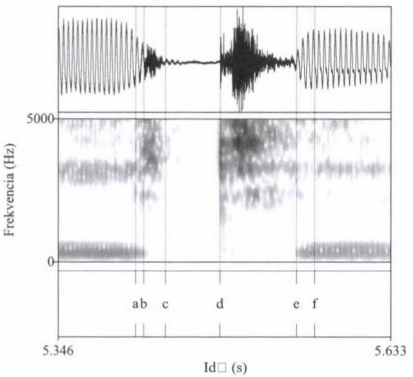
Kísérleti személyek, anyag és módszer

A kérdések megválaszolásához laboratóriumi beszéd elemzését választottuk. A vizsgálat anyagához olyan CVCVC logatomokat állítottunk össze, ahol az első és az utolsó mássalhangzó azonos (minden esetben /l/), a két magánhangzó azonos, /ɔ/, /i/ vagy /u/, a vizsgált konszonáns pedig középen, tehát azonos magánhangzók közötti intervokális helyzetben szerepelt. Minden, a magyar nyelvben meglévő explozíva és affrikáta szerepelt a vizsgálatban, tehát a /b, p, d, t, ɟ, c, g, k, ʣ, ʦ, ʤ, ʧ/ fonémák megvalósulásait elemeztük. A logatomokat (pl. *labal, libil, lubul, lapal, lipil, lupul*) az „A képernyőn a *l*VCV*l* alak látható." mondatba ágyaztuk. Minden mondatot összesen négyszer olvastattunk fel az adatközlőkkel, randomizált sorrendben a SpeechRecorder szoftver (Draxler–Jänsch 2004) segítségével. A felvételeket AT 4040 mikrofonnal, 40 kHz-en, 16 bites mintavételezéssel rögzítettük egy csendesített helységben.

A mondatokat négy női és két férfi adatközlő felolvasásában rögzítettük. Mindannyian a budapesti köznyelvet beszélő, ép hallású és beszédképzésű 20–29 éves beszélők voltak.

A hanganyagot a Praat 5.1–5.3 szoftver használatával címkéztük és elemeztük (Boersma–Weenink 2011). A jelen tanulmányban ismertetendő adatokhoz a mássalhangzó kezdetét és végét, a belső időszerkezet viszonyait jelöltük (1. ábra). Címkéztük a megelőző magánhangzó második formánsának lecsengését (a), a mássalhangzó zárszakaszának kezdetét (c), a zárfeloldás kezdetét (d), a követő magánhangzó második formánsának indulását (f), illetve a mássalhangzó képzése alatt a zönge leállásának (b) és újbóli megindu-

lásának esetleges időpontját (e). Adatoltuk, hogy jelent-e meg 1. a magánhangzó-zár átmenet során előaspiráció vagy záródási zörej (felpattanás, réses elem), 2. a zárszakasz alatt átmeneti zárfelnyílás vagy 3. réses áramlás. Néhány esetben a realizáció nem zárhang, hanem rés- vagy közelítőhang volt. Ezekben az esetekben is jelöltük a mássalhangzó kezdetét és végét a hangátmenet alapján.



1. ábra
Címképelda: a *li*[tʃ]*il* hangsor címkézése

Elemeztük a mássalhangzó realizációjának módját, az időtartamát (f–a), a zár (a zár kezdetétől a feloldás kezdetéig eltelt idő; d–c) és annak feloldásának (a feloldás kezdetétől a követő magánhangzó második formánsáig eltelt idő; f–d) időtartamát. Azokban a realizációkban, amelyek során leállt a zöngeképzés, elemeztük a zönge leállása és a zár kezdete között eltelt időt (b–c) és a zöngekezdési időt (e–d).

A statisztikai elemzéseket az SPSS 19.0 szoftver segítségével végeztük. Az adatok a Shapiro–Wilk-próba alapján több esetben nem tekinthetők normál eloszlásúnak, ezért nemparametrikus (Mann–Whitney és Kruskal–Wallis) próbát alkalmaztunk.

Eredmények

A megvalósult hang képzési módja

A képzési mód ritkán tért el az alaprealizációban elvárt képzési módtól. Az esetek többségében zárhangként valósultak meg az elemzett fonémák. Approximánsként vagy réshangként a zöngések közül a /d/, /ʤ/, /ɟ/ és /g/ jelent meg. Az alveolárisok 1-2 (1,4% és 2,9%), míg a palatális és a veláris 5 és 8 előfordulása (6,8% és 11,4%) jelent így meg. A zöngétlenek között a /c/ és a /ʧ/ (1,4% és 2,6%) jelent meg réshangként. Néhány realizáció esetében a képzés teljes időtartama alatt jelen volt gyenge réses áramlás, a szerkezet

azonban egyértelműen felbontható volt (szivárgó levegővel képzett) zárszakaszra és -feloldásra. Ez a /tʃ/, a /ʤ/ és a /ɟ/ esetében volt adatolható (2,6–5,7%).

A zár időtartama alatt néhány esetben megjelentek korábbi felpattanások is, azaz a zár alatt nem folyamatos levegőszivárgás történt, hanem a mögötte felgyülemlő levegő néhány esetben korábban, mint a zár feloldása felnyitotta egyszer vagy többször átmenetileg a zárat. Ezt elsősorban az affrikáták jellemzőjeként tartják számon, Szende szerint (1974) ez a nem tökéletes zárképzés eredménye ezekben a hangzókban. A zöngés explozívák zárszakasza alatt összesen háromszor (a /g/ és a /ɟ/ esetében) fordult elő. A zöngétlen explozíváknál gyakrabban jelent meg. A /t/ esetében ötször (négyszer az /i/, egyszer az /ɔ/ környezetében), a /k/ esetében háromszor (az /ɔ/ és az /i/ kontextusában), a /c/ esetében pedig 11-szer (ebből nyolcszor az /ɔ/-s környezetben) fordult elő. A zár-réshangok között gyakoribb volt a zár alatti átmeneti felnyílásoknak az előfordulása. A /ʤ/ esetében hatszor (ebből négyszer az /ɔ/-s környezetben), a /ʣ/-ében 2-3-szor mindhárom jelent meg mindhárom kontextusban. A zöngétlen affrikáták között az alveolárisnál ötször (három esetben az /i/-k között), a posztalveolárisnál pedig az /u/-s környezetben kétszer fordult elő. A hangkörnyezet kismértékben befolyásolta az előfordulás gyakoriságát azokban a fonémamegvalósulásokban, amelyekben gyakoribb volt a jelenség. A /t/ és a /ʦ/ esetében az /i/-s, a /c/ és a /ʤ/ esetében az /ɔ/-s kontextus váltotta ki gyakrabban a zár alatti felnyílásokat. A /ʣ/ esetében mindhárom magánhangzónál hasonló értéket kaptunk. Elemeztük, hogy a jelenség tekinthető-e egyéni ejtésbeli jellegzetességnek, avagy egyértelműen csak magukat a konszonánsokat jellemzi. Két beszélő esetében egy-két fonéma realizációjában jelent meg, három kísérleti személy esetében négy fonémát érintett a jelenség, míg az f2-es adatközlőnél nyolc fonéma realizációjánál is adatolhattuk a jelenséget. Az összes zárhangként történt ejtéshez viszonyított arány is ennek megfelelően alakult. Míg a többi beszélőnél 0,7–3,5%-ban, ennél az adatközlőnél 15,3%-ban fordult elő a jelenség.

A zár feloldása néhány esetben egyáltalán nem adatolható a regisztrátumon (sem felpattanás, sem rés nem jelenik meg). A /b/ esetében 2,8%, a /ɟ/-ében 1,4%, a /g/-ében pedig 6,8% volt ennek az előfordulása, a zöngétlenek között a /c/ esetében 1,4%, a /tʃ/-ében pedig 2,6% volt. Más esetekben előfordult többszöri felpattanás, illetve egy esetben kétszeri réses felnyílás. Az elemzések során nem csak a várt módú felnyílásokat (explozívák esetében egy vagy több felpattanás, affrikáták esetében résbe való átmenet – esetleg felpattanással; illetve a palatálisok esetében mindkettő) vettük figyelembe, hanem a feloldás ezektől eltérő módját is. Tehát például a felpattanó zárhangok esetén a ritkább réses nyitódást.

A mássalhangzó időtartama

A zöngés explozívák időtartama két kiugró értéktől eltekintve 33–136 ms között volt. A képzési hellyel hátrafelé haladva 86±16 ms, 70±14 ms, 83±19

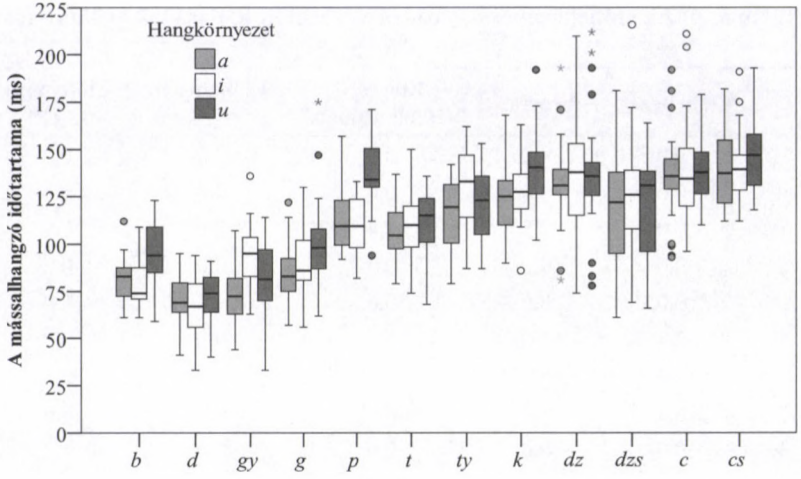
ms és 92±21 ms volt az átlagos időtartamuk. A /d/, /ɟ/ és /g/ esetében adatolt nem zárhang-realizációk jelentősen rövidebbek voltak, 31±4 ms, 55±17 ms és 77±11 ms.

A zöngétlen explozívák a vártaknak megfelelően minden esetben hosszabbak voltak a zöngés párjuknál, 68–171 ms között jelentek meg (ismét a kiugró értéket nem figyelembe véve). A képzési hellyel hátrafelé haladva 108±18 ms, 122±20 ms, 131±20 ms és 134±32 ms volt az átlagos időtartamuk.

Az affrikáták esetében a zöngések időtartama 61–216 ms, a zöngétleneké pedig 93–211 ms között alakult. Az előbbiek esetében ez 134±32 ms-os, illetve 122±30 ms-os, az utóbbiakéban pedig 138±24 ms-os és 142±19 ms-os átlagot jelent. (Az approximáns /ʤ/ időtartama 72 ms volt, azaz az explozívákhoz hasonlóan rövidebb, mint a zár-réshangként realizálódott társai.)

Az egyes magánhangzós környezeteket összevetve változóan alakult annak hatása az időtartamra (2. ábra). Az explozívák esetében jellemzően hatással volt arra, míg az affrikáták esetében nem. A /b/ és /p/ esetében az /u/-s környezetben jellemzően (18–19%-kal) hosszabb az átlagos időtartam, mint a másik két magánhangzó között mért értékek, amelyek közel azonosak. A /ɟ/ esetében az /ɔ/-k között mértük a legrövidebb, az /i/-k között a leghosszabb időtartamot, az eltérés 21%. A /g/ és /k/ esetében ismét az /ɔ/-s környezetben valósult meg a legrövidebb időtartamban az explozíva, míg az /u/-sban a leghosszabban, ez 10% (/k/) és 16% (/g/). Jellemzően a zöngétlen felpattanók esetében kisebb, a zöngésekében pedig nagyobb a hangkörnyezet alapján kapott különbség. A Kruskal–Wallis-próba alapján ezek a különbségek szignifikánsnak bizonyultak (a χ^2 7,341 és 29,477 között, a $p \leq 0{,}027$ alakult). A /d/, /t/ és a /c/ és az affrikáták esetében az időtartamok nem tekinthetőek eltérőnek a vizsgált hangkörnyezetekben.

Az egyes magánhangzó-környezetekben külön-külön vetettük össze, hogy a képzési hely és a zöngésség mentén eltérően alakultak-e a mássalhangzóidőtartamok. Az explozívák esetében mindhárom kontextusban a zöngés és a zöngétlen fonémák realizációiban is szignifikáns eltérést kaptunk az időtartamok között (Kruskal–Wallis-próba: a χ^2 12,559 és 28,876 között, a $p \leq 0{,}006$ alakult), míg az affrikáták esetében egyik magánhangzós környezetben sem tért el (Mann–Whitney-próba: a Z −1,697 és −0,866 között, a $p \geq 0{,}090$ alakult). Ez azonban nem jelenti azt, hogy a hátsóbb képzési hely hosszabb időtartammal járna. Ahogyan a 2. ábra is mutatja, a bilabiálisok viszonylag hosszúak voltak, az alveolárisoktól a velárisig ugyan megfigyelhető emelkedő tendencia, ami azonban a hangkörnyezet függvényében elmaradhat, mivel a palatálisok az /i/ környezetében jellemzően hosszabbak lehetnek a velárisoknál. Ennek magyarázata nyilván a magánhangzó és a mássalhangzó képzési helyeinek közelségében keresendő, azaz abban, hogy a zár megképzéséhez mennyi utat kell a nyelvnek megtenni.



2. ábra
A mássalhangzók teljes időtartama (ms)

A zöngésségi párokat összevetve a felpattanó zárhangok esetében minden képzési helyen mindhárom hangkörnyezetben szignifikáns eltérést kaptunk a mássalhangzó teljes időtartamában (Mann–Whitney-próba: a Z −5,468 és −5,184 között, a $p \leq 0{,}001$ alakult). Ez az eltérés 26–38% között mozgott. A zár-rés hangok esetében az alveolárisok egyik környezetben sem, a posztalveolárisok pedig a felső nyelvállásfokú magánhangzók között mutattak szignifikáns (11,1% és 17,9%) különbséget (Mann–Whitney-próba: a Z = −2,023; p = 0,020 és Z = −3,494; a $p \leq 0{,}001$ ebben a két esetben). Elemeztük, hogy az egyes beszélők ejtésében külön-külön vizsgálva az affrikátákat kimutatható-e valamilyen tendencia a zöngésségi párok esetében. Az alveolárisok esetében egy beszélőnél (n2) a zöngés jellemzőbben rövidebb időtartamú volt zöngétlen párjánál, de mindhárom hangkörnyezetben csak az ő ejtésében jelent meg ilyen tendencia. A posztalveolárisoknál az alsó nyelvállásfokú magánhangzónál ismét ennek a beszélőnek a beszédében található meg ez az eltérés, míg a másik két környezetben további három kísérleti személy ejtésében is felfedezhető ez a tendencia.

A mássalhangzó belső időszerkezete: szupraglottális jegyek

Elemeztük a mássalhangzók belső időszerkezetét is. A szupraglottális jegyek között elsősorban a zárszakasz és a –feloldás időtartamát vizsgáltuk, azonban a zárképzésbe való átmenet során néhány esetben átmeneti jellemzők is jelentek meg: **előhehezet,** illetve **záródási réses** vagy **„felpattanásos" zörej** (1. táblázat). Néhány esetben a lenyomat jellege miatt kérdéses az előhehezet vagy rés elkülönítése.

1. táblázat: Az előhehezet és a záródási zörej előfordulásának gyakorisága (%)

C	Előhehezet	Előrés	Előrés + "előfelpattanás"	Előhehezet vagy előrés	"Előfelpattanás(ok)"
/b/	0,0	0,0	0,0	0,0	0,0
/d/	0,0	0,0	0,0	0,0	1,4
/ɟ/	0,0	4,3	1,4	0,0	2,9
/g/	0,0	0,0	0,0	0,0	0,0
/p/	0,0	0,0	0,0	0,0	0,0
/t/	2,7	11,0	0,0	0,0	0,0
/c/	14,9	18,9	0,0	0,0	2,7
/k/	15,5	1,4	0,0	0,0	0,0
/dz/	0,0	0,0	0,0	0,0	1,4
/dʒ/	0,0	8,2	0,0	0,0	4,1
/ts/	4,1	29,7	0,0	1,4	1,4
/tʃ/	14,5	56,6	3,9	3,9	1,3

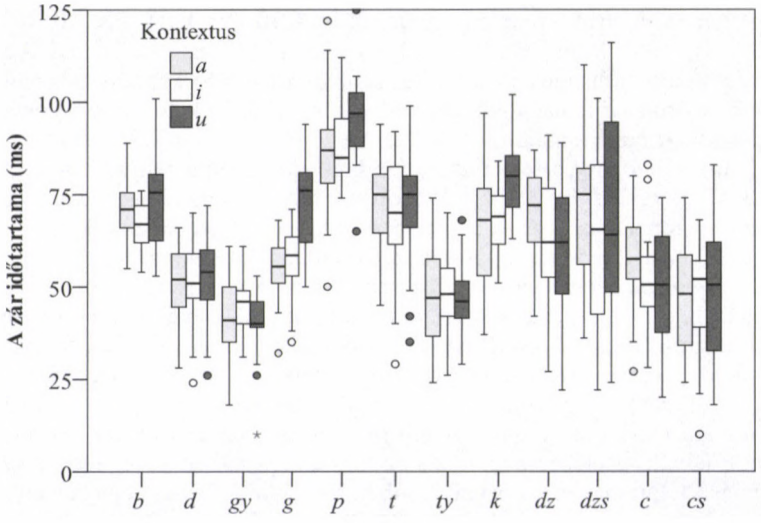
Előaspiráció leggyakrabban a hátsóbb képzési helyeken fordult elő mind az explozívák, mind az affrikáták esetében, és – ahogyan az várható volt – a zöngétlen fonémák megvalósulásaiban adatolhattuk. Záródási zörej a zöngéseknél is megjelent, és általában az affrikáták realizációiban gyakoribb volt. Érdemes megfigyelni, hogy a két palatális explozíva átmeneti jellegéből adódhatóan a felpattanók között ezeknél jelenik meg elsősorban, habár a /t/-nél is adatolhattunk néhány előfordulást. Minden beszélőnél előfordult legalább egy-két esetben előzörej vagy előhehezet, a /ts/, /tʃ/ és a /c/ mindenkinél érintett volt. Beszélőnként 3–8 fonéma összesen 15–50 realizációját érintette a két jelenség, amely beszélőnként az összes realizáció 8,8–36,0%-a volt. A két veláris magánhangzó kontextusában ejtett beszédhangok 18,0%-ánál és 20,9%-ánál, míg az /i/-ében 15,5%-uknál jelent meg. Az időtartamuk 3–49 ms között szóródott, az átlag19±9 ms volt. Az egyes mássalhangzók esetében 14–26 ms között alakult az előhehezet vagy a záródási zörej időtartama.

Mind a zöngés, mind a zöngétlen explozívák esetében jellemző volt, hogy a képzési hellyel hátrafelé haladva csökkent **a zárszakasz időtartama,** majd a veláris képzési helyűek realizációiban ismét az alveoláris és a bilabiális közötti átlagértéket vett fel. Ez a zöngéseknél a /b/-től a /g/-ig haladva 70±10 ms, 52±10 ms, 41±10 ms és 60±12 ms, a zöngétleneknél a /p/-től a /k/-ig haladva 90±13 ms, 72±15 ms, 47±12 ms és 70±13 ms volt.

Az affrikáták esetében a zöngések zárszakaszának átlagos időtartama közel azonos volt, a posztalveoláris valamivel nagyobb szóródásával: 65±16 ms és

67±24 ms. A zöngétlenek esetében a hátsóbb képzési helyen valamivel rövidebb volt az átlagos időtartam: 52±14 ms és 48±16 ms.

Elemeztük a mássalhangzó zárszakaszának időtartamát a kontextus függvényében is (3. ábra). A legtöbb konszonáns esetében nem mutatható ki az átlagértékekben eltérés, és a szóródás is csak ritkán változik a hangkörnyezettel nagyobb mértékben. Jelentős eltérést a felpattanó zárhangok esetében a két bilabiális és a két veláris hangzónál adatolhattunk. A két zöngés felpattanó zárhang realizációi az /u/ kontextusában 10,1%-kal és 21,2%-kal hosszabbak voltak, mint az /i/-ében, illetve az adatok szóródása is jellemzően nagyobb volt (az eltérés: 48,6% és 24,8%). A bilabiális esetében az /ɔ/-ra kapott adatok átlaga az /u/-s környezethez volt hasonló, alacsonyabb szóródással, míg a veláris esetében inkább az /i/-hez közelített ennek a kontextusnak a hatása. A két zöngétlen mássalhangzó esetében az /u/ kontextusában kaptuk a legmagasabb záridőtartamokat (8,2%, 16,5% eltérés az /i/-hez képest), a legnagyobb szóródást azonban az /ɔ/-s környezet eredményezte (35,4% és 43,2% eltérés az /i/-hez képest). Ennél a két mássalhangzónál az /i/ környezetében kaptuk a legalacsonyabb szóródást. Ennél a négy konszonánsnál ezek az eltérések statisztikailag relevánsnak bizonyultak (/b/, /p/, /g/, /k/: Kruskal–Wallis-próba: a χ^2 6,449 és 18,685 között, a $p \leq 0{,}039$ alakult; a többi konszonáns esetében a χ^2 0,041 és 3,438 között, a $p \geq 0{,}150$ alakult).



3. ábra
A mássalhangzók zárszakaszának időtartama (ms) a kontextus függvényében

A zár időtartama a képzési hellyel hátrafelé haladva jellemzően változott a felpattanó zárhangok esetében (3. ábra). Mindhárom magánhangzó kontextusában a bilabiálistól a palatális területig csökkent az időtartam, míg a veláris képzési helyű konszonánsok esetében ismét megemelkedett. Ez a hatás az /u/-s hangkörnyezetben bizonyult a leghangsúlyosabbnak. A képzési hely hatása a zár időtartamára statisztikailag is szignifikánsnak bizonyult (Kruskal–Wallis: a χ^2 40,796 és 66,092 között, a $p \leq 0{,}001$ alakult). Ezzel szemben az affrikáták esetében nem tér el a zárszakasz időtartama az egyes képzési helyek között (kivétel: a zöngétlenek az a kontextusában: $Z = -2{,}102$; $p = 0{,}036$, a többi esetben: a Z −1,847 és −0,043 között alakult, $p \geq 0{,}061$ volt).

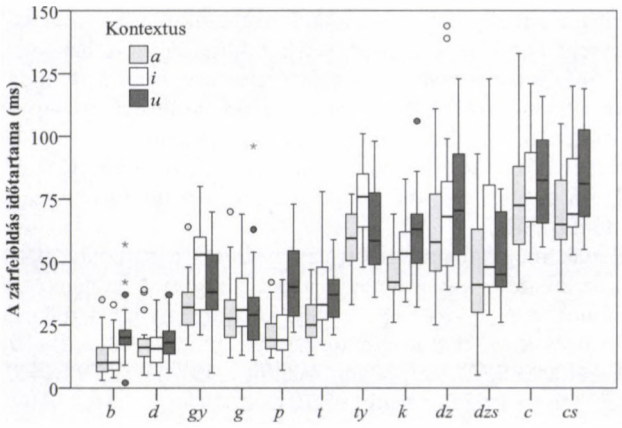
A zöngésségi párok tagjait összevetve majdnem minden esetben elmondható, hogy a zár időtartama a zöngés képzés során rövidebb a felpattanó zárhangok esetében (15,5–29,9%; Mann–Whitney: a Z −5,803 és −2,343 között, $p \leq 0{,}019$ volt). A palatális pár az /ɔ/ és az /i/, a veláris pedig az /u/ kontextusában nem tekinthető szignifikánsan eltérőnek (4,6–15,6%; Mann–Whitney: a Z −1,651 és −0,539 között, $p \geq 0{,}099$ volt). A zár-réshangok esetében a zárszakaszok időtartama mindhárom hangkörnyezetben fordítva alakult, a zöngések képzésében 18,5–32,5%-kal hosszabb időtartamot adatolhattunk (Mann–Whitney: a Z −3,726 és −2,023 között, $p \leq 0{,}043$ volt).

A zár felnyitásának időtartama a bilabiális és az alveoláris területen rövidebb, a palatális és veláris területen pedig hosszabb volt mind a zöngés, mind a zöngétlen explozívák esetében (zöngések: 15±10 ms; 18±9 ms; 40±16 ms; 33±17 ms; zöngétlenek: 30±15 ms; 34±14 ms; 66±16 ms; 54±16 ms). A zöngés felpattanó zárhangok megvalósulásaiban a hátsóbb képzési helyeken az értékek szóródása is nagyobb volt, illetve a palatálisok felnyílása hosszabb átlagos időtartamot mutatott.

A zár-rés hangok esetében azonos átlagos időtartamú felnyílással realizálódtak a két képzési helyen mind a zöngés, mind a zöngétlen fonémák, a /ʤ/ esetében azonban az értékek szóródása ötöde a /ʣ/-hez viszonyítva (zöngések: 26±25 ms; 23±5 ms; zöngétlenek: 20±40 ms; 19±39 ms).

A vizsgált magánhangzó-környezetek néhány mássalhangzó esetében gyakoroltak hatást a felnyílás időtartamára, míg mások esetében nem (4. ábra). A felpattanó zárhangok esetében jellemzően eltérően alakult a zár felnyitásának ideje az egyes magánhangzók környezetében. A /d/ esetében lényegében nincs ilyen hatás, a /g/ esetében csak tendenciaszerű, a /b/, /ɟ/ és a zöngétlen felpattanók megvalósulásaiban pedig minden esetben szignifikáns az eltérés. A két bilabiális esetében az /ɔ/-s és az /i/-s környezetben közel azonos volt a felnyitás időtartama, az /u/-sban azonban több mint 40%-kal emelkedett meg az átlagérték. Az alveoláris és palatális zöngétlen felpattanók zárnyitása a két felső nyelvállású magánhangzó környezetében közel azonos volt, az alsó nyelvállásfokú esetében azonban 22–30%-kal rövidebb értéket kaptunk. A zöngés palatális esetében ugyancsak az /ɔ/-s kontextusban volt a legrövidebb a feloldás időtartama, az /i/-ében viszont a leghosszabb (32,7%-os elté-

rés). A /k/ esetében ismét az /ɔ/-s környezetben kaptuk a legrövidebb, de az /u/-s esetében a leghosszabb felnyílási átlagos időtartamot (eltérés: 25,4%). A /g/ esetében ugyancsak jelentősen alacsonyabb volt az /ɔ/-k között mért érték, mint a másik két magánhangzó között (13,4%), de – mint fentebb említettük – ez nem érte el a szignifikancia szintjét. A Kruskall–Wallis-próba eredményei alapján a szignifikáns esetekben a χ^2 8,275 és 28,013 között, a $p \leq 0{,}016$ volt, a többi esetben a χ^2 0,905 és 1,280, a p 0,636 és 0,057 volt.



4. ábra
A mássalhangzók zárfeloldásának időtartama (ms) a kontextus függvényében

A zár-rés hangok esetében csak a /ʧ/ realizációiban volt a hangkörnyezet hatással a felnyílás időtartamára (Mann–Whitney: $Z = 6{,}072$; $p = 0{,}048$). E hangzó és a többi affrikáta esetében is az /ɔ/-s környezetben volt a legrövidebb a felnyílás időtartama, az /u/-ében a leghosszabb, de ez az eltérés a többi esetben csak kismértékű volt, a /ʧ/-nél elérte a 16,5%-ot.

Összegezve tehát elmondható, hogy az alsó nyelvállásfokú magánhangzó esetében a legrövidebb a zár felnyitása, míg a felső nyelvállásfokúak közül a mássalhangzó képzési helyéhez közelebb eső esetében a leghosszabb, vagy hasonló ezek között.

Az egyes kontextusokban külön összevetve a képzési hely alapján a zár feloldásának időtartamát lényegében a fentebb is említett tendencia rajzolódik ki: a felpattanók esetében az elülső két képzési helyen rövidebb, hátrébb pedig hosszabb időtartamokat kapunk, a leghosszabb értékeket a palatálisok esetében mérhettük. Az eltérés az /i/ kontextusában a legnagyobb (> 70%), mivel a két palatális zárhang felnyílása ebben a helyzetben a leghosszabb, a két palatálisé pedig az /ɔ/-éhoz hasonló értékű marad. Az eltérés minden magánhangzó környezetében szignifikáns volt mind a zöngés, mind a zöngétlen

explozívák esetében (Kruskal–Wallis-próba: χ^2 26,466 és 66,269 között, a $p < 0{,}001$ volt).

A zár-rés hangok megvalósulásaiban nem volt ilyen egyértelmű eltérés. A zöngések esetében kimutatható egy tendencia, miszerint a /ʣ/ zárfeloldása hosszabb mindhárom környezetben a /ʤ/-énél (17–31%), ez azonban csak az /u/ kontextusában éri el a szignifikanciaszintet (Mann–Whitney: a Z −3,184 és −1,476 között, a p pedig 0,001 és 0,081 között alakult). A zöngétlenek esetében ez a különbség 10% alatt maradt (Mann–Whitney: a Z −0,602 és −0,246 között, a p pedig 0,547 és 0,806 között alakult).

A zöngésségi párok tagjait összevetve a felpattanó zárhangoknál minden esetben 31–58%-kal hosszabb a pár zöngétlen tagjának zárfeloldásának időtartama. Az eltérés az első két képzési helyen jellemzően nagyobb, mint a hátsóbbak esetében, egy-egy kontextus (a /b/ esetében az /u/, a /d/ esetében az /ɔ/) kivételével. Az eltérés minden esetben szignifikáns (Mann–Whitney-próba: a Z −5,205 és −3,612 között alakult, a p pedig minden esetben kisebb volt, mint 0,001).

Az affrikáták megvalósulásaiban ez az eltérés a posztalveolárisok között megtalálható: 23–31%-kal hosszabb a pár zöngétlen tagjának zárfeloldása (Mann–Whitney: a Z −4,744 és −2,863 között, $p \leq 0{,}004$ volt). Az alveolárisok esetében csak az /ɔ/-s kontextusban mutatható ki eltérés (19,0% szemben a 9,5%-kal és a 9,6%-kal; Mann–Whitney: $Z = -2{,}079$; $p = 0{,}038$; illetve a Z −1,597 és −1,625, p pedig 0,110 és 0,104).

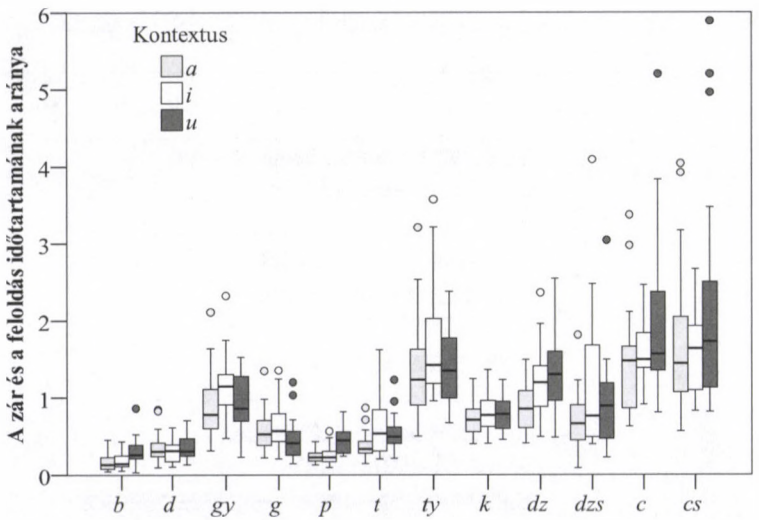
A zár feloldásának és a zárszakasz időtartamának arányát is elemeztük. A felpattanó zárhangok esetében a zár feloldásának időtartama jellemzően rövidebb a zárszakasz időtartamánál, hiszen a zár feloldása annak felpattintásával történik, míg a zár-rés hangoknál egy réses elem jelenik meg. A nyomásviszonyok és az esetleges többszöri felpattanások miatt azonban az explozívák között is előfordulhat hosszabb feloldás, továbbá a két palatális átmeneti és variábilis jellege miatt nagyobb szóródás feltételezhető. Ezt a feltételezést követték is a felnyílás adatai. A zárszakasz pedig a nyomásviszonyok miatt hosszabb lehetett az elülső képzési helyeken. A két veláris felpattanó azonban kivételt képzett a tendencia alól. Az affrikáták esetében a zárszakasz közel hasonló volt a két képzési helyen. Mindezek alapján várható, hogy a felpattanók esetében a képzési hellyel hátrafelé haladva egyre magasabb feloldás/zár arányt adatolhatunk mind a zöngés, mind a zöngétlen képzésű hangok esetében, az affrikátáknál azonban mindkét érték közel azonos volt, így ott az explozíváknál jóval magasabb, de a képzési helytől nem függő eredmények várhatóak. A két palatális zárhang esetében átmeneti arányt várhatunk (vö. pl. Kovács 2000, 2002).

A felpattanó zárhangok esetében a képzési hellyel hátrafelé haladva 21,8±14,3%, 35,3±18,4%, 107,1±88,0%, 56,3±28,5% a zöngések, illetve 33,6±20,4%, 51,4±27,4%, 150,1±63,1%, 77,1±23,5% volt a zöngétlenek esetében a feloldás időtartama a zárszakaszéhoz viszonyítva. A képzési hellyel

hátrafelé haladva tehát emelkedett az átlag és a szóródás is, a palatálisok esetében mind a zöngés, mind a zöngétlen sorban a legmagasabb (a velárishoz képest közel kétszeres értéket) kaptunk, a szóródás pedig ugyancsak közel háromszor magasabb a velárisokhoz hasonlítva.

Az affrikáták esetében a feloldás a vártaknak megfelelően általában hosszabb volt a zárszakasznál vagy azzal azonos időtartamú, emiatt az átlagos arány és a szóródás is jóval magasabb a felpattanókénál 113,1±47,4%, 95,6±70,4% (zöngések), 165,9±77,1% és 199,1±161,8% (zöngétlenek). A két palatális értékeit a két képzési mód adataival összevetve – ugyancsak a vártaknak megfelelően – köztes tartományban jelentek meg.

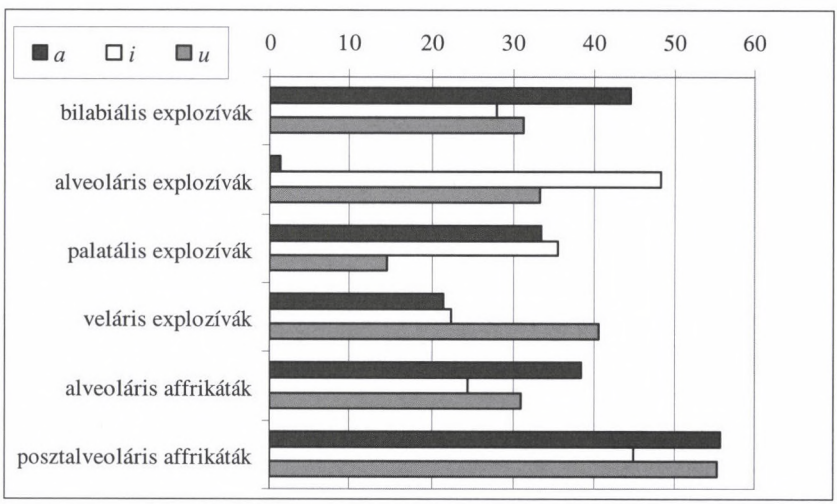
Az elemzett kontextusnak nincs jelentős hatása a feloldás és a zár időtartamának arányára (5. ábra). A zár-rés hangok esetében jellemző, hogy az /ɔ/-s kontextusban a legkisebb, az /u/-sban a legnagyobb ez az arány, a /ɟ/, /c/, /g/ esetében pedig az /i/-ében a legnagyobb, de ezek az eltérések statisztikailag nem szignifikánsak. A két bilabiális esetében az /u/-környezetben, a /t/ esetében az /i/-ében mért értékek szignifikánsan magasabbak a másik két környezetben kapott eredményekhez képest (Kruskal–Wallis-próba: a Z −24,471 és −9,987 között, $p \leq 0{,}007$ volt). Ez a nyolc explozíva fonémának kevesebb, mint a fele, így lényegében azt mondhatjuk, hogy erre az arányra nem gyakorolt jelentős hatást a három elemzett hangkörnyezet.



5. ábra
A mássalhangzók zárszakaszának és -feloldásának aránya a kontextus függvényében

Az egyes kontextusokban összevetve a mássalhangzók feloldás/zár arányát a képzési hely szerint az explozíváknál minden esetben szignifikáns eltérést kapunk a két palatális jellemzőiből eredően. A (mindig legkisebb arányt adó) /b/-hez viszonyítva 68–85% a különbség. A három további képzési helyet összevetve az eltérés 15–60% között alakul. A legkisebb a bilabiális-alveoláris párok között. Az affrikáták esetében csak a zöngések /ɔ/-kontextusában volt szignifikáns az eltérés (29,0%, szemben: 1,6–25,8%). Az átlagértékek közötti nagyobb eltérések ellenére, feltehetően az értékek nagy szórása miatt nem igazolható különbség. Az explozívák esetében a Kruskal–Wallis-próba eredményeiben a χ^2 35,438 és 71,493 között volt, a p pedig minden esetben kisebb volt 0,001-nél. Az affrikáták megvalósulásaira kapott Mann–Whitney-eredményekben a Z −1,420 és −0,064 közötti, a p pedig 0,156 fölötti értéket vett fel, a zöngések /u/-s kontextusára pedig $Z = -2{,}767$, $p = 0{,}006$.

A feloldás és a zár arányára a zöngésségi párok tagjait összevetve 14,6–55,7% közötti eltérést kapunk az alveoláris explozívák /ɔ/-s kontextusának kivételével (1,5%). Az egyes kontextusok hatását elemezve csak óvatos következtetéseket tehetünk a párok tagjainak eltérésére gyakorolt hatásukról. Az /ɔ/-környezetben a képzési hellyel hátrafelé haladva csökkenő különbséget figyelhetünk meg a felpattanó zárhangok között (a /d/ kivételével), az /u/-éval pedig növekvőt az összes vizsgált mássalhangzót együtt elemezve (6. ábra). A Mann–Whitney-próba alapján a Z −4,365 és −2,405 között, $p \leq 0{,}016$ volt. Az alveoláris explozívák esetében $Z = -0{,}979$, $p = 0{,}328$.



6. ábra
A zöngésségi párok tagjainak feloldás/zár arányának eltérése (%) az egyes magánhangzó-kontextusokban

A zöngésség fonetikai vetülete

A zöngés mássalhangzók alatt néhány esetben leállt a zöngeképzés a korábban említett aerodinamikai okok miatt, majd a zárfeloldás folyamán indult újra.

A nem approximánsként vagy réshangként megvalósult beszédhangok közül a felpattanó zárhangok esetében ez az esetek 18–49%-ában fordult elő, és a képzési hely hátrább haladtával gyakoribb volt, míg a zár-rés hangoknál azonban szinte minden esetben (/ʣ/: 97,2%, /ʤ/: 93,0%) bekövetkezett. A magánhangzó-kontextus hatása változó volt, a bilabiális kivételével az explozívák esetében az /ɔ/-s környezetben gyakrabban zöngétlenedtek részlegesen a megvalósulások (2. táblázat). A beszélők egyéni artikulációs sajátosságai is megfigyelhetőek, n3 és f1 esetében jellemzően gyakoribb volt a zöngétlenedés, n1 esetében ritkább, mint a többiek esetében. Mind a kontextus, mind a beszélők sajátosságainak hatása mellett jellemző a felpattanó zárhangok esetében a képzési hely hatása. A zár-rés hangok esetében kismértékben változó, hogy előfordul-e, és ha igen, mely képzési helyre jellemző 9–18%-ban a zönge fenntartása.

A zöngétlen fonémák realizációi alatt (természetesen) minden esetben leállt a zár ideje alatt a zönge, azonban az időtartama nagy szóródást mutatott.

2. táblázat. A zöngés fonémák zöngétlenedésének gyakorisága (%) a hangkörnyezet és a beszélő függvényében

	Kontextus			Beszélő					
	/ɔ/	/i/	/u/	n1	n2	n3	n4	f1	f2
/b/	8,7	25,0	20,0	8,3	16,7	8,3	0,0	75,0	0,0
/d/	41,7	17,4	21,7	25,0	10,0	33,3	16,7	50,0	25,0
/ɟ/	50,0	44,4	36,4	25,0	33,3	54,5	33,3	66,7	44,4
/g/	75,0	25,0	47,6	27,3	66,7	75,0	41,7	58,3	25,0
/ʣ/	100,0	95,7	95,8	100,0	100,0	100,0	91,7	100,0	91,7
/ʤ/	95,8	91,7	91,3	100,0	81,8	91,7	100,0	100,0	83,3

A zönge lecsengésének és újraindulásának időviszonyait azokban a konszonánsokban elemeztük, melyek valamely orális zárhangként jelentek meg, és a zöngeképzés legalább rövid időre leállt a képzés során.

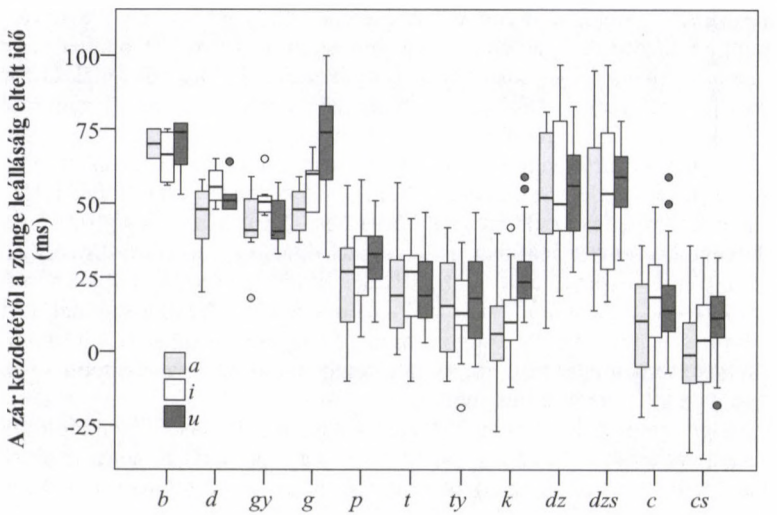
A zönge lecsengésének időtartama, azaz a zár kezdetétől a zönge leállásáig mért időtartam a képzési hellyel hátrafelé haladva jellemzően csökkent a zöngés explozívák esetében (68±9 ms, 49±11 ms, 45±10 ms, 56±16 ms), a veláris képzés kivételével. A zöngétlen fonémák megvalósulásaiban kisebb arányban, de megjelenik hasonló tendencia (29±14 ms, 21±14 ms, 16±15 ms, 15±16 ms).

Az affrikáták esetében a zöngés képzés során a zönge leállásának átlagos időtartama azonosnak tekinthető a két képzési helyen (55±19 ms, 53±21 ms). A két zöngétlen fonéma realizációi között az eltérés nagyobb mértékű volt (15±16 ms, 5±16 ms).

Az elemzett kontextusok között csak kevés esetben találtunk eltérést a zönge leállásának időtartamában (7. ábra). A két palatális explozíva esetében a palatális felső nyelvállásfokú magánhangzó között jellemzően kisebb az adatok szórása, mint a két veláris vokális között, az adatok eltérése azonban nem szignifikáns. A két veláris megvalósulásaiban a veláris felső nyelvállásfokú magánhangzó kontextusában hosszabb időtartamban maradt fenn a zönge, mint a másik két vokális esetében (/g/: 17,6% és 30,9%; /k/: 57,0% és 71,7%). A zár-rés hangok megvalósulásaiban csak a /ʧ/ esetében kaptunk szignifikáns eltérést a magánhangzó-környezetek hatásában, de megfigyelhető, hogy az /u/ közötti realizációkban kisebb az értékek szóródása mind a négy fonéma esetében. A három szignifikáns eltérés esetében a Mann–Whitney-próbában a Z értéke −20,433 és −7,290 között, a $p \leq 0{,}026$ volt. A magánhangzó-környezet tehát nem befolyásolja nagymértékben a zönge lecsengésének idejét. A zár és a mássalhangzó időtartamának hatása is felmerül. Elemeztük, hogy azokban a képzésekben, melyekben a zönge leáll, illetve amelyek során nem, eltér-e a zárszakasz, illetve a mássalhangzó teljes időtartama. A /b/ esetében nem különbözik egyik sem (Mann–Whitney: Z = −0,404, ill. −0,201, p = 0,444, ill. 0,214), míg a /d/ és a /g/ esetében a konszonáns teljes időtartama (Mann–Whitney: Z = −3,912, ill. −2,768, $p \leq 0{,}006$; ill. Z = −0,937, ill. −0,315, p = 0,349 és 0,752), a /ʝ/ esetében pedig mindkettő szignifikánsan hosszabb azokban az esetekben, ahol a zöngeképzés leállt az ejtés során (Mann–Whitney: Z = −2,400, ill. −3,264, $p \leq 0{,}016$). Az értékek azonban gyakorta esnek azonos tartományba, hiszen a beszélő egyéni sajátosságai is hatással vannak a zönge leállására (üregméretek, hangszalagok jellegzetességei, zöngeképzési jellemzők és az aktív kompenzációs stratégiák).

A zöngétlen fonémák megvalósulásaiban néhány esetben (13,5%) a zár kezdete előtt, az előaspiráció vagy a záródási zörej alatt állt le a zöngeképzés. A zöngések realizációiban ez nem fordult elő. A /p/ esetében egyszer sem (nem jelent meg ilyen szakasz a képzése során), a /t/ esetében egyszer, a /c/ esetében hétszer (a kontextus nem befolyásolta a gyakoriságot), a /k/ esetében kilencszer állt le a zönge ezen átmeneti szakaszok időtartama alatt. A velárisnál az /ɔ/-kontextusban volt a leggyakoribb (hat adat), az /u/-ében pedig nem fordult elő. A /ʦ/ esetében is kilenc ilyen esetet adatolhattunk, amelyek közül hat ismét az alsó nyelvállásfokú magánhangzó kontextusában jelent meg, az /i/-ében kétszer, az /u/-ében pedig egyszer. A /ʧ/ képzése során ugyancsak az /ɔ/ környezetében volt a leggyakoribb (12 előfordulás), az /i/-ében ritkább (nyolc eset), az /u/-ében pedig a legritkább (négy adat). Ezek az eredmények természetesen részben azzal is összefüggésben állnak, hogy egyáltalán hányszor jelent meg előaspiráció és/vagy záródási zörej (réses vagy

felpattanásos lenyomat). A zöngések esetében – mint korábban bemutattuk – ritka volt az ilyen szakasz megjelenése, míg a zöngétlenek között a képzési hellyel hátrafelé haladva egyre gyakoribb volt, a /c/ kivételével, amely az explozívák között a legtöbbször tartalmazott ilyen elemet. Ezen szakasz és az ez alatt leálló zönge előfordulásainak arányát összevetve a /t/ esetében ritka (2,7%), a /ʧ/ esetében relatíve gyakori (31,6%), hogy a zár teljes beállta előtt áll le a zöngeképzés (többi zöngétlen esetében: 11–13%).



7. ábra
A zönge lecsengési időzítése a zár kezdetéhez képest (ms)

Mindhárom magánhangzó-kontextus esetében elmondható, hogy a zöngétlen felpattanó zárhangok esetében a képzési hellyel hátrafelé haladva a zár kezdetéhez viszonyítva rövidebb ideig marad fenn a zöngeképzés. A zöngések esetében a ritkább és változó gyakoriságú előfordulás miatt csak tendenciaszerű megfigyelést tehetünk, azonban a két felső nyelvállásfokú vokális esetében a bilabiális képzés során jellemzően hosszabb ideig maradt fenn a zöngeképzés, az /ɔ/-s környezetben hasonló tendenciát kaptunk, a szignifikancia szintjét azonban nem érik el az eltérések (Kruskal–Wallis-próba: a χ^2 7,165 és 18,401 között, a $p \leq 0{,}067$ volt).

A zöngétlen affrikáták esetében az /u/-s környezet kivételével szignifikánsan hamarabb állt le a zönge a zár kezdetéhez képest a hátsóbb képzési helyen (Mann–Whitney: /ɔ/- és /i/-kontextus: $Z = -2{,}056$, $p = 0{,}040$ és 0,012, /u/-kontextus: $Z = -1{,}026$; $p = 0{,}305$). A zöngések esetében nem mutatható ki jellemző eltérés (Mann–Whitney: a Z 1,022 és $-0{,}610$ között, $p \geq 0{,}307$ volt).

A zöngésségi párok tagjai között a felpattanó zárhangok esetében fontos megemlíteni, hogy mivel a zöngés tag megvalósulásainak többségében a teljes időtartam alatt jelen van zönge, a zöngétlen alatt azonban leáll, ez nyilvánvalóan megkülönbözteti a párok tagjait. Mindemellett a zöngétlenedett előfordulások és a zöngétlenfonéma-realizációk számának eltérése miatt elsősorban tendenciaszerű következtetéseket vonhatunk csak le. A bilabiális és a veláris párokat összevetve csak kevés eset található, ahol a zöngés és a zöngétlen fonémák megvalósulásaiban hasonló értéket kaptunk a zár kezdetétől a zönge leállásáig eltelt időre, az alveolárisok esetében a két felső nyelvállásfokú magánhangzó közötti realizációk is ezt a mintázatot követik. A palatális pár esetében a zöngés tag megvalósulásai a zöngétlenek felső kvartilisának tartományában jelentek meg (7. ábra). Az affrikáták esetében valamivel gyakrabban fordul elő, hogy a zöngésségi párok tagjai azonos értékhalmazban jelennek meg, az átfedés azonban ezek esetében sem éri el az 50%-ot. A Mann–Whitney-próba alapján mindhárom kontextusban mind a hat zöngésségi pár esetében szignifikáns az eltérés (a Z –5,912 és –2,307 között, a $p \leq 0{,}007$ volt).

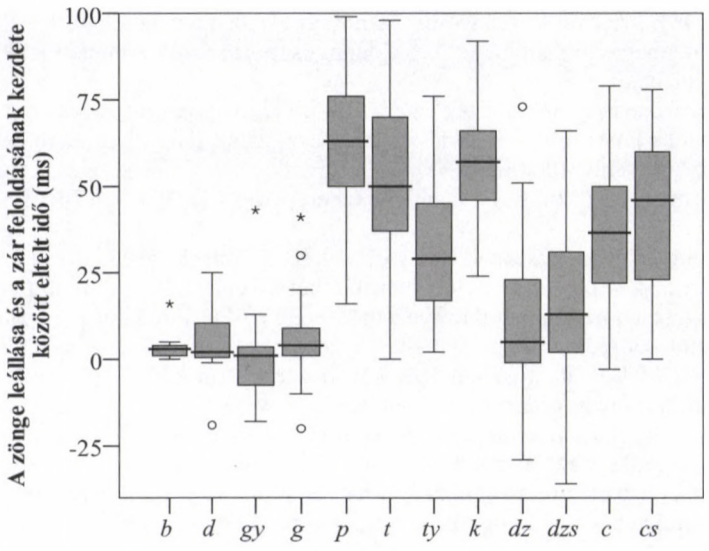
Elemeztük **a zönge leállása és a zár feloldásának kezdete között eltelt időt.** A zöngés explozívák esetében ismét kevés, a képzési hellyel hátrafele haladva növekvő számú adatunk van a zöngétlenedés gyakoriságának következtében. Továbbá mind a zöngés, mind a zöngétlen explozívák lenyomatában kevés esetben nem volt adatolható a zár feloldása – vö. fentebb –, ezekben ezt az értéket nem tudtuk elemezni.

A zöngés felpattanó zárhangok esetében a zár feloldása általában a zönge leállása után kezdődik (8. ábra). A bilabiális képzés esetében nem, az alveolárisában pedig kétszer (10,5%) fordult elő, hogy a zár feloldásának kezdete után állt le a zöngeképzés. A /g/ megvalósulásaiban háromszor (8,8%). A /ɟ/ előfordulásaiban jóval gyakrabban (tíz eset, 37,0%) következett be a zöngének a zárfeloldás megindulása után való leállása. Ritkán (2-3 előfordulás fonémánként) adatolhattunk olyan megvalósulást, ahol ez az érték pontosan 0 ms volt. A zönge leállása és a feloldás kezdete között eltelt idő 4±4 ms, 5±9 ms, 0±12 ms és 6±21 ms volt a képzési hely szerint hátrafele haladva. A zöngétlen explozívák esetében a kevés esetben hosszabb ideig fennálló passzív zönge megjelenése ellenére is minden esetben előbb állt le a zönge képzése, mint hogy megkezdődött a zár feloldása. Egy előfordulást (/t/) adatolhattunk, ahol pontosan 0 ms volt az érték. Az átlagos értékek 61±17 ms, 51±23 ms, 31±17 ms, 56±15 ms a bilabiálistól a velárisig haladva.

A zöngés alveoláris affrikáta esetében ismét gyakrabban fordult elő 18 esetben, hogy a zár feloldása során állt le a zöngeképzés (26,1%), a posztalveoláris esetében pedig nyolcszor (8,8%). A zöngétlenek esetében összesen egyszer, az alveoláris képzéshelyen fordult elő (1,4%). A zöngések esetében nyolcszor-nyolcszor fordult elő, hogy a zönge leállása és a zár feloldásának kezdete pontosan egybe esett, a zöngétlenek esetében ez összesen négyszer

következett be. Az átlagosan eltelt idő az előbbiek esetében 9±19 ms és 17±22 ms, az utóbbiakéban pedig 37±20 ms és 42±21 ms.

A zöngés explozívák esetében a fonetikai kontextus hatása nem elemezhető statisztikailag, mivel az befolyásolta a zárfeloldás elmaradásának, illetve a zöngétlenedésnek a gyakoriságát is, így az egyes hangkörnyezetek között nagymértékben eltérő adatmennyiséget kaptunk. A zöngétlen explozívák esetében a legnagyobb kontextusbeli eltérés 13,0%. Az affrikáták esetében az adatok mennyisége alapján elemezhető a kontextus hatása. A zöngétlen alveoláris fonéma megvalósulásai esetében mondható el, hogy az eltérés szignifikáns (29,3% szemben 9,8%-kal). A zöngés zár-rés hangok esetében az átlagok nagymértékű eltérése (62,9% és 27,1%) ellenére az adatok szóródása miatt nem igazolható kontextuális hatás. A Kruskal–Wallis-próba alapján a /ts/ esetében $\chi^2 = 6{,}385$; $p = 0{,}041$; a többi esetben pedig a χ^2 0,494 és 4,091 között, a p pedig $\geq 0{,}129$ volt.



8. ábra
A zönge lecsengése és a zár feloldása között eltelt idő (ms)

A kontextusok alapján külön elemeztük a képzési hely hatását, így a zöngés felpattanó zárhangokat kizártuk a vizsgálatnak ezen részéből is. A zöngétlen explozívák esetében a bilabiálisra jellemzőek a legmagasabb adatok, a palatálisra pedig a legalacsonyabbak a gyakran megjelenő negatív értékek miatt, ezért az eltérés statisztikailag szignifikáns. Ugyancsak jellemző a szó-

ródás csökkenése a képzési hellyel hátrafelé haladva. A Kruskal–Wallis-próba alapján a χ^2 23,763 és 33,392 között, a $p \leq 0{,}001$ volt.

Az affrikáták esetében csak a zöngések mutatottak egy, az /u/ környezetben, a zöngeleállás-zárfeloldás szakasz időtartama és a képzési hely között összefüggést (Mann–Whitney-próba: $Z = -2{,}767$; $p = 0{,}006$; a többi esetben: a Z −1,420 és −0,064 között, $p \geq 0{,}156$ volt).

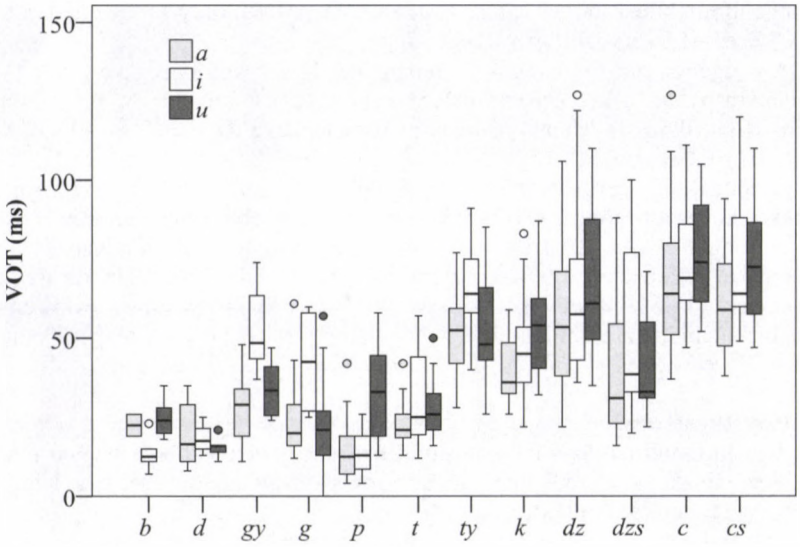
A zár feloldásának kezdetétől a zönge megindulásáig eltelt időt csak azokban az esetekben elemezzük, amelyekben a hangszalagok rezgése leállt a képzés során. Emiatt a zöngés felpattanók esetében csak kevés és eltérő számú előfordulást vizsgáltunk. Elemeztük azokat a beszédhangokat is, amelyek során a zönge a zárfeloldás alatt állt le.

A zöngés felpattanó zárhangok esetében a zönge újraindulásának ideje az elülsőbb képzési helyeken rövidebb (19±8 ms, 18±8 ms), a hátsóbbakon pedig hosszabb (37±15 ms és 27±16 ms). A zöngétlen fonémák megvalósulásaiban már az elülsőbb képzési helyek között is nagyobb az eltérés (21±14 ms, 27±11 ms, 57±15 ms, 47±15 ms). Mindkét esetben a palatálisok VOT-értéke volt a legmagasabb, ami a két beszédhangcsoport közötti átmeneti jellegűkkel magyarázható.

Az affrikáták esetében ezek az értékek jóval magasabbak, amely egyrészt a jellemzően hosszabb zárfeloldással, másrészt az az alatti réses elem és a zöngeképzés további ellentmondásának a következménye. A zöngések esetében 62±24 ms és 45±22 ms, a zöngétlenekében pedig 74±19 ms és 70±19 ms volt a VOT.

Az elemzett kontextusok hatással voltak a zöngekezdési idő alakulására (9. ábra). Az explozívák esetében csak a két alveoláris nem mutatott szignifikáns eltérést a három hangkörnyezetben, meg kell ugyanakkor jegyezni, hogy az adatok szóródása jelentősen eltért az egyes esetekben. A zöngés bilabiális esetében az /i/ kontextusában 46%-kal rövidebb volt a VOT, mint a két veláris között, míg zöngétlen párja esetében az alsó nyelvállásfokú magánhangzó hatására tért el a többi hangkörnyezettől (57,4%-kal magasabb az átlag ebben a helyzetben). A két hátsóbb képzési helyű zöngés felpattanó zárhang esetében ismét a palatális magánhangzó hatása mutatható ki, azonban ezekben a realizációkban ez a hangkörnyezett okozta a leghosszabb VOT-értékeket (46,4%-kal és 42,9%-kal hosszabbakat az alsó nyelvállásfokú vokálishoz képest). A zöngétlen palatális zöngés párjához hasonlóan ebben a kontextusban mutatta a legmagasabb VOT-t, de az eltérés kisebb (23,8%). A zöngétlen veláris zöngés párjától eltérően az /u/-kontextusban vette fel a legmagasabb zöngekezdési időket (25,1%). Ez utóbbi és az affrikáták esetében elmondható, hogy az egyes hangkörnyezetek között fokozatos emelkedés tapasztalható, azaz nem egy-egy magánhangzó esetében tér el nagymértékben a két másik közel azonos hatást gyakorló vokális környezetében kapott adatoktól, hanem egy „mintázatsort" állíthatunk fel. A zöngétlen veláris explozíva és a két alveoláris affrikáta esetében az /ɔ/ < /i/ < /u/ sor, a két posztalveoláris esetében

pedig az /ɔ/ < /u/ < /i/ jellemző. Az affrikáták közül a magánhangzókontextus csak a zöngétlen alveoláris esetében bizonyult szignifikáns hatásúnak. A szignifikáns esetekben a Kruskal–Wallis-próba alapján a χ^2 6,258 és 27,851 között, a $p \leq 0{,}044$ volt. A többi esetben a χ^2 0,513 és 4,894 között, a $p \geq 0{,}087$ volt.



9. ábra
A zárfeloldás kezdetétől a zönge megindulásáig eltelt idő (ms)

A képzési hely a felpattanó zárhangok esetében jellemzően hatással volt a zönge újraindulásának idejére. A zöngés explozívák esetében az alsó nyelvállásfokú magánhangzó környezetében nem tért el szignifikánsan a VOT-érték e tényező mentén, de a többi kontextushoz hasonlóan a palatális mutatta a legmagasabb értéket. A két felső nyelvállásfokú magánhangzó hatására nagyobb volt az egyes képzési helyek között a VOT-ben mért eltérés. Ezekben az esetekben is a palatális mássalhangzó esetében mértük a leghosszabb zönge-újraindulási időt (19–74%-kal haladta meg a többi konszonánsnál mért értéket). A palatális vokális esetében a veláris mássalhangzó adatai is megemelkedtek (57–68%-kal hosszabb a két elsőbb képzési helyhez képest), de a palatálisnál így is rövidebbek maradtak. A Kruskal–Wallis-próba alapján a χ^2 9,665 és 17,421, a $p \leq 0{,}022$ volt a két felső nyelvállásfokú magánhangzó esetében, az alsóéban pedig $\chi^2 = 3{,}204$; $p = 0{,}361$.

A zöngétlen felpattanó zárhangok esetében általában a bilabiális < alveoláris < veláris < palatális sor állítható fel. A felső nyelvállásfokú veláris kontex-

tus esetében az alveoláris valamivel rövidebb VOT-t vesz fel, mint a bilabiális. Az eltérések mindhárom hangkörnyezetben szignifikánsak. A Kruskal–Wallis-próba alapján a χ^2 44,086 és 71,301 között, a $p < 0{,}001$ volt.

A zár-rés hangok esetében a zöngések a két veláris esetében szignifikánsan, a palatális magánhangzó környezetében pedig tendenciaszerűen mutattak eltérést a két képzési hely értékei között. Az alveoláris 18–40%-kal hosszabb zönge-újraindulási idővel valósult meg. A Mann–Whitney-próba alapján a Z −3,832 és −1,925 között, $p \leq 0{,}054$ volt.

A zöngétlen zár-rés hangok esetében 0,7–11,7% különbség volt az egyes kontextusokban a két képzési helyre kapott adatok között. Ez egy esetben sem szignifikáns (a Mann–Whitney-próba alapján a Z −1,070 és −0,330 között, $p \geq 0{,}285$ volt).

A zöngésségi párokat az explozívák esetében nem vetettük össze, mivel sokkal kevesebb zöngés előfordulás volt (a kevesebb zöngétlenedés miatt), mint zöngétlen. Az affrikáták esetében a posztalveolárisok mindhárom kontextusban jellemzően eltérnek egymástól (29–33%), az alveolárisok az /u/-ében nem (11,8% szemben: 13,5% és 22,4%). A Mann–Whitney-próba eredményei alapján ez utóbbi esetben a $Z = -1{,}600$; $p = 0{,}110$; a többiben pedig a Z −4,771 és −2,090 között; $p \leq 0{,}037$ volt.

Következtetések

A jelen tanulmányban laboratóriumi beszédben elemeztük a magyar nyelv explozíváinak és affrikátáinak időtartamviszonyait a zöngésség, a képzési hely és a fonetikai kontextus függvényében.

Összességében elmondható, hogy az explozívák között gyakorta egyértelmű mindhárom paraméter hatása, míg az affrikáták esetében többször csak tendenciaszerű összefüggés mutatható ki, esetleg az adatok alapján nincs összefüggés az elemzett értékek és a vizsgált jellemzők között.

A mássalhangzó teljes időtartama, belső szupraglottális jegyeinek időzítése és azoknak a glottális jellemzőkkel való összevetése is jellemzően összefüggést mutatott mind a képzési hellyel, mind a magánhangzó-környezettel, mind pedig a fonológiai zöngésséggel. A konszonáns, zárszakaszának és zárfeloldásának időtartama, valamint a két utóbbi aránya is eltérő volt a képzési hely függvényében. A hátsóbb képzési területeken általában hosszabb időtartamokat adatolhattunk. A zárszakasz időtartamának a képzési hellyel hátrafele haladva történő csökkenése, a feloldásnak pedig a növekedése a kisebb zár mögötti térfogattal és kisebb passzívan vagy aktívan tágítható felülettel magyarázható, hiszen gyorsabban emelkedik meg a képzés alatt a nyomás, a zár felnyitását eredményezve. A velárisok esetében az ismét hosszabb ideig tartó zár feltehetően a zárképzésben rész vevő nyelvterület nagyobb tehetetlenségének a következménye.

A palatálisokra kapott értékek általában meghaladták a velárisokét is, ezzel az affrikátákhoz hasonló szerkezetet jelezve. Ez a korábban említett átmeneti

fonetikai jellegük következménye, vagyis a részvevő nyelvterület mérete és lomhasága miatt a zár és annak feloldása a hangkörnyezet függvényében az affrikátákhoz hasonlóbb szerkezetet eredményezhetett.

Az affrikáták esetében ritkán kaptunk a képzési hely függvényében eltérő értékeket az egyes vizsgált paraméterekben. Ennek feltehetően az az oka, hogy a távolság kisebb, így a hatása nem jelentős mértékű. A fonológiai zöngésség tekintetében a fonetikai zöngésség jellemzői eltérőek voltak, de a szupraglottális időviszonyok csak a posztalveolárisok néhány esetében. Több ok is szerepet játszhat az eredmények alakulásában. Egyrészt a zöngés alveoláris affrikáta ejtése gyakorta hosszú beszédhangként történik, habár a korábbi adatok variabilitása alapján (feszes ejtésben) nem minden esetben tapasztalható ez a jelenség (vö. pl. Magdics 1969; Kassai 1982).

A magánhangzó-környezet hatását összegezve elmondható, hogy az alsó nyelvállásfokú magánhangzó általában rövidebb időtartamokat eredményezett, és valamely (általában a mássalhangzó képzési helyéhez közelebb eső) felső nyelvállásfokú vokális pedig hosszabbakat. Így például az alsó nyelvállásfokú magánhangzó esetében a legrövidebb a zár felnyitása, míg a felső nyelvállásfokúak közül a mássalhangzó képzési helyéhez közelebb eső esetében a leghosszabb, vagy hasonló ezek között. A VOT-értékek alakulásában is a mássalhangzó képzési helyének függvényében változott a magánhangzó képzési jegyeinek hatása a korábbi eredményekhez (Gósy 2000) hasonlóan. Ennek oka feltehetően az, hogy a mássalhangzó-magánhangzó átmenet között megtett nagyobb távolság következtében gyorsabban tud távozni a zár mögött felgyűlt levegő, azaz hamarabb csökken a nyomás, és állnak fel a magánhangzóra jellemző aerodinamikai feltételek.

A szupraglottális jellemzők, így a képzési mód is jellemző hatással volt a fonetikai zöngésség alakulására. A zöngeképzés a zöngétlen képzés során jellemzően a zárfeloldás előtt hosszabb idővel állt le mindkét képzési módban, a zöngés explozívák esetében ritkán az után, az affrikáták esetében pedig valamivel gyakrabban esett ez a zár felnyitásának indulása utánra. Ennek oka feltehetően, hogy a zárszakasz rövidebb ezekben a mássalhangzókban. A zöngés explozívák esetében ritkább volt a zöngétlenedés (amely azonban a képzési hellyel hátrafelé haladva gyakoribbá vált), mint az affrikáták esetében. Ennek oka a zár utáni képzési rész eltérésében keresendő. Feltehetően az intenzív spirantikus zörej mint artikulációs-akusztikai cél és a zöngeképzés során a hangszalagok felől áramló alacsonyabb légnyomás ellentmondása áll a háttérben.

Az eredmények alapján további kérdések is felmerülhetnek. Ilyen például az akusztikai szerkezet spektrális elemzésének szükségessége, amely a palatális konszonánsok két csoport közötti helyzetét részletesebben leírhatják; vagy a fonetikai kontextusban rejlő másodlagos kulcsok vizsgálata.

Irodalom

Alphen, Paul M. van – Smith, Roel 2004. Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: the role of prevoicing. *Journal of Phonetics* 32/4. 455–449.

Berg, Jur van den 1958. Myoelastic-aerodynamic theory of voice production. *Journal of Speech and Hearing Research* 1. 227–244.

Bickley, Christine A. – Stevens, Kenneth N. 1986. Effects of a vocal tract constriction on the glottal source: experimental and modeling studies. *Journal of Phonetics* 14. 373–382.

Boersma, Paul – Weenink, David 2011. *Praat: doing phonetics by computer* (Version 5.1.13). Elérhető: http://www.fon.hum.uva.nl/praat/download_win.html

Bóna Judit 2011. A [p, t, k] mássalhangzók zöngekezdési ideje idősek és fiatalok spontán beszédében és felolvasásában. *Beszédkutatás 2011*. 61–72.

Cseresnyési László 1992. An outline of Hungarian phonology. *The Journal of Intercultural Studies. Extra Series* No. 2. 79–104.

Draxler, Christoph – Jänsch, Klaus 2004. SpeechRecorder, a universal platform independent multi-channel audio recording software. In: *Proceedings of LREC*, Lisbon. 559–562.

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

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

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

Gráczi Tekla Etelka 2011a. Intervokális explozívák a zöngésségi oppozíció függvényében. *Beszédkutatás 2011*. 46–60.

Gráczi, Tekla Etelka 2011b. Voicing contrast of intervocalic plosives in Hungarian. In Lee, Wai-Sum – Zee, Eric (ed.): *17th International Congress of Phonetic Sciences.* Hong Kong. 759–762.

Gráczi Tekla Etelka – Kohári Anna 2012. A zöngekezdési idő egy módszertani kérdés függvényében. In Markó Alexandra (szerk.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngekezdési időig*. ELTE BTK–MTA NyTI, Budapest. 228–248.

Gráczi Tekla Etelka – Markó Alexandra – Beke András 2009. Zöngekezdési idő a spontán beszédben. Elhangzott: *Beszédkutatás 2009* konferencia.

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

Kázmér Miklós 1961. *A magyar affrikátaszemlélet.* Nyelvtudományi Értekezések 21. Akadémiai Kiadó, Budapest.

Keating, Patricia – Linker, Wendy – Huffman, Marie 1983. Patterns in allophone distribution for voiced and voiceless stops. *Journal Phonetics* 11. 277–290.

Kovács Magdolna 2002. Az affrikáták akusztikai szerkezetéről. In Hunyadi László (szerk.): *Kísérleti fonetika, laboratóriumi fonológia.* Debreceni Egyetem Kossuth Egyetemi Kiadó, Debrecen. 39–54.

Ladefoged, Peter 1941. *Preliminaries to linguistic phonetics.* The University of Chicago Press, Chicago–London.

Maddieson, Ian 1997. Phonetical universals. Hardcastle, William J. – Laver, John (eds.): *The handbook of phonetic sciences*. Blackwell, Oxford. 619–639.

Magdics Klára 1969. A magyar beszédhangok időtartama. *Nyelvtudományi Közlemények* 68. 125–139.

Ohala, John J. 1976. *A model of speech aerodinamics*. Report of the phonological Laboratory. Berkeley.

Ohala, John J. 1983. The origin of the sound patterns in vocal tract constraints. In MacNeilage, Paul F. (ed.): The production of speech. Springer Verlag, New York. 189–216.

Ohala, John J. 1990. There is no interface between phonetics and phonology. A personal view. *Journal of Phonetics* 18. 153–171.

Ohala, John J. 1997. Aerodynamics of phonology. In: *Proceedings 4th Seoul International Conference on Linguistics*. Seoul. 92–97.

Ohala, John J. – Riordan, Carol J. 1979. Passive vocal tract enlargement during voiced stops. In Wolf, Jared J. – Klatt, Dennis H. (eds.): *Speech Communication Papers*. Acoustical Society of America, New York. 89–92.

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

Pape, Daniel – Mooshammer, Christin – Hoole, Philip – Fuchs, Susanne 2006. Devoicing of word-initial stops: A question of the following vowel? In Harrington, John – Tabain M. (eds.): 211–226.

Perkell, Joseph S. 1969. *Physiology of speech production: Results and implications of a quantitative cineradiographic study*. Research monograph 53. MIT Press, Cambridge, MA.

Shadle, Christine H. 1997. The aerodynamics of speech. In Hardcastle, William J. – Laver, John (eds.): *The handbook of phonetic sciences*. Oxford, Blackwell. 33–64.

Siptár Péter 1995. *A magyar mássalhangzók fonológiája*. Linguistica, Series A, Studia et Dissertationes 18. MTA Nyelvtudományi Intézet, Budapest.

Solé, Maria J. 2011. Articulatory adjustments in initial voiced stops in Spanish, French and English In Lee, Wai-Sum – Zee, Eric (ed.): *17th International Congress of Phonetic Sciences*. Hong Kong. 1878–1883.

Solé, Maria J. – Sprouse, Ronald L. 2011. Voice-initiating gestures in Spanish: Prenasalization. In Lee, Wai-Sum – Zee, Eric (ed.): *17th International Congress of Phonetic Sciences*. Hong Kong. 72–75.

Stevens, Kenneth N. 1971. Airflow and turbulence noise for fricative and stop consonants: Static considerations. *Journal of Acoustical Society of America* 50. 1180–1992.

Stevens, Kenneth N. 1997. Articulatory–acoustic–auditory relationships. In Hardcastle, William J. – Laver, John (eds.): *The handbook of phonetic sciences*. Blackwell, Oxford. 462–506.

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 Acoustic Society of America* 91/5. 2979–3000.

Svirsky, Mario – Stevens, Kenneth N. – Matthies, Melanie L. – Manzella, Joyce – Perkell, Joseph S. – Wilhelms-Tricarico, Reiner 1997. Tongue surface displacement during bilabial stops. *Journal of the Acoustical Society of America* 102. 562–571.

Szende Tamás 1974. Magánhangzóközi affrikátáink természetéről. *Magyar Nyelv* 71. 432–438.

Szende Tamás 1997. *Alapalak és lazítási folyamatok.* Linguistica Series A Studia et dissertationes 22. MTA Nyelvtudományi Intézet, Budapest.

Westbury, John R. – Keating, Patricia A. 1986. On the naturalness of stop consonant voicing. *Journal of Linguistics* 22. 145–166

A KITÖLTÖTT SZÜNETEK FONETIKAI SAJÁTOSSÁGAI AZ ÉLETKOR FÜGGVÉNYÉBEN

Gósy Mária – Bóna Judit – Beke András – Horváth Viktória

Bevezetés

A beszédben különböző funkciókban, különféle gyakorisággal fordulnak elő a kitöltött szünetek (vagy másképpen hezitálások, hezitációk). Megjelenésüket előidézheti az, hogy a beszéd sebessége gyorsabb, mint a mondandó tartalmának kialakítása, avagy diszharmónia alakul ki a beszédtervezés és a szóelőhívás között. A spontán beszédben főként azért fordulnak elő a kitöltött szünetek, mert időt biztosítanak a gondolkodásra és a beszédprodukciós folyamatok összehangolására (pl. Levelt 1989; Berger–Jordan 1992). A kitöltött szünetek megjelenhetnek az önmonitorozás jeleként vagy a társalgásokban a beszédlépések határán (vö. Christenfeld 1996; Ward 2000; Shilcock et al. 2001; Shriberg 2001; Clark–Fox Tree 2002; Eklund 2004). Az eddigi kutatások szerint a kitöltött szünetek ejtése független a beszélő életkorától, gyermekek spontán beszédében éppúgy tapasztalhatók, mint időseknél (vö. Ragsdale–Sisterhen 1984; Yaruss et al. 1999; Esposito et al. 2004).

A kitöltött szünet a nemzetközi szakirodalomban esernyőterminusként funkcionál (lásd Postma–Kolk 1993; Lickley–Bard 1996; Shriberg 2001; Eklund 2004; Thordardottir–Weismer 2002; Corley–Hartsuiker 2003; Ferreira et al. 2004; Hudson Kam–Edwards 2008; Corley–Stewart 2008), ami azt jelenti, hogy számos, különféle megakadásjelenséget foglal magában, például ismétlés, újraindítás, nyújtás, töltelékszó, egyes diskurzusjelölők stb. (vö. Shriberg 1994; Hokkanen 2001). A jelen tanulmányban a *kitöltött szünet* terminust azokra a jelenségekre használjuk, amelyek egy magánhangzóból vagy egy mássalhangzóból állnak, illetve egy magánhangzó és egy vagy több mássalhangzó kapcsolatából jönnek létre, és amelyek nem lexémák a nyelvünkben. Közülük a leggyakoribb az *ö*-féle hangzás (innen ered az *özés* elnevezés is). A kitöltött szünetek artikulációs sajátosságai nyelvenként különbözők. Az angolban az *uh, uhm* vagy az *er* hangsorokat ejtik (Roberts et al. 2009), a portugálban az *uum,* az [ɐ] és ritkábban az [ə] magánhangzót (Veiga et al. 2011), a japánban *az ano, e, eto* és *ma* fordul elő leggyakrabban (Watanabe et al. 2008), a franciában a magyarhoz hasonlóan az [ø] (*euh,* l. Duez 2001), a dominikai spanyolban az *eh* (Roggia 2012), a törökben pedig a *şey* (Furman–Özyürek 2007). A kitöltött szünetek formái összefüggést mutathatnak a funkcióval (Levelt 1989; Clark 1994; Smith–Clark 1993; Duez 2001; Clark–Fox

Tree 2002). A bilabiális zár gyakran megjelenik a kitöltött szünet végén, jelezve, hogy a beszélő befejezte a beszédfordulót (Shriberg 2001).

A kitöltött szüneteket számos aspektusból elemezték mind a beszédprodukcióban, mind a beszédpercepcióban. Olyan kutatások is voltak, amelyekben a hezitálások gyakorisága alapján azt ítélték meg, hogy vajon a beszélő mennyire jártas az adott témában, avagy mennyire megbízható a közlése (Smith–Clark 1993; Eklund 2004; Shriberg 2001; Berger–Jordan 1992; Christenfeld 1995; Brennan–Williams 1995; Lickley 1995; Fox Tree 2001; Brennan–Schober 2001; Clark–Fox Tree 2002; Bailey–Ferreira 2003; Ferreira et al. 2004; Corley et al. 2007; Corley–Stewart 2008; Finlayson–Corley 2012 stb.).

A különböző szerzők általában egyetértenek a kitöltött szünetek funkcióiban és előfordulásuk okaiban (pl. Smith–Clark 1993; Arnold et al. 2003; Simpson 2006). A hezitációs jelenségek beszédtervezési és beszédkivitelezési problémát jelezhetnek anélkül, hogy a felszínen hiba jelenne meg (Levelt 1983; Hartsuiker–Kolk 2001); előfordulhatnak az önmonitorozás részeként (Christenfeld 1996; Swerts 1998; Clark–Fox Tree 2002; Simpson 2006; Finlayson–Corley 2012), illetve utalhatnak arra, hogy a beszélő bizonytalan abban, hogyan folytassa a mondanivalóját, milyen grammatikai szerkezetet használjon (Levelt 1989; Shriberg 2001; Watanabe et al. 2008). A kitöltött szünetek gyakran jelennek meg hibajelenség kísérőjeként; a párbeszédekben és a társalgásokban jelezhetik, hogy a beszélő még nem kívánja átadni a szót, avagy éppen át akarja adni azt, utóbbiak diskurzusjelölői funkciók lehetnek (pl. Bortfeld et al. 2001; Local 2004; Horváth 2010). Levelt (1989) szerint az *er* formájú hezitálás univerzális szóként funkcionál a szerkesztési szakaszban. A kitöltött szünetek lehetnek a beszédstílus (Finlayson–Corley 2012) vagy a kommunikációs helyzet nehézségének a következményei (Bóna 2013). Jelezhetik az egyén mentális állapotát, különösen a szorongást, habár ezen a téren ellentmondó eredmények születtek (vö. Pope et al. 1970; Ragsdale 1976; Schachter et al. 1994; Christenfeld–Creager 1995). A kétnyelvűek kódváltásában speciális funkciójuk van (pl. Hlavac 2011).

A kitöltött szünetek megjelenhetnek teljes beszédaktus, mondat, szó előtt vagy után, rendszerint azonban grammatikai szerkezetek határán fordulnak elő (vö. Shriberg 1994, 1996; Clark–Wasow 1998; Duez 2001; Watanabe et al. 2008). A hezitációs jelenségek gyakoriságát számos tényező befolyásolja (a beszélő életkora, neme, a téma nehézsége, a grammatikai szerkezet, a beszédstílus, a megnyilatkozás témája és hossza vagy a beszélők közötti kapcsolat stb.), és mindez változhat a beszélőktől és a kommunikációs szituációtól függően is (Shriberg 1996; Yaruss et al. 1999; Merlo–Mansur 2004; Bortfeld et al. 2001; Watanabe et al. 2008; Roberts et al. 2009; Horváth 2010).

Nagy mennyiségű adaton végzett kutatások azt mutatták, hogy a kitöltött szünetek az összes megakadás harmadát, akár több mint a felét is kitehetik

(Shriberg 1994). Svéd beszélők társalgásai során az *eh* típusú kitöltött szünet 3,5%-ban jelent meg az összes szószámhoz viszonyítva (Eklund 2004). Kanadai angol beszélők különböző témájú narratíváiban átlagosan 4 kitöltött szünet fordult elő mintegy 70 szóra vetítve (Roberts et al. 2009). Az *um*-féle hezitálásokat vizsgálták fiatal felnőtteknél, akiket a szorongásuk foka szerint osztottak két csoportba. A szorongó beszélők percenként 7,03 db *um* szünetet tartottak, míg a nem szorongók átlagosan 4,07 darabot. Ugyancsak a mentális állapottól függően adatoltak percenként 2,07–5,6 db *um* hezitálást egy másik kísérletben (Christenfeld–Creager 1996). Formális előadásokban 3,17 db *um*-féle szünetet találtak percenként (Christenfeld et al. 1991). Alkoholos állapotban a beszélők kevesebb *um*-típusú hezitálást produkáltak, mint a józan adatközlők, ezt a kutatók az önmonitorozás hiányával magyarázzák az előbbieknél (l. Christenfeld–Creager 1996). Magyar beszélőknél azt találták, hogy az összes szó arányában mintegy 4,4% volt a kitöltött szünet (Gósy 2003), egy másik vizsgálat szerint a fiatal felnőttek spontán beszédében átlagosan 3,8 kitöltött szünet fordult elő percenként (Horváth 2010). Fiatal magyar beszélők történet-visszamondásaiban percenként 9,9 hezitálást találtak, ez a vizsgált szövegekben azt jelentette, hogy 11,2 szavanként fordult elő egy kitöltött szünet (Gósy–Bóna 2011). Fiatal felnőttek 25 szavanként produkáltak egy kitöltött szünetet narratívában és közel 52 szavanként társalgásban (Bóna 2013).

A kitöltött szünetek különféle időtartamokban jelenhetnek meg: lehetnek rövidebbek, mint 100 ms, de akár 750 ms-ban vagy még hosszabban is realizálódhatnak (pl. Ford 1982; Shriberg 2001; Horváth 2010). Francia spontán beszédben átlagosan 500 ms-os kitöltött szüneteket mértek (Duez 2001), svéd beszélőknél 480 ms volt az átlag (Eklund 2004), míg Merlo és Mansur (2004) angol anyanyelvű adatközlőknél átlagosan 776 ms-ot adatolt. Magyar beszélők narratíváiban az átlag 374 ms, társalgásban pedig 334 ms volt (Bóna 2013).

A hezitációs jelenségeket többnyire fiatal felnőttek beszédprodukciójában vizsgálták, kevesebb kutatás foglalkozott a tipikus nyelvfejlődésű gyermekek, avagy az idősek kitöltött szüneteivel. Több vizsgálat adatai szerint nincs nagy különbség a fiatal és az idős felnőttek megakadásainak gyakoriságában (pl. Leeper–Culatta 1995), még a százévesek is hasonló típusú és gyakoriságú megakadásokat produkálnak, mint a náluk fiatalabb idős beszélők (Searl et al. 2002). Más kutatások ugyanakkor arra az eredményre jutottak, hogy az idős személyek több kitöltött szünetet produkálnak, mint a fiatal felnőttek (Yairi–Clifton 1972; Bortfeld et al. 2001; Menyhárt 2001; Gayraud et al. 2011); illetve hogy a kitöltött szünetek gyakoribbá és hosszabb időtartamúvá váltak az életkor előrehaladtával (Manning–Monte 1981; Pindzola 1990; áttekintést nyújt a témában Kemper 1992). A spanyol anyanyelvű, Santiagóban élő idősek beszédprodukcióiban gyakrabban fordul elő az *eh*, mint a fiatalokéban (Roggia 2012). Hasonló eredményeket kapott Kemper (1992), aki idősödők

(60–74 évesek) és idősek (75–90 évesek) beszédprodukcióit hasonlította össze, és azt találta, hogy az idősek gyakrabban hezitálnak, mint az idősödők. Az életkor hatása különösen 80 éves kor után mutatkozik meg a hezitációs jelenségekben (Furquim de Andrade–de Oliveira Martins 2010). A stressz jobban befolyásolja az idősek beszédprodukcióját, akik kevésbé folyamatosan beszélnek, mint a fiatalok (Caruso et al. 1997). Az eddigi vizsgálatok tehát többé-kevésbé ellentmondó eredményekre jutottak.

Az anyanyelv-elsajátítás során a gyermekek jól formált és hibás nyelvi mintákat egyaránt hallanak, amelyekben különböző típusú megakadások fordulnak elő. A gyermekeknél hasonló típusú megakadásokat tapasztaltak, mint a felnőtteknél (pl. MacWhinney–Osser 1977; Hudson Kam–Edwards 2008). A kitöltött szünetek már két-hároméves korban megjelenhetnek (l. Furman–Özyürek 2007; Hudson Kam–Edwards 2008). A gyermekek a szemantikai, szintaktikai, fonológiai stb. szabályokkal együtt elsajátítják a kitöltött szünetek funkcióját, formáit, használatát is (MacLachlan–Chapman 1988; Pak et al. 1996; Montes 1999; Schiro 2003; Furman–Özyürek 2007; Reuterskiöld et al. 2011). Az iskoláskort közvetlenül megelőző életkorban a gyermekek beszéde a felnőttekéhez hasonló megakadásmintázatot mutat (DeJoy–Gregory 1985). A megakadások gyakorisága folyamatosan csökken hatéves kortól felnőtt korig, ugyanakkor a különböző megakadástípusok aránya életkorfüggő sajátosságokat mutat (Ito 1986). Kilencéves török anyanyelvű gyermekek a kitöltött szünetet (*şey*) 62%-ban beszédtervezési és beszédszándék funkcióban használták, 35%-ban pedig töltelékelemként (Furman–Özyürek 2007). Számos kutatásban igazolták a kitöltött szünetek (és más szünetek) összefüggéseit a grammatikai struktúrákkal a gyermekek beszédében is (MacLachlan–Chapman 1988; Rispol–Hadley 2001; Fiestas et al. 2005; Farantouri et al. 2008). A szorongás hatása az előfordulás növekedésére gyermekeknél is igazolható volt (Levin–Silverman 1965).

A kitöltött szünetek gyakorisága a gyermekek beszédében növekszik a megnyilatkozások nyelvi komplexitásának és hosszúságának függvényében (Yaruss et al. 1999; Thordardottir–Weismer 2002), továbbá a kommunikációs feladat nehézsége is befolyásolja a megjelenésüket (Levin et al. 1965). Angol anyanyelvű háromévesek és négyévesek már megértették a kitöltött szünetek használatának alapjait, de még nem tudtak különbséget tenni közöttük (Hudson Kam–Edwards 2008). Ugyancsak angol anyanyelvű öt- és hatéves gyermekek spontán beszédében az *ah* realizációjú kitöltött szünet volt a leggyakoribb megakadástípus (Ragsdale–Sisterhen 1984). Kilencéves spanyol anyanyelvű gyermekek összes megakadásának 28%-át tette ki a kitöltött szünet (Esposito 2005). Török anyanyelvű gyermekek beszédében nem találtak szignifikáns különbséget a *şey* előfordulásában a 3,5 évesek és a 9 évesek között (Furman–Özyürek 2007). Spanyolul beszélő hatévesek narratíváiban az összes szószámhoz viszonyítva átlagosan 2,3% volt a kitöltött szünetek aránya, míg ugyanilyen korú angol gyermekeknél ez az arány 3,9%-nak adódott

(Fiestas et al. 2005). A kitöltött szünetek időtartama a gyermekek beszédében is több tényező függvénye; 4,6 évesek spontán közléseiben átlagosan 800 ms-ot mértek (MacWhinney–Osser 1977), 6-7 éves magyar anyanyelvű óvodásoknál 343 ms volt az átlag (Horváth 2009). Egy másik kutatásban 6 éves magyar gyermekek hezitálásainak átlagos időtartama 377 ms-nak adódott, 7 éveseknél 427 ms-ot, 9 éveseknél 347 ms-ot, 11 éveseknél 359 ms-ot, míg a 13 éveseknél 385 ms-ot mértek; a csoportok között azonban nem volt szignifikáns különbség (Neuberger 2013).

A kitöltött szünetek formája a magyarban is többféle, mint például *ö*, *m*, *öh*, *öm*, *öhm*, időtartamuk is különböző (Gósy 2003; Menyhárt 2001; Horváth 2010; Gósy–Bóna 2011; Bóna 2013). A legjellemzőbb kitöltött szünetek az /ø/ realizációihoz hasonló magánhangzók, beleértve a semleges magánhangzókat ([ə]). Ezek a hezitálások a fiatal felnőttek összes hezitációs jelenségeinek közel 80%-át teszik ki (Horváth 2010). Az [ø] magánhangzó nem csupán beszédhang, hanem szó is (egyes szám, első személyű személyes névmás, ill. birtokos névmás); gyakoriságuk azonban a spontán beszédben relatíve alacsony (0,045%-ban fordult elő egy közel 750 000 szót tartalmazó korpuszban). (Megjegyezzük, hogy a semleges [ə] más magánhangzóknak gyakori realizációja a spontán beszédben, vö. Beke–Szaszák 2010.)

A jelen kutatás célja az, hogy megvizsgáljuk az [ø]-féle, illetve a semleges magánhangzóként realizálódó kitöltött szünetek akusztikai fonetikai sajátosságait az életkor függvényében. Arra kerestünk választ, hogy milyen összefüggés van e hezitációs jelenség időtartama és formánsstruktúrája, valamint a beszélő életkora (gyermek, fiatal felnőtt és idős) között. Kérdéseink megválaszolásához kilencéves gyermekek, a húszas éveikben járó fiatal felnőttek és hetven évnél idősebb beszélők beszédprodukcióit elemeztük. Hipotéziseink szerint 1. a kitöltött szünetek különböző gyakorisági mintázatokat mutatnak az életkor függvényében; 2. az időtartamokban nem lesznek jelentős különbségek az életkori csoportok között; 3. a kitöltött szünetek formánsstruktúrája az életkor és a nem függvényében eltérést fog mutatni. Feltételeztük, hogy az *ö*-féle hezitálások az életkor előrehaladtával egyre jobban a semleges magánhangzó formánsstruktúrájához lesznek hasonlók. Feltevésünk szerint az időseknél gyakrabban fordulnak majd elő kitöltött szünetek, mint a gyermekeknél és a fiatal felnőtteknél.

Kísérleti személyek, anyag, módszer

A kutatásban 48 beszélő spontán beszédfelvételeit elemeztük. Az adatközlők három életkori csoportot alkottak; az elsőben 16 kilencéves gyermek, a másodikban 16 fiatal felnőtt (22–28 évesek), a harmadikban pedig 16 idős beszélő (75–90 évesek) vett részt. Mindegyik csoportban az adatközlők fele nő, fele férfi volt. Valamennyien egynyelvű, magyar anyanyelvű, köznyelvi beszélők. A gyermekeket véletlenszerűen választottuk ki egy átlagos általános iskola harmadik osztályából. Középosztálybeli családból származtak, ti-

pikus beszédfejlődésűek voltak, és egyiküknek sem volt hallásproblémája. A felnőtteket az életkoruk alapján véletlenszerűen választottuk ki a BEA magyar beszélt nyelvi adatbázisból (Gósy 2012). Hallásuk és beszédprodukciójuk az életkoruknak megfelelő volt.

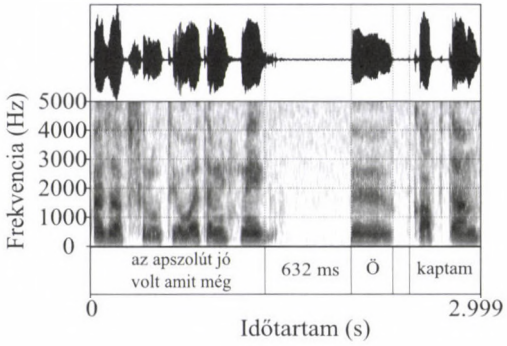
A gyermekek a felvételek során a családjukról, a tanulmányaikról, illetve a szabadidejükről beszéltek (a felvételkészítő csak nagyon ritkán tett fel kérdést, amikor úgy tűnt, hogy a gyermek nem tudja folytatni az elbeszélést). A beszédprodukciókat digitálisan rögzítettük ugyanolyan típusú mikrofonnal, mint amilyennel a felnőttek beszédét vették fel. A beszédfelvételek az iskolában, csendesített szobában, délelőtt készültek. A gyermekek átlagosan hét percet beszéltek, ebből az elemzéshez gyermekenként öt percet választottunk ki a spontán beszédük közepéről.

A BEA adatbázis (több más beszédtípus mellett) olyan narratívákat tartalmaz, amelyekben a beszélők a családjukról, a tanulmányaikról, a jelenlegi vagy korábbi munkájukról, szabadidejükről beszélnek. A vizsgálathoz mind a fiatalok, mind az idősek narratíváinak közepéből beszélőnként átlagosan 8,5 percnyi hanganyagot választottunk ki. A BEA hangfelvételeit csendesített szobában, közvetlenül számítógépre rögzítették, a felvételkészítő minden esetben ugyanaz a személy volt. Összesen 216 percnyi (mintegy 3,6 órányi) spontán beszédet elemeztünk a három korcsoportban.

A kitöltött szüneteket a definíciónak megfelelően azonosítottuk a beszédmintákban. Minden beszédmintát két szerző egymástól függetlenül, manuálisan annotált a Praat szoftverrel (Boersma–Weenink 2005), majd a másik két szerző ellenőrizte az annotálásokat. A hanganyagokban felcímkéztük a beszédszakaszokat, a néma szüneteket, illetve a hezitálásokat, amelyeket a nekik megfelelő beszédhangokkal jelöltünk (1. ábra). A kitöltött szünetek azonosításában gyakorlatilag nem volt különbség a szerzők között (az egyezés 98%-os volt).

A gyermekek és az idősek többféle hangot és hangkapcsolatot produkáltak kitöltött szünetként, mint a fiatalok (gyermekek: 16 típus, idősek: 14 típus, fiatalok: 7 típus, például *m*, *öm*, *öh*, *ömh*, *möh*). Anyagunkban az *ö*-féle hezitálások az összesnek a 71,5%-át tették ki. (A továbbiakban csak ezekkel foglalkoztunk.) Összesen 1054 (gyermekeknél 249, fiataloknál 523, időseknél 282) *ö*-féle kitöltött szünet időtartamát és formánsszerkezetét elemeztük. A hezitálások megjelentek önállóan két szó között, avagy néma szünettel kombinálva: két néma szünet között, vagy egy néma szünetet megelőzve vagy követve, illetve hozzátapadva a szomszédos szóhoz. Elemeztük az *ö*-féle hezitálások gyakoriságát, időtartamát és az első két formánsukat. A kitöltött szünetek időtartamát a zöngeindulástól a zönge lecsengéséig mértük. A formánsokat a hezitálás időtartamának közepén adatoltuk. Mind az időtartamokat, mind a formánsértékeket egy erre a célra írt script segítségével automatikusan nyertük ki a hangfelvételekből. Ezután az időtartamokat és a formán-

sokat két szerző manuálisan is ellenőrizte. A kitöltött szünetek funkcióit nem vizsgáltuk.



1. ábra
Spektrogram és oszcillogram: annotált beszédminta *ö*-típusú kitöltött szünettel

Tekintettel arra, hogy a kitöltött szünetek időtartama nem követte a normál eloszlást, ezért 12 különböző, nevezetes eloszlástípust teszteltünk az adatokon. A cél annak a megállapítása volt, hogy a mért adatok mely eloszláshoz hasonlítanak legjobban. Az összehasonlítás mértékét a NLogL (Negative of the log likelihood), a BIC (Bayesian information criterion), az AIC (Akaike information criterion) és az AICc (AIC with a correction for finite sample sizes) tesztekkel mértük. Az eredmények szerint az időadatok a gamma eloszlást követik a legnagyobb mértékben, ezért ezeken nem-parametrikus teszteket alkalmaztuk. További statisztikai elemzést végeztünk az SPSS 19.0 verziójával (ismétléses ANOVA-t, MANOVA-t, regresszióelemzést, diszkriminanciaelemzést, független mintás *t*-próbát és háromdimenziós euklideszi távolságmérést alkalmaztunk, vö. 2. ábra).

Eredmények

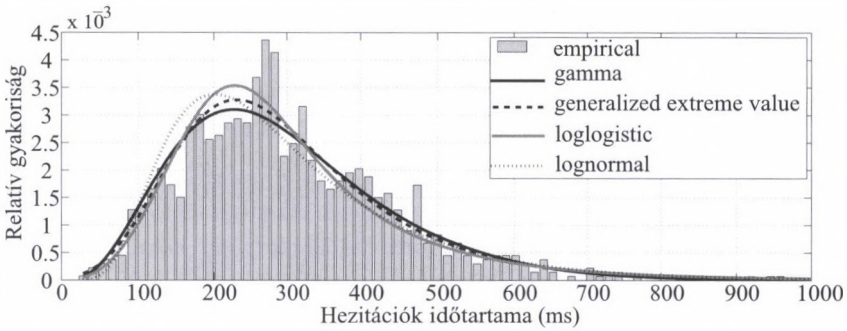
Három példát mutatunk be a kitöltött szünetekre egy-egy közlésrészletben (felnőtt adatközlőktől; a hezitálást félkövér ***öö*** betűk jelzik).

(1) *nekem mondjuk **öö** én mondjuk fel akarok használni*
(2) *hogy a tanyá **öö** a tanároknak milyen könnyű*
(3) ***öö** én úgy általában nem nagyon szoktam inni*

Gyakoriság

Az *ö*-féle kitöltött szünetek 78,5%-ban fordultak elő a fiatal felnőtteknél, 67,6%-ban az időseknél és 68,8%-ban a gyermekeknél (az összes hezitálást tekintve 100%-nak). Az arányok azt mutatják, hogy a gyermekek és az idősek beszédprodukciója hasonlít egymásra, a fiataloké mindkettőtől különbö-

zik, utóbbiak jobban preferálják az *ö*-féle hezitálásokat (szemben az egyéb hangokat tartalmazó kitöltött szünetekkel). A fiatalok beszédében percenként 2,97 előfordulást, a gyermekeknél percenként 2,76-ot, az időseknél pedig percenként 2,60 hezitálást találtunk. A kilencéves lányok és fiúk között nem volt eltérés az előfordulásban. A fiatal nők valamivel többször hezitáltak ugyanazon időegység alatt (2,99 hezitálás/perc), mint a fiatal férfiak (2,49 hezitálás/perc). Különösen nagy volt azonban a különbség az idős nők (3,50 hezitálás/perc) és az idős férfiak (1,60 hezitálás/perc) között. Ez azt jelenti, hogy az idős női beszélők 1 órányi beszédidő alatt mintegy harminccal több kitöltött szünetet produkáltak, mint az idős férfiak.



2. ábra
Különböző eloszlástípusok tesztelési eredménye a mért időtartamokon

Említettük, hogy a hezitálások előfordulhatnak néma szünetek között, illetve szavakhoz tapadva, így sajátos koartikulációt alkotnak. A gyermekeknél a szünetek között előforduló hezitálások 31,8%-ot tettek ki, a fiatal felnőttek beszédében 16,6%-ot, az időseknél pedig 18,8%-ot. Ebből következően a szó elejéhez vagy végéhez tapadók aránya 68,2% volt a gyermekeknél, 83,4% a fiataloknál és 81,2% az időseknél.

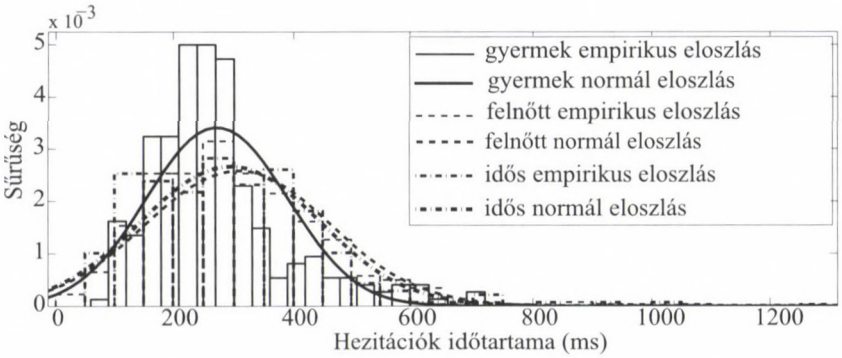
Időtartam

A kitöltött szünetek időtartama tág tartományban szóródott mindhárom életkori csoportban (1. táblázat).

1. táblázat: A kitöltött szünetek temporális jellemzői az életkori csoportokban

Korcsoport	Átlagos időtartam (ms)	Átlagos eltérés (ms)	Minimum (ms)	Maximum (ms)
Gyermek	269	110	101	633
Felnőtt	340	163	42	1290
Idős	322	163	72	1004
Összes	318	154	42	1290

A fiatal felnőttek produkálták a leghosszabb kitöltött szüneteket, míg a gyermekek a legrövidebbeket. Az idősek hezitációinak időtartamai a másik két életkori csoport közé estek (3. ábra), a különbség szignifikáns (Kruskal–Wallis-próba: χ^2 (2) = 10,813, p = 0,004).



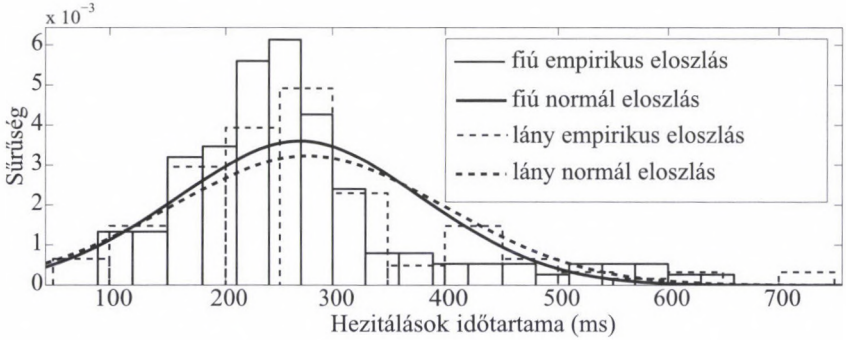
3. ábra
A kitöltött szünetek időtartamának és relatív gyakoriságának összefüggései az életkor függvényében

Mindhárom életkori csoportban elemeztük az időtartamokat a nemek függvényében. A kilencéves fiúk kitöltött szüneteinek átlaga 270 ms volt (átl. elt.: 110 ms), a lányoké 276 ms (átl. elt.: 123 ms), a különbség nem szignifikáns (4. ábra). A fiatal felnőttek és az idősek adatai ugyanakkor szignifikáns különbséget mutattak a nők és a férfiak hezitálásai között (Mann–Whitney-próba a fiataloknál: Z = –4,947, p < 0,001; az időseknél: Z = –2,109; p = 0,035), lásd az 5. és 6. ábrát. Noha a fiatal nők több kitöltött szünetet produkáltak, mint a fiatal férfiak, az utóbbiak hezitálásai hosszabbak voltak a nőknél mérteknél (a férfiak átlaga 340 ms, a szórás 163 ms; a nőké 275 ms, a szórás 137 ms). Ugyanezt találtuk az idős nők és az idős férfiak beszédében is, csak az idősek hosszabb kitöltött szüneteket tartottak, mint a fiatalok (a férfiak átlaga 322 ms, a szórás 163 ms; a nők átlaga 280 ms, a szórás 139 ms).

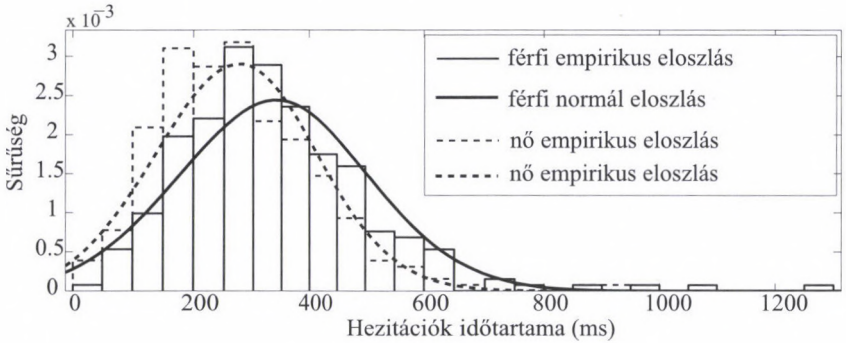
A hezitálások időtartama a fiúgyermekeknél szignifikánsan rövidebb volt, mint a fiatal és az idős férfiaknál (Kruskall–Wallis-próba: χ^2 (2) = 24,5453, p < 0,001), lásd a 7. ábrát is.

Az átlag a fiúknál 270 ms (átl. elt.: 110 ms), a fiatal férfiaknál 340 ms (átl. elt.: 163 ms), az idős férfiaknál 322 ms (átl. elt.: 163 ms) volt (az utóbbiak között nem volt szignifikáns különbség). Mindkét felnőtt csoport szignifikánsan különbözött a gyermekektől (Mann–Whitney-próba: gyermekek és fialatok között: Z = –4,982; p < 0,001; gyermekek és idősek között Z = –2,941; p = 0,003). Az idős nők tendenciaszerűen valamivel hosszabb hezitá-

lásokat produkáltak, mint a fiatal nők és a 9 éves lányok; a különbség azonban nem volt szignifikáns az életkor függvényében. A kitöltött szünetek átlagos időtartama a kilencéves lányoknál 276 ms-nak (átl. elt.: 123 ms), a fiatal nőknél 275 ms-nak (átl. elt.: 137 ms), az idős nőknél 280 ms-nak (átl. elt.: 139 ms) adódott (8. ábra).



4. ábra
A kitöltött szünetek időtartama a gyermekeknél a nem függvényében

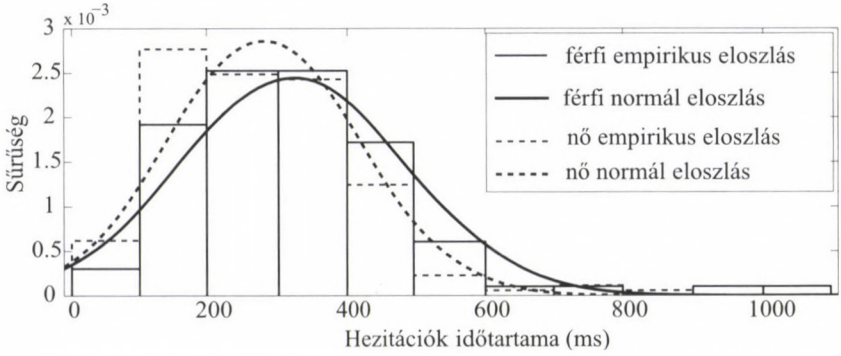


5. ábra
A kitöltött szünetek időtartama a fiatal felnőtteknél a nem függvényében

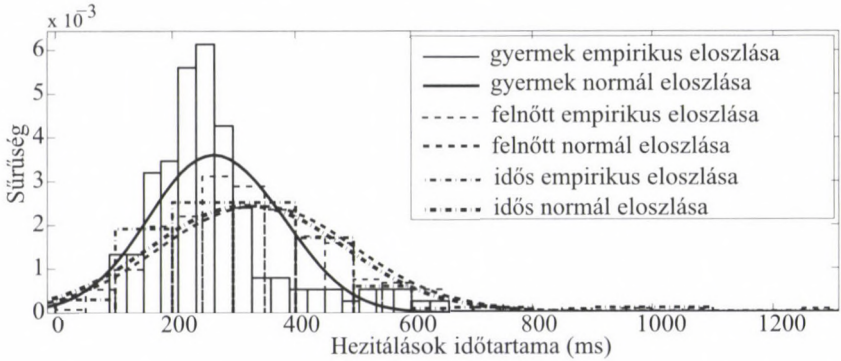
A formánsszerkezet

A formánsszerkezetet értelemszerűen az életkor és a nem függvényében elemeztük (2. táblázat). Mind az első, mind a második formáns életkorfüggő sajátosságokat mutatott. Az F_1 értékei az életkor előrehaladtával csökkentek; az F_2-értékek a gyermekeknél és az időseknél magasabbak voltak, mint a fiatalokéi.

Az életkor és a nem hatását MANOVA-val elemeztük. Mindkét tényező hatással van mind az első, mind a második formánsra (az életkor hatása az F_1-re: $F(2, 1044)= 277{,}985$, $p < 0{,}001$; a második formánsra: $F(2,1044) = 141{,}686$, $p < 0{,}001$; a nem hatása az F_1-re: $F(1, 1044) = 107{,}591$, $p < 0{,}001$; a második formánsra: $F(1,1044) = 235{,}131$, $p < 0{,}001$). A két tényező együttes hatása csak a második formáns esetében volt szignifikáns [$F(2, 1044) = 9{,}571$, $p < 0{,}001$], lásd a 9. ábrát is.



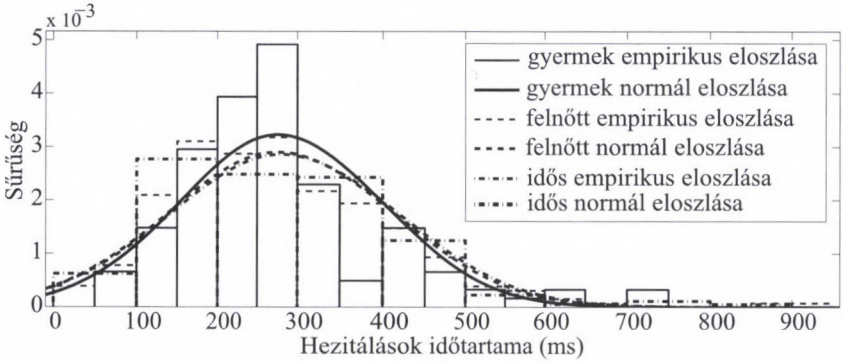
6. ábra
A kitöltött szünetek időtartama az időseknél a nem függvényében



7. ábra
A kitöltött szünetek időtartama a férfi beszélőknél a három életkori csoportban

A formánsokat az életkor függvényében a Tukey post hoc teszttel is megvizsgáltuk, amely szerint mindhárom életkori csoport szignifikánsan különbözött egymástól ($p = 0{,}001$ minden esetben). A fiatal és az idős felnőtteknél

az elvárásainknak megfelelően szignifikáns különbséget találtunk mindkét formánsban a nők és a férfiak között (egytényezős ANOVA a fiatal felnőttek első formánsainál: $F(1, 519) = 99{,}606$, $p < 0{,}001$; a második formánsoknál: $F(1{,}519) = 347{,}479$, $p < 0{,}001$); az idősek első formánsainál: $F(1, 274) = 35{,}126$, $p < 0{,}001$; a második formánsoknál: $F(1{,}274) = 40{,}331$, $p < 0{,}001$). Noha a 9 éves lányok és fiúk első és második formánsainak értékei nem különböztek olyan mértékben, mint a felnőtt nőké és férfiaké, a különbség náluk is szignifikáns volt (az első formánsokra: $F(1, 245)=11{,}009$, $p < 0{,}001$; a második formánsokra: $F(1{,}245) = 29{,}269$, $p < 0{,}001$). A parciális η^2 szerint az életkor 34,9%-ban magyarázza az első formánsok különbségét, míg a második formánsokét csak 21,4%-ban. A beszélők neme elhanyagolhatónak tűnik az első formánsok magyarázatában (9,4%), kissé nagyobb mértékű a második formánsoknál (18,5%). Az életkor tehát sokkal meghatározóbb a formánsszerkezet alakulásában, mint a nem.



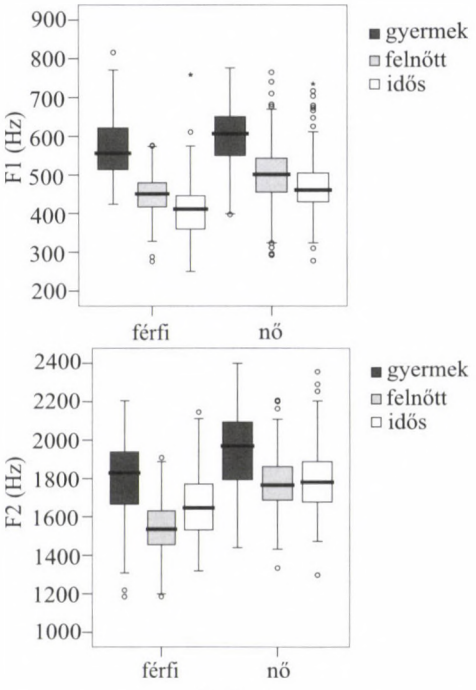
8. ábra
A kitöltött szünetek időtartama a női beszélőknél a három csoportban

2. táblázat: Az *ö*-féle hezitálás formánsai az életkor és a nem függvényében

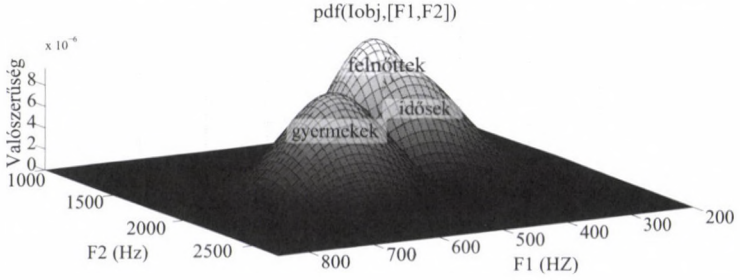
Életkori csoportok	F_1 (Hz)				F_2 (Hz)			
	Nők		Férfiak		Nők		Férfiak	
	Átlag	Átl. elt.	Átlag	Átl. elt.	Átlag	Átl. elt.	Átlag	Átl. elt.
Gyermekek	600	77	568	74	1946	212	1803	203
Fiatalok	503	74	447	53	1773	142	1543	140
Idősek	474	75	416	85	1793	165	1658	180

Az első és a második formánsokat Gauss-görbével modelleztük, hogy megfigyelhessük, illetve szemléltessük a három életkori csoport közötti különbséget. A modellben a csoportok közötti távolságot euklideszi távolsággal

jellemeztük, amelyet az első két formánsérték eloszlásának középértékétől számítottunk. E mérés szerint az idős és a fiatal felnőttek hezitálásainak megvalósulásai állnak egymáshoz közel, míg a gyermekekéi mindkét csoporttól jelentősen távol esnek (10. ábra).

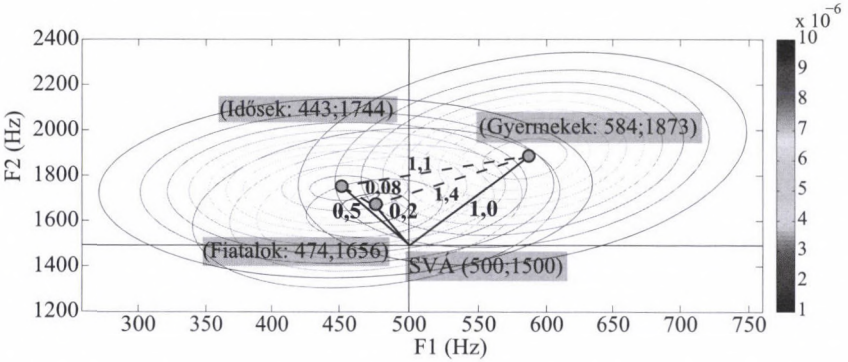


9. ábra
Az első két formáns az életkor és a nem függvényében (medián és szórás)



10. ábra
A három életkori csoport formánsai Gauss-görbével modellezve

Választ kerestünk arra, hogy vajon a három életkori csoport beszélői közül kiknek a hezitálásai realizálódnak a legközelebb a semleges ejtéshez. A semleges magánhangzó (l. Pickett 1980) első formánsának értéke 500 Hz, a másodiké 1500 Hz (a harmadiké pedig 2500 Hz). Adatainknak a semleges magánhangzótól való távolságát négyzetes euklideszi távolsággal mértük az egyes eloszlások középértékétől. Az elemzésből kiderült, hogy a fenti értékekkel jellemzett semleges magánhangzóhoz a fiatal felnőttek hezitálásai állnak a legközelebb, majd az idősekéi. A gyermekek hezitálásai viszont nagymértékben eltérnek mindkettőtől (11. ábra).



11. ábra
A három életkori csoport *ö*-féle magánhangzóinak négyzetes euklideszi távolsága a (klasszikus) semleges magánhangzótól

Következtetések

Kutatásunk célja az volt, hogy megvizsgáljuk a magyar *ö*-típusú kitöltött szünetek gyakoriságát, időtartamát és formánsszerkezetét az életkor függvényében. Eredményeink megerősítik azokat a korábbi szakirodalmi megállapításokat, hogy a kitöltött szünetek gyakoriak a spontán beszédben függetlenül a beszélő életkorától (Gósy 2005; Horváth 2010; Bóna 2013). Adataink igazolták, hogy a magyarra az *ö*-féle hezitálások a legjellemzőbbek, és életkortól függően nagyobb mértékben tapadnak a megelőző vagy a követő szóhoz szemben a két néma szünet közötti előfordulással. A kilencévesek kitöltött szünetei voltak a legrövidebbek, a fiatal felnőttekéi pedig a leghosszabbak. A hezitálások első és második formánsainak értékei szignifikánsan különböztek az életkor és a nem tekintetében. Bár az időtartamokban szignifikáns különbség volt az életkori csoportok között, a részletes elemzés szerint nem volt matematikailag alátámasztott eltérés a kilencéves fiúk és a felnőtt férfiak, illetve a kilencéves lányok és a felnőtt nők között. Az a hipotézisünk, amely szerint a kitöltött szünetek időtartama nem mutat életkorfüggő különbségeket, csak részben igazolódott.

Másik hipotézisünk az volt, hogy eltérő mintázatot találunk a kitöltött szünetek előfordulásában az életkor függvényében (vö. Yairi–Clifton 1972; Roggia 2012). A nemtől függetlenül az idősek valamivel ritkábban hezitáltak, mint a fiatalok, de az idős nők hezitálásainak aránya meghaladta valamennyi adatközlői csoportét. A kilencéves lányok és fiúk, illetve a fiatal nők és férfiak között nem találtunk jelentős különbséget a hezitálások gyakoriságában. Ez feltehetően azzal magyarázható, hogy a kísérletünkben részt vevő fiatal felnőttek komplexebb grammatikai szerkezeteket produkáltak, mint a másik két életkori csoport beszélői. Ennek igazolásához azonban a megnyilatkozások morfológiai és szintaktikai struktúrájának elemzése szükséges (lásd Rochester–Gill 1973). Lehetséges magyarázat az is, hogy a gyermekek és az idősek kevésbé kontrollálják a közléseik tartalmát, mint a fiatalok, és ez is eredményezhet gyakoribb hezitálást az utóbbiaknál. Ennek megerősítése is további, funkcionális és tartalmi elemzéseket igényel. Angol anyanyelvű, 10-11 éves lányok beszédében mintegy kétszer annyi kitöltött szünet fordult elő, mint a fiúkéban (Farantouri et al. 2008). Magyar anyanyelvű 9 évesek spontán beszédében nem volt szignifikáns különbség a kitöltött szünetek előfordulásában a nemek szerint (Horváth 2013). Hasonlóan nem találtak matematikailag igazolható eltérést a kitöltött szünetek gyakoriságában a nemek függvényében 7 éveseknél, 11 és 13 éveseknél sem (Neuberger 2013).

A szakirodalom egy része szerint az idősek ugyanolyan folyamatosan, avagy akár folyamatosabban beszélnek, mint a fiatalok; más része szerint a kitöltött szünetek gyakorisága az idős korra nő (Leeper–Culatta 1995; Pindzola 1990). Saját adataink ellentmondanak azoknak a megállapításoknak, amelyek az időseknek a fiatalokénál gyakoribb és hosszabb hezitálását adatolták (l. Yairi–Clifton 1972; Manning–Monte 1981; Pindzola 1990; Roggia 2012). Eredményeink szerint nagy átfedés van a fiatal és az idős csoport adatai között; anyagunkban az idősek kitöltött szünetei rövidebbek voltak, mint a fiatalokéi. Azt feltételezzük, hogy ennek elsősorban nyelvspecifikus okai lehetnek: a magyar *ö*-féle hezitálás stabil formája és gyakori koartikulációs megvalósulása különbözik a szakirodalomban elemzett más nyelvek kitöltött szüneteitől. Az a tény, hogy a vizsgált beszélők – életkortól függetlenül – legnagyobb mértékben a kiejtett szóval együtt artikulálják a kitöltött szünetet, egy sajátos stratégiát feltételez. Igyekeznek minél kevésbé feltűnővé tenni a hezitálásokat. Ezért azt a stratégiát alkalmazzák, hogy nem két szünet közé iktatják őket, hanem mintegy a lexéma részeként produkálják. A szó időtartamának ilyenfajta megnövelése nyújt időt a rejtett beszédtervezési folyamatokra, kontrollálásra, javításra. Az *ö*-típusú kitöltött szüneteknek ez a megvalósulási formája kevésbé feltűnő, így produkálásuk észrevétlenebb marad (l. a stigmatizáltságról Horváth 2009).

Kutatásunkban igazoltuk, hogy az életkor és a nem meghatározó tényezők a magyar *ö*-féle hezitálások első és második formánsának alakulásában. Ezek az eredmények alátámasztják a harmadik hipotézisünket, ugyanakkor részben

ellentmondanak a korábbi szakirodalomnak. A gyermekek formánsai magasabbak, mint a felnőttekéi (l. Busby et al. 1995; a magyarra Bóna–Imre 2010) a különböző formájú és méretű beszédszervek és a fejlődésben lévő hangszalagok következtében (Sorenson 1989; Baken–Orlikoff 2000). Adatainkban a gyermekek formánsai szignifikánsan különböztek mindkét felnőtt csoport formánsaitól. Az idősek esetében a módosult vagy megváltozott anatómiai felépítés, fiziológiai működés, illetve az eltérő beszédtervezési folyamatok és artikulációs programozás lehet magyarázat a megváltozott artikulációs gesztusokra (Caruso 1997; Hooper–Cralidis 2009). Mindennek ellenére a vizsgálatunkban részt vett idősek formánsai nagy átfedést mutattak a fiatalok formánsaival, és nem is volt szignifikáns különbség közöttük. Egy amerikai angol anyanyelvűekkel végzett kutatásban azt találták, hogy az F_1 szignifikánsan eltér az idősek és a fiatalok ejtésében minden magánhangzónál; míg az F_2-ben a hátul képzett magánhangzók esetében adódott csak szignifikáns különbség (Watson–Munson 2007). Egy másik, szintén amerikai angol anyanyelvűekkel végzett vizsgálatban csak az első formáns értékeiben találtak szignifikáns eltérést, de az egyes paraméterek erősen függtek a vizsgált beszédhangok minőségétől is (Torre–Barlow 2009). A magyarban fiatal és idős (70 év feletti) női beszélők magánhangzóit vetették össze, az eredmények szerint az idősek szűkebb frekvenciatartományban, illetőleg valamivel hátrébb képezik a magánhangzókat, mint a fiatalok (Bóna 2009; Bóna 2012). A kitöltött szünet magánhangzójának formánsszerkezete azonban nem mutatja a fenti eltéréseket. Ez adódhat abból, hogy az izoláltan ejtett hang artikulációs jellemzői stabilabbak a beszédhangsorozatokban ejtettekénél, azaz jobban megőrződnek az életkor előrehaladtával is. Magyarázat lehet továbbá az is, hogy egy [ø] vagy egy svászerű magánhangzó képzése sokkal egyszerűbb artikulációs gesztusokat igényel, mint más magánhangzóké, mivel a szájüreg elülső vagy középső részén képződik, középső nyelvállással és kevésbé kerekített ajkakkal. Ráadásul az [ø]–[ə] kontinuum különböző realizációi közötti eltérések nem észlelhetők (illetve nem azonosíthatók) a magyar hallgatók számára, mivel a svá nem része a magyar magánhangzórendszernek.

Feltételeztük, hogy az életkor előrehaladtával a kitöltött szünet a semleges magánhangzóra jellemző formánsfrekvenciák irányába fog eltolódni. Ez részben igazolódott: a felnőttek valóban a neutrális magánhangzóhoz közeli artikulációs gesztusokat hoztak létre a kitöltött szünetek produkciója során, a gyermekek ejtésében azonban az /ø/ realizációi voltak jellemzők.

Ellentmondó megállapítások olvashatók a szakirodalomban a gyermekek formánsaira vonatkozóan. Lee és munkatársai (1999) kutatásai szerint nincs eltérés a nemek között a formánsértékekben 15 éves korig. Busby és Plant (1995) azt találták, hogy a lányok formánsai magasabbak, mint a fiúkéi (a vizsgált gyermekek életkora 5 és 11 év közé esett). Vorperian és Kent (2007) különbségeket talált a nemek szerint már a 4 évesek formánsfrekvenciáiban, az eltérések 8 éves korban tovább növekedtek. Néhány magánhangzó for-

mánsai szignifikáns különbséget mutattak 6-7 éves magyar lányok és fiúk között, de az [ø] formánsaiban nem találtak eltérést a nemek függvényében (Deme 2012). Mivel a gyermekek beszédszervei, beleértve a hangszalagokat is, egyformának tekinthetők 9 éves korig (Simpson 2009), ezért nem vártunk nagy eltéréseket a két nem között a jelen kutatásban részt vevő gyermekek hezitálásainak formánsaiban. Noha a szupraglottális felépítés, illetve a toldalékcső konfigurációja (különösen a vokális traktus hossza és a szájnyílás mérete) meghatározók a formánsfrekvenciák alakulásában, más tényezők, mint a tanult viselkedések és az adott magánhangzó-minőség az életkortól függetlenül is mutathatnak különbségeket a nemek szerint (Busby–Plant 1995; Diehl et al. 1996; Huber et al. 1999; Simpson 2009).

Anyagunkban a nem hatása a formánsértékekre mindegyik életkori csoportban szignifikáns különbségeket mutatott. A nők formánsértékei általában magasabbak, mint a férfiakéi (vö. Bachorowski–Owren 1999); a magánhangzótér is nagyobb a nőknél, mint a férfiaknál (Henton 1995; Simpson 2009; Igeta–Arai 2011). Az első formánsok szórása a 9 éves lányoknál valamivel nagyobb, mint a fiúknál (l. Busby–Plant 1995), és ugyanez figyelhető meg az idős nők és férfiak esetében is. A második formánsok a 9 éves lányoknál ugyancsak szélesebb tartományban szóródnak, mint a fiúknál, az idős beszélők esetében azonban a nőknél szűkebb a szórástartomány a férfiakéhoz képest. Az F_2-k magasabb értékei az idős férfiaknál az ajakkerekítés mértékének csökkenésére is utalhatnak. Ha mindkét formánst tekintetbe vesszük, akkor a fiatal nőknél szélesebb a szórástartomány, szemben a férfiakéval, míg az idősek esetében a férfiaknál nagyobb a szórás. Ezek az eredmények a beszédszervek működésének időskori megváltozására és a kevésbé kontrollált artikulációs gesztusokra egyaránt visszavezethetők (Torre–Barlow 2009; Bóna 2009).

A kitöltött szünetek gyakoriságát és időtartamát illetően eredményeink tehát részben alátámasztják a szakirodalomban olvasható állításokat, némely tekintetben azonban azoknak ellentmondó eredményekre jutottunk. A különbségeket elsősorban a magyar *ö*-féle hezitálásnak a többi nyelvtől eltérő hangzásával és alkalmazási stratégiáival magyarázhatjuk. A kitöltött szünetek univerzális és természetes jelenségei a spontán beszédnek, artikulációs, illetve akusztikai sajátosságaik jellegzetes nyelvspecifikus mintázatot mutatnak.

Irodalom

Arnold, Jennifer E. – Fagnano, Maria – Tanenhaus, Michael K. 2003. Disfluencies signal theee, um, new information. *Journal of Psycholinguistic Research* 32. 25–36.

Bachorowski, Jo-Anne – Owren, Michael J. 1999. Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech. *The Journal of the Acoustical Society of America* 106. 1054–1063.

Bailey, Karl G. B. – Ferreira, Fernanda 2003. Disfluencies influence syntactic parsing. *Journal of Memory and Language* 49. 183–200.

Baken, Ronald – Orlikoff, Robert F. 2000. *Clinical measurement of speech and voice*. Singular Publishing Group, San Diego.

Beke, András – Szaszák, György 2010. Automatic recognition of schwa variants in spontaneous Hungarian speech. *Acta Linguistica Hungarica* 57/2–3. 329–353.

Berger, Charles R. – Jordan, Jerry M. 1992. Planning sources, planning difficulty and verbal fluency. *Communication Monographs* 59. 130–149.

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

Bóna Judit 2012. A rövid-hosszú magánhangzók realizációi idősek spontán beszédében. *Beszédkutatás 2012*. 43–57.

Bóna Judit 2013. A beszédszünetek fonetikai sajátosságai a beszédtípus függvényében. *Beszédkutatás 2013*. 60–75.

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.

Bortfeld, Heather – Leon, Silvia D. – Bloom, Jonathan E. – Schober, Michael F. – Brennan, Susan E. 2001. Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech* 44. 123–147.

Busby, P. A. – Plant, G. L. 1995. Formant frequency values of vowels produced by preadolescent boys and girls. *The Journal of the Acoustical Society of America* 97. 2603–2606.

Blass, Thomas – Siegman, Aron W. 1975. A psycholinguistic comparison of speech, dictation and writing. *Language and Speech* 18. 20–34.

Boersma, Paul – David Weenink 2009. *Praat: doing phonetics by computer (Version 5.0.1)*. http://www.fon.hum.uva.nl/praat/download_win.html.

Brennan, Susan E. – Schober, Michael F. 2001. How listeners compensate for disfluencies in spontaneous speech. *Journal of Memory and Language* 44. 274–296.

Brennan, Susan E. – Williams, Maurice 1995. The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language* 34. 383–398.

Caruso, Anthony J. – McClowry, Troy M. – Ludo, Max 1997. Age-related effects on speech fluency. *Seminars in Speech and Language* 18/2. 171–180.

Christenfeld, Nicholas 1995. Does it hurt to say um? *Journal of Nonverbal Behavior* 19. 171–186.

Christenfeld, Nicholas 1996. Effects of a metronome on the filled pauses of fluent speakers. *Journal of Speech and Hearing Research* 39. 1232–1238.

Christenfeld, Nicholas – Schachter, Stanley – Bilous, Frances 1991. Filled pauses and gestures: It's not coincidence. *Journal of Psycholinguistic Research* 20. 1–10.

Christenfeld, Nicholas – Creager, Beth 1995. Anxiety, alcohol, aphasia, and urns. *Journal of Personality and Social Psychology* 70. 451–460.

Clark, Herbert H. 1994. Managing problems in speaking. *Speech Communication* 15. 243–250.

Clark, Herbert H. – Wasow, Thomas 1998. Repeating words in spontaneous speech. *Cognitive Psychology* 37. 201–242.

Clark, Herbert H. – Fox Tree, Jean E. 2002. Using *uh* and *um* in spontaneous speaking. *Cognition* 84. 73–111.

Corley, Martin – Hartsuiker, Robert J. 2003. Hesitation in speech can... um... help a listener understand. In Alterman, Richard – Kirsh, David (eds.): *Proceedings of the 25th meeting of the Cognitive Science Society*. Mahwah NJ, Erlbaum. 276–281.

Corley, Martin – Stewart, Oliver W. 2008. Hesitation disfluencies in spontaneous speech: The meaning of *um*. *Language and Linguistics Compass* 4. 589–602.

Corley, Martin – MacGregor, Lucy J. – Donaldson, David I. 2007. It's the way that you, er, say it: Hesitations in speech affect language comprehension. *Cognition* 105. 658–668.

DeJoy, Daniel A. – Gregory, Hugo H. 1985. The relationship between age and frequency of disfluency in preschool children. *Journal of Fluency Disorders* 10. 107–122.

Deme Andrea 2012. Óvodások magánhangzóinak akusztikai jellemzői. In Markó Alexandra (szerk.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngekezdési időig*. ELTE BTK–MTA NYTI, Budapest. 77–99.

Diehl, Randy L. – Lindblom, Björn – Hoemeke, Kathryn A. – Fahey, Richard P. 1996. On explaining certain male-female differences in the phonetic realization of vowel categories. *Journal of Phonetics* 24. 187–208.

Duez, Danielle 2001. Acoustico-phonetic characteristics of filled pauses in spontaneous French speech: Preliminary results. In *DISS'01*. 41–44. http://www.isca-speech.org/archive_open/archive_papers/diss_01/dis1_041.pdf.

Eklund, Robert 2004. Disfluency in Swedish human–human and human–machine travel booking dialogues. *Linköping Studies in Science and Technology*. Dissertation No. 882. Unitryck, Linköping.

Esposito, Anna 2005. Children's organization of discourse structure through pausing means. Nonlinear analyses and algorithms for speech processing. *Lecture Notes in Computer Science* 3817. 108–115.

Esposito, Anna – Marinaro, Maria – Palombo, Giulia 2004. Children speech pauses as markers of different discourse structures and utterance information content. In: *Proceedings of From Sound to Sense Conference*. MIT, Cambridge, MA. 139–144.

Farantouri, Vassiliki – Potamianos, Alexandros – Narayanan, Shrikanth 2008. Linguistic Analysis of Spontaneous Children Speech. *Proceedings of the Workshop on Child, Computer and Interaction*. sail.usc.edu/aigaion2/index.php/attachments/

Ferreira, Fernanda – Lau, Ellen F. – Bailey, Karl G. D. 2004. Disfluencies, parsing, and tree-adjoining grammars. *Cognitive Science* 28. 721–749.

Fiestas, Christine E. – Bedore, Lisa M. – Peña, Elizabeth D. – Nagy, Vanessa J. 2005. Use of mazes in the narrative language samples of bilingual and monolingual 4- to 7-year old children. In Cohen, James – McAlister, Kara T. – Rolstad, Kellie – MacSwan, Jeff (eds.): *Proceedings of the 4th International Symposium on Bilingualism*. Cascadilla Press, Somerville. 730–740. http://www.lingref.com/isb/4/057ISB4.PDF.

Finlayson, Ian R. – Corley, Martin 2012. Disfluency in dialogue: An intentional signal from the speaker? *Psychonomic Bulletin & Review* 19. 921–928.

Ford, Marlyn. 1982. Sentence planning units: Implications for the speaker's representation of meaningful relations underlying sentences. In Bresnan, Joan (ed.): *The mental representation of grammatical relations*. MIT Press, Cambridge, MA. 798–827.

Fox Tree, Jean E. 2001. Listeners' uses of um and uh in speech comprehension. *Journal of Memory and Language* 29. 320–326.

Fox Tree, Jean E. 2002. Interpreting pauses and ums at turn exchanges. *Discourse Processes* 34. 37–55.

Furman, Reyhan – Özyürek, Asli 2007. Development of interactional discourse markers: Insights from Turkish children's and adults' oral narratives. *Journal of Pragmatics* 39. 1742–1757.

Furquim de Andrade, Claudia Regina – de Oliveira Martins, Vanessa 2010. Speech fluency variation in elderly. *Pró-Fono Revista de Atualização Científica* 22. 13–18.

Gayraud, Frederique – Lee, Hye-Ran – Barkat-Defradas, Melissa 2011. Syntactic and lexical context of pauses and hesitations in the discourse of Alzheimer patients and healthy elderly subjects. *Journal of Clinical Linguistics and Phonetics* 25. 198–209.

Gósy Mária 2003. A spontán beszédben előforduló megakadásjelenségek gyakorisága és összefüggései. *Magyar Nyelvőr* 127. 257–277.

Gósy Mária 2012. Multifunkcionális beszélt nyelvi adatbázis – BEA. *Általános Nyelvészeti Tanulmányok* XXIV. 329–349.

Gósy Mária – Bóna Judit 2011. Beszédfolyamatok monitorozása. *Magyar Nyelvőr* 135. 393–414.

Hartsuiker, Robert J. – Kolk, Herman H. J. 2001. Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive Psychology* 42. 113–157.

Hartsuiker, Robert J. – Corley, Martin – Lickley, Robin – Russell, Melanie. 2003. Perception of disfluency in people who stutter and people who do not stutter: Results from magnitude estimation. In Robert Eklund (ed.): *Proceedings of Disfluency in Spontaneous Speech Workshop.* Gothenburg Papers in Theoretical Linguistics 90. 35–38.

Henton, Caroline 1995. Cross-language variation in the vowels of female and male speakers. In Elenius, Kjell – Branderud, Peter (eds.): *Proceedings of the XIIIth International Congress of Phonetic Sciences.* Royal Institute of Technology, Stockholm. 420–423.

Hlavac, Jim 2011. Hesitation and monitoring phenomena in bilingual speech: A consequence of code-switching or a strategy to facilitate its incorporation? *Journal of Pragmatics* 43. 3793–3806.

Hokkanen, Tapio 2001. *Slips of the tongue. Errors, repairs and a model.* Finnish Literature Society, Helsinki.

Hooper, Celia R. – Cralidis, Ann 2009. Normal changes in the speech of older adults: You've still got what it takes; it just takes a little longer! *Perspectives on Gerontology* 14. 47–56.

Horváth Viktória 2009. *Funkció és kivitelezés a megakadásjelenségekben.* PhD-disszertáció. ELTE, Budapest.

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

Huber, Jessica E. – Stathopoulos, Elaine T. – Curione, Gina M. – Ash, Theresa A. – Johnson, Kenneth 1999. Formants of children, women, and men: The effects of vocal intensity variation. *The Journal of the Acoustical Society of America* 106. 1532–1543.

Hudson Kam, Carla L. – Edwards, Nicole A. 2008. The use of uh and um by 3- and 4-year-old native English-speaking children: Not quite right but not completely wrong. *First Language* 28. 313–327.

Igeta, Takako – Arai, Takayuki 2011. Case study on comparison of male and female vowel formants by native speakers of Korean. http://www.icphs2011.hk/resources/OnlineProceedings/RegularSession/Igeta/Igeta.pdf

Ito, T. 1986. Speech dysfluencies and acquisition of syntax in children 2–6 years old (abstract). *Folia Phoniatrica* 38. 310.

Kemper, Susan. 1992. Adults sentence fragments: Who, what, when, where and why? *Communications Research* 19. 444–458.

Lee, Sungbok – Potamianos, Alexandros – Narayanan, Shrikanth. 1999. Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *The Journal of the Acoustical Society of America* 105. 1455–1468.

Leeper, Linda H. – Culatta, Richard 1995. Speech fluency: Effect of age, gender and context. *Folia Phoniatrica et Logopaedica* 47. 1–14.

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.

Levin, Harry – Silverman, Irene. 1965. Hesitation phenomena in children's speech. *Language and Speech* 8/2. 67–85.

Lickley, Robin J. 1995. Missing disfluencies. In Elenius, Kjell – Branderud, Peter (eds.): *Proceedings of the XIIIth International Congress of Phonetic Sciences.* Royal Institute of Technology, Stockholm. 192–195.

Lickley, Robin J. – Bard, Ellen Gurman 1996. On not recognizing disfluencies in dialog. In *Proceedings, International Congress on Spoken Language Processing.* Philadelphia. 1876–1879

Local, John. 2004. Getting back to prior talk: and-uh(m) as a back-connecting device. In Couper-Kuhlen, Elizabeth – Ford, Cecilia E. (eds.): *Sound production in interaction.* John Benjamins, Amsterdam. 377–400.

MacLachlan, Barbara G. – Chapman, Robin S. 1988. Communication breakdowns in normal and language learning-disabled children's conversation and narration. *Journal of Speech and Hearing Disorders* 53. 2–7.

MacWhinney, Brian – Osser, Harry 1977. Verbal planning functions in children's speech. *Child Development* 48. 978–985.

Manning, Walter H. – Monte, Karen L. 1981. Fluency breaks in older speakers: Implications for a model of stuttering throughout the life cycle. *Journal of Fluency Disorders* 6. 35–48.

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.

Merlo, Sandra – Mansur, Letícia Lessa. 2004. Descriptive discourse: topic familiarity and disfluencies. *Journal of Communication Disorders* 37. 489–503.

Montes, Rosa Graciela 1999. The development of discourse markers in Spanish: interjections. *Journal of Pragmatics* 31. 1289–1319.

Neuberger Tilda 2013. *Életkor-specifikus sajátosságok az anyanyelv fejlődésében.* PhD-disszertáció. ELTE, Budapest.

Pak, Maria – Sprott, Richard – Escelera, Elena 1996. Little words, big deal: the development of discourse and syntax in child language. In Slobin, Dan Isaac (ed.): *Social*

Interaction, Social Context and Language: Essays in Honor of Susan Ervin-Tripp. Lawrence Erlbaum Associates, Mahwah, NJ. 287–305.

Pickett, James M. 1980. *The sounds of speech communication.* Unversity Park Press, Baltimore.

Pindzola, Rebekah H. 1990. Dysfluency characteristics of aged, normal-speaking black and white males. *Journal of Fluency Disorders* 15. 235–243.

Pope, Benjamin – Blass, Thomas – Siegman, Aron W. – Raher, Jack 1970. Anxiety and depression in speech. *Journal of Consulting and Clinical Psychology* 35. 128–133.

Ragsdale, J. 1976. Relationship between hesitation phenomena, anxiety and self-control in a normal communication situation. *Language and Speech* 19. 257–265.

Ragsdale, Donald J. – Sisterhen, Daniel H. 1984. Hesitation phenomena in the spontaneous speech of normal and articulatory-defective children. *Language and Speech* 27. 235–244.

Reuterskiöld, Christina – Hansson, Kristina – Sahlén, Birgitta 2011. Narrative skills in Swedish children with language impairment. *Journal of Communication Disorders* 44. 733–744.

Rispol, Matthew – Hadley Pamela 2001. The leading-edge: The significance of sentence disruptions in the development of grammar. *Journal of Speech, Language, and Hearing Research* 44. 1131–1143.

Roberts, Patricia M. – Meltzer, Ann – Wilding, Joanne 2009. Disfluencies in non-stuttering adults across sample lengths and topics. *Journal of Communication Disorders* 42. 414–427.

Rochester, S. R. – Gill, Judith 1973. Production of complex sentences in monologues and dialogues. *Journal of Verbal Learning and Verbal Behavior* 12. 203–210.

Roggia, Aaron B. 2012. *Eh* as a polyfunctional discourse marker in Dominican Spanish. *Journal of Pragmatics* 44. 1783–1798.

Schachter, Stanley – Christenfeld, Nicholas J. S. – Ravina, Bernard – Bilous, Frances 1991. Speech disfluency and the structure of knowledge. *Journal of Personality and Social Psychology* 20. 362–367.

Schachter, Stanley – Rauscher, Frances – Christenfeld, Nicholas – Tyson Crone, Kimberly 1994. The significance of sentence disruptions in the development of grammar. The vocabularies of Academia. *Psychological Science* 5. 37–41.

Schiro, Martha 2003. Genre and evaluation in narrative development. *Journal of Child Language* 30. 165–195.

Searl, Jeffrey P. – Gabel, Rodney M. – Fulks, Steven J. 2002. Speech disfluency in centenarians. *Journal of Communication Disorders* 35. 383–392.

Shillcock, Richard – Kirby, Simon – McDonald, Scott – Brew, Chris 2001. Filled pauses and their status in the mental lexicon. *Proceedings of DiSS '01 Disfluency in Spontaneous Speech.* University of Edinburgh, Scotland. 53–56.

Shriberg, Elisabeth 1994. *Preliminaries to a theory of speech disfluencies.* PhD Dissertation. University of California. ftp://www.speech.sri.com/pub/papers/

Shriberg, Elizabeth 1996. Disfluencies in Switchboard. In: *Proceedings of International Conference on Spoken Language Processing.* Addendum, Philadelphia. 11–14.

Shriberg, Elizabeth 2001. To „errrr" is human: ecology and acoustic of speech disfluencies. *Journal of the International Phonetic Association* 31. 153–169.

Simpson, Adrian P. 2006. Phonetic processes in discourse. In Brown, Keith (ed.): *Encyclopedia of language and linguistics*. Elsevier, Amsterdam. 284–288.

Simpson, Adrian P. 2009. Phonetic differences between male and female speech. *Language and Linguistics* 3. 621–640.

Smith, Vicki L. – Clark, Herbert H. 1993. On the course of answering questions. *Journal of Memory and Language* 32. 25–38.

Sorenson, David N. 1989. A fundamental frequency investigation of children ages 6–10 years old. *Journal of Communication Disorders* 22. 115–123.

Swerts, Marc 1998. Filled pauses as markers of discourse structure. *Journal of Pragmatics* 30. 485–496.

Thordardottir, Elin T. – Weismer, Susan Ellis 2002. Content mazes and filled pauses on narrative language samples of children with specific language impairment. *Brain and Cognition* 48. 587–592.

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

Veiga, Arlindo – Candeias, Sara – Lopes, Carla – Perdigão, Fernando 2011. Characterization of hesitations using acoustic models. In: *Proceedings of the International Congress of Phonetic Sciences*. University of Hong Kong, Hong Kong. 2054–2057. http://www.icphs2011.hk/resources/OnlineProceedings/RegularSession/veiga

Vorperian, Houri K. – Kent, Ray 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.

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*. Saarbrücken. 561–564. (http://www.icphs2007.de/conference/Papers/1453/1453.pdf)

Yairi, Ehud – Clifton, Noel F. Jr. 1972. Disfluent speech behavior of preschool children, high school seniors, and geriatric persons. *Journal of Speech and Hearing Research* 15. 714–719.

Yaruss, J. Scott – Robyn M. Newman – Tracy Flora 1999. Language and disfluency in nonstuttering children's conversational speech. *Journal of Fluency Disorders* 24. 185–207.

Ward, Nigel 2000. The challenge of non-lexical speech sounds. In: *Proceedings of the International Conference on Spoken Language Processing (ICSLP)*. Beijing, China, (2.) 571–574.

Watanabe, Michiko – Hirose, Keikichi – Den, Yasuharu – Minematsu, Nobuaki 2008. Filled pauses as cues to the complexity of upcoming phrases for native and non-native listeners. *Speech Communication* 50. 81–94.

A tanulmány a Bolyai János kutatási ösztöndíj támogatásával készült.

TEMPORÁLIS SZERVEZŐDÉS KILENCÉVES GYERMEKEK SPONTÁN BESZÉDÉBEN

Horváth Viktória

Bevezetés

A beszélő életkora nagymértékben befolyásolja a beszéd számos jellemzőjét. A beszéd tervezése és artikulációs kivitelezése, a közlés grammatikai megformáltsága, a megakadásjelenségek, a temporális jellegzetességek, a szóhasználat, a diskurzusjelölők használata mind függenek attól, hogy a beszélő idős vagy fiatal felnőtt, óvodás vagy kisiskolás gyermek, avagy középiskolás (pl. Yairi–Clifton 1972; Caruso et al. 1997; Bortfeld et al. 2001; Vousden–Maylor 2006; magyarra: Balázs 1993; Gocsál 2000; Menyhárt 2003; Szabó 2008; Laczkó 2009, 2010; Markó–Dér 2011; Bóna 2012a, b; Deme 2012; Neuberger 2012b).

A beszéd fejlődésében meghatározó az első hat életév. A kutatások alapján feltételezhető, hogy az első biológiai sorompó 6–7 éves kor között van; ez az utolsó időpont a maradéktalanul sikeres anyanyelv-elsajátítás megkezdéséhez (Gósy 2005). Az első hat évben zajlanak le az anyanyelv-elsajátítás legfontosabb folyamatai; ezután látványos minőségi változások már nem mennek végbe. A hazai és a külföldi kutatások mindezek miatt a gyermekek első hat életévének nyelvi fejlődését állítják középpontba (MacWhinney–Osser 1977; DeJoy–Gregory 1985; Stemberg 1989; Adams–Gathercole 1995; Yaruss et al. 1999; Robinson et al. 2000; magyarra: Gósy 1981, 1997; Lengyel 1981; Horváth 2006; Markó et al. 2010; Csákberényiné–Hajdu 2011; Markó–Dér 2011; Neuberger 2011, 2012b; Deme 2012). A nemzetközi CHILDES korpusz is elsősorban hatéves korig tartalmaz hanganyagokat, jóval kevesebb a kisiskolás korú gyermekekkel készült felvétel (http://childes.psy.cmu.edu/).

Az ezt követő, 7–10 éves kor közötti szakaszról jóval kevesebb adatunk van, noha fontos lenne tudni, hogy a spontán beszéd milyen változásokon megy át az első biológiai sorompó után. Az anyanyelv elsajátítása nem zárul le a hetedik életévvel, gyakorlatilag egész életen át tartó folyamat. A kisiskolás korban történő változások érintik a beszéd szegmentális szintjét is. Angol gyermekekkel végzett kutatás szerint 9 és 12 éves kor között a beszédhangok időtartamának tartománya és változatossága az életkor előrehaladtával szignifikánsan csökken, 12 éves kor körül már a felnőtt beszédben mért értékekhez hasonlít (Lee et al. 1999). Magyar nyelvre igazolták, hogy 9 éves korra már kialakul a fonológiailag rövid–hosszú magánhangzópárok fizikai időtartamának elkülönülése (Bóna–Imre 2010). Az ejtésbeli variabilitás a mássalhangzók

esetében is csökken az életkor előrehaladtával. Az artikulációs gesztusok megszilárdulásának folyamata a felpattanó zárhangok zöngekezdési idejében is nyomon követhető: 9 éves korban még jóval nagyobb szórást mutatnak a gyermekek VOT-értékei, mint két évvel később (Whiteside et al. 2003). A beszéd szegmentális szintje mellett a szupraszegmentális tényezők is változnak kisiskolás korban. Az alaphangmagasság 7–8 éves korban szignifikánsan csökken a fiúknál, a lányoknál már nem változik nagymértékben ebben az életkorban (Ferrand–Bloom 1996). Más kutatások szerint azonban a lányoknál ekkor megy végbe az alaphangmagasság csökkenése, a fiúknál pedig 8–9 éves kor között zajlik (Hacki–Heitmüller 1999).

Még ebben az életkorban is fejlődik a grammatika, folytatódik bizonyos nyelvtani szerkezetek megszilárdulása (Crystal 2003). Kisiskolás korban továbbá évről évre nő a beszédegységeken belüli szavak száma, a gyermekek egyre több szóból álló beszédszakaszokat hoznak létre (Loban 1976). A szókincs – részben az írott nyelv elsajátításának hatására – nagymértékben növekszik (Nippold 2006).

A magyar kisiskolások beszédét csupán néhány szempontból elemezték az elmúlt években; ilyen például a lexikális folyamatok működése (Neuberger 2012b). A gyermekek az életkor előrehaladtával egyre több szót mondanak az asszociációs tesztben, már minden szófaj előfordul, vagyis egyre hatékonyabban működik az előhívás. A kilencévesek közlései szerkezetileg is komplexebbek, mint az óvodásoké; továbbá a spontán beszédükben növekszik a lexikai eredetű megakadások aránya. A kilencéveseknél már gyakoriak az összetett mondatok, közléseikben sok és változatos kötőszókat használnak. Beszédükben a főnevek dominálnak (Nagy J. 1978), de már előfordulnak a ragozott főnévi igenevek; megnő a határozószók, a névmások és melléknevek száma (Neuberger 2011). A ragozás szintén tökéletesedik, a gyermekek már nyolc éves korban képesek bizonyos morfológiai kombinációk helyes alkalmazására (gyakorító és ható képző, segédigék), az agrammatikus mondatok aránya pedig jelentősen csökken (Laczkó 2011).

A grammatika mellett azonban nagyon fontos jellemzők a temporális jegyek, mert információval szolgálnak a tervezési folyamatokról, a tervezés és kivitelezés összehangolásáról. Széleskörű leírásuk ezért fontos adatokkal szolgál egy adott korosztály/csoport spontán beszédéről, és annak életkori fejlődéséről.

A jelen kutatás célja a temporális jellemzők vizsgálata kilencéves gyermekek spontán beszédében. A kutatás fő kérdése, hogy milyen paraméterek jellemzik a beszédet ebben az életkorban, mert ezeket összevetve a kisebb gyermekek és a felnőttek adataival, képet kaphatunk az anyanyelv-elsajátítás egy állomásáról, az esetleges minőségi változásokról és tendenciákról. A kutatás során arra a kérdésre is kerestük a választ, hogy vannak-e különbségek a fiúk és a lányok spontán beszédének temporális sajátosságai között.

A kutatás hipotézisei a következők voltak: (i) a kilencéves gyermekek temporális jellemzői jellegzetesen eltérnek az óvodásokra és a felnőttekre jellemzőktől; (ii) nagy a különbség a fiúk és lányok beszédének folyamatossága között: a lányok hosszabban és folyamatosabban beszélnek.

Anyag és módszer, kísérleti személyek

A kutatáshoz narratívákat rögzítettünk. A kísérletvezető előre közölte a témákat a gyermekekkel, és csakis akkor szólalt meg, amikor segítő, a felvételt továbbvivő kérdésre volt szükség. A téma a gyermekek családja, lakókörnyezete, hobbija, iskolai elfoglaltságai voltak. A teljes korpusz időtartama 67 perc (gyermekenként 3–5 perc).

A kisiskolásokkal készült felvételek a narratívák hasonló tematikája, protokollja és időtartama szempontjából is hasonlóak a BEA adatbázis spontán narratíváihoz; így a gyermekek adatai összevethetők a felnőttekéivel (Gósy et al 2012).

A gyermekekkel készült hangfelvételek a megszokott iskolai környezetben, de csendes körülmények között készültek Sony ICD-SX700 típusú hangfelvevővel; a kísérletvezető minden esetben ugyanaz a személy volt.

A kutatáshoz 18 gyermek (9 lány és 9 fiú) beszédét rögzítettük. A gyermekeket véletlenszerűen választottuk ki egy budapesti általános iskola harmadik osztályából. Mindannyian egynyelvűek, tipikus beszédfejlődésűek voltak, és egyiküknek sem volt hallásproblémája.

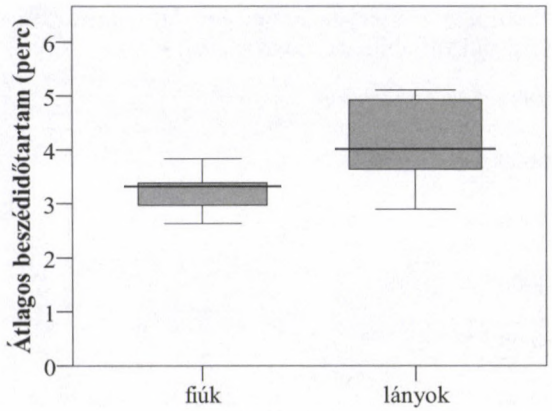
A korpuszt négy szinten annotáltuk a Praat 5.3 programban (Boersma–Weenink 2011): beszédszakaszok és szünetek, szavak és szünetek, tematikus egységek és megakadásjelenségek. Elemeztük a beszédszakaszok (szünettől szünetig tartó beszédegységek), a néma és a kitöltött szünetek időtartamát; a kitöltött szünetek realizációját; az artikulációs és beszédtempót. Az eredményeket – ahol lehetett – összevetettük a felnőttekkel készült hasonló kutatás adataival (Gósy et al. 2011), és minden esetben kitértünk a lányok és fiúk spontán beszédének összehasonlítására. A statisztikai elemzéseket (független mintás *t*-próba, többtényezős varianciaanalízis, Mann–Whitney U-próba, Kruskal–Wallis-próba) az SPSS 13.0 programban végeztük.

Eredmények

A gyermekek átlagosan 3'42"-es narratívákat hoztak létre (a legrövidebb időtartama 2'36", a leghosszabbé 5'06" volt). A lányokra általában az volt a jellemző, hogy szívesebben, hosszabban beszéltek, a fiúk szűkszavúbbak voltak. A fiúk narratíváinak átlagos időtartama 3'12"; a lányoké egy perccel több (4'10", vö. 1. ábra). A különbség statisztikailag is szignifikáns (független mintás *t*-próba: $t(16) = -3{,}119$; $p = 0{,}009$).

Elemeztük a beszédszakaszok (1695 darab), azaz a szünettől szünetig terjedő közlésegységek időtartamát. Az összes beszédszakasz átlagos időtartama 1241 ms (átl. elt. 907,73 ms). A legrövidebb egy *a* névelő volt, alig több mint

10 ms-os megvalósulással; a leghosszabb beszédszakasz időtartama pedig 6213 ms volt: *amikor a jobb szélen el akarnak nálam futni akkor énnekem kell kitolni a pályáról és megszerezni a labdát.* A kilencévesek átlagosan 540 ms-mal rövidebb egységeket hoztak létre két szünet között, mint a felnőttek (Gósy et al. 2011).

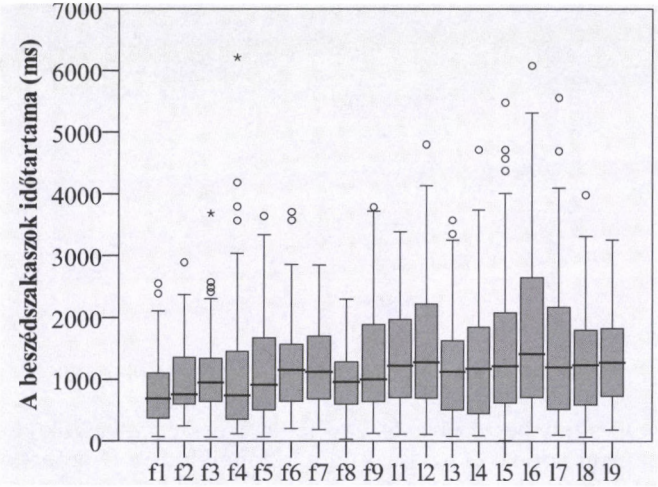


1. ábra
A fiúk és a lányok narratíváinak átlagos időtartama

A lányok és fiúk beszédében összevetettük a beszédszakaszok időtartamát. Az adatok nem normál eloszlásúak, így Mann–Whitney U-próbát alkalmaztunk. Az eredmények szerint statisztikailag szignifikáns különbség van a lányok és a fiúk beszédszakaszainak időtartamában ($Z = -6{,}728$; $p < 0{,}001$). A fiúk beszédszakaszainak átlagos időtartama 1056 ms (átl. elt. 736,7); a lányoké nagyjából 350 ms-mal több (1404 ms, átl. elt. 1008,3). A lányok és a fiúk beszédszakaszainak időtartama között majdnem akkora a különbség, mint a felnőttekre és a gyermekekre jellemző értékek között. Az egyéni különbségeket a 2. ábra mutatja. Volt olyan gyermek, akinek beszédszakaszai 30 ms és 2300 ms időtartam között szóródtak; a legnagyobb variabilitás egy beszélőn belül pedig 145 ms és 6077 ms közötti tartományban valósult meg.

Az összes beszédszakaszban átlagosan 2,66 szó valósult meg (átl. elt: 1,59). A gyermekek beszédszakaszainak 50,1%-a csupán egyetlen szóból állt (például *a, én, enni, ezt, épület, és, gereblyézni, focizni, játszok, juharszirup, kúttal, mindig, Németországban, régi, rózsaszín* stb.). A példák azt mutatják, hogy a gyermekek nagyon sokszor egyáltalán nem szerkezethatáron tartottak szünetet, elválasztották akár a jelzős vagy igeneves szerkezetek egymáshoz szorosan kapcsolódó tagjait is. Ez az artikulációs tervezés és megvalósítás kevésbé gyakorlott voltára utal. A legtöbb szó 18 volt egy adott szakaszban. A kilencéveseknél különbség mutatkozott a nemek között abban is, hogy át-

lagosan hány szóból álló beszédszakaszokat produkáltak. A fiúk átlagosan 2,29 lexémát ejtettek két szünet között (átl. elt. 0,57), a lányok ennél többet, 3,03 szót (átl. elt. 0,32). A különbség statisztikailag is szignifikáns: a többtényezős varianciaanalízis azt igazolta, hogy nemcsak a beszédszakaszok hosszában, hanem – ezzel nyilvánvalóan összefüggésben – az ezeken belüli szószámban is eltérés van a fiúk és a lányok beszéde között ($F(1, 1694) = 43{,}676$; $p < 0{,}001$). A beszédszakaszonkénti szószám az egyes gyermekek között is nagy egyéni variabilitást mutat (3. ábra).

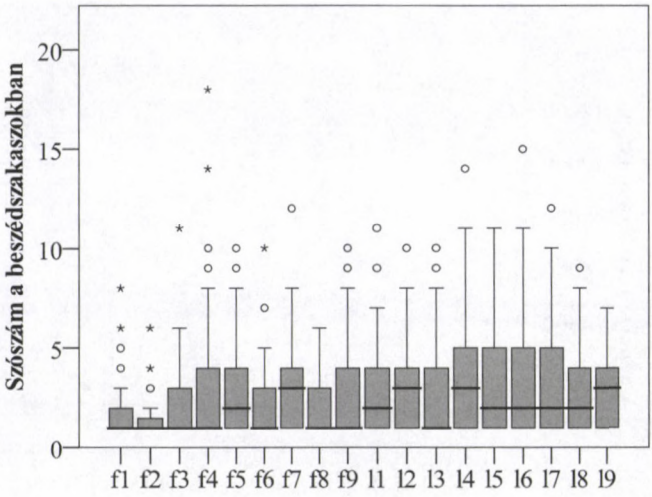


2. ábra
A beszédszakaszok időtartama egyénenként (ms)

A spontán beszéd egyik fontos jellemzője, hogy a néma és a kitöltött szünetek milyen gyakorisággal és időtartamban valósulnak meg. A teljes korpuszban 1842 darab néma szünet fordult elő, és 374 darab kitöltött szünet. A szünetek a teljes korpusz időtartamának 46%-át tették ki, ez az arány a felnőttek narratíváiban 20–30% körüli (vö. pl. Markó 2005, Menyhárt 2010, Váradi 2010). Az 1842 darab néma szünet átlagos időtartama 944 ms (átl. elt. 1380); a felnőttek narratíváiban ez az átlag 510 ms (Gósy et al. 2011). A kilencéveseknél a legrövidebb időtartamú néma szünet 31 ms-os (a felnőtteknél 23 ms-os) volt; a leghosszabb néma szünet pedig 14513 ms (egy lány beszédében adatoltuk egy új téma indítása előtt). A felnőttek narratíváiban adatolt leghosszabb szünet 3036 ms volt.

A néma szünetek időtartamai nem normál eloszlásúak, így Mann–Whitney U-próbát alkalmaztuk. Az eredmények szerint statisztikailag szignifikáns különbség van a lányok és a fiúk beszédszüneteinek időtartamában ($Z = -3{,}280$;

p = 0,001). A lányok néma szünetei átlagosan 881 ms-os időtartamban valósultak meg (átl. elt. 1345); a fiúkéi 1071 ms-osak voltak (átl. elt. 1415). A fiúk tehát hosszabb szüneteket tartanak a spontán beszédben, mint a lányok. Ez a különbség felnőttkorban is megmarad; a férfiak beszédére hosszabb néma szünetek jellemzők (Gósy et al. 2011).



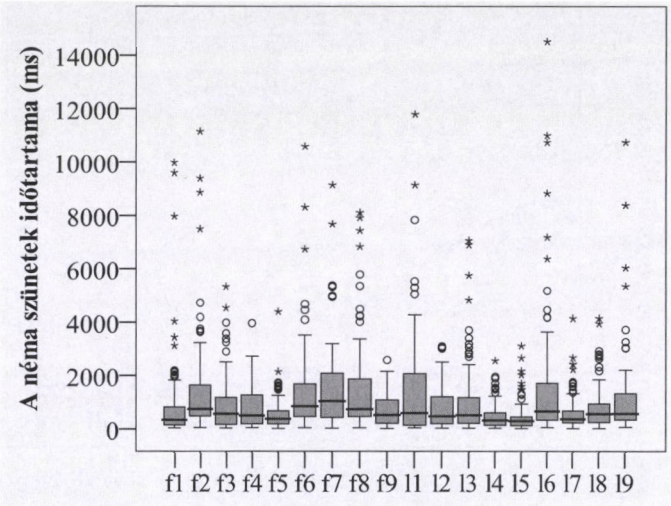
3. ábra
Szószám a beszédszakaszokban

A kilencéveseknél nemcsak a fiúk és lányok csoportja között van eltérés a beszédszünetek időtartamában; nagy szóródást mutatnak az értékek egy beszélőn belül és a beszélők között is (4. ábra).

Mindösszesen 12 olyan néma szünetet adatoltunk, amelyek szón belül fordultak elő; ezek átlagos időtartama 319 ms. A legrövidebb 55 ms (*hogy be* szünet *jussunk a célba*), a leghosszabb pedig 660 ms időtartamban valósult meg (*já* szünet *tszani*). Az ilyen szünetek 67%-át fiúk produkálták.

Az összes néma szünet 36%-ában megjelent valamilyen egyéb, nonverbális hangjelenség. Az esetek 67,3%-ában (az összes ilyet tartalmazó szünetet véve 100%-nak) hangos ki/belégzést, sóhajszerű hangjelenséget adatoltunk a néma szünet teljes vagy részleges időtartamában. Jóval kisebb volt az aránya azoknak a hangjelenséget tartalmazó néma szüneteknek, amelyek alatt a gyermekek nyelvcsettintést produkáltak (16,9%). Hasonló arányban fordult elő, amikor ez a két hangjelenség kombinálódott a szünetben (15,2%). Egy-egy példát találtunk olyan néma szünetekre (0,2%), amelyekben a következők jelentek meg: nevetés, torokköszörülés, hangos nyelés, *fff*-szerű hangjelenség. Egy kutatás a BEA-adatbázis alapján felnőttek narratíváiban vizsgálta

részletesen ezeket a hangjelenségeket (Neuberger 2012a). A felnőttek korpuszában szinte azonos a hangos légzések aránya (69%), mint a gyermekeknél; a csettintések pedig náluk is a második leggyakoribbak (17%). A nevetésben azonban nagy különbséget találtunk: a felnőtteknél 11%-ban jelennek meg (szemben a 0,2%-kal).



4. ábra
A néma szünetek időtartama (ms)

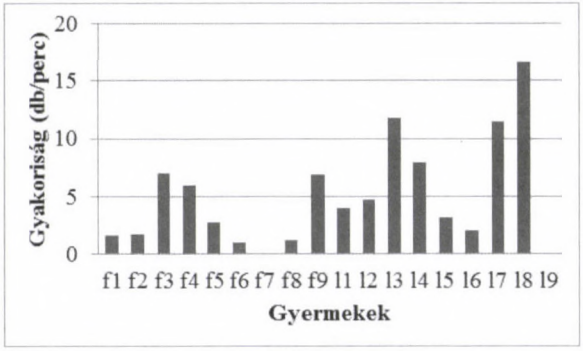
A korpuszban 374 darab kitöltött szünetet (hezitálást) adatoltunk; percenként átlagosan 5,6 fordult elő a hanganyagban. Ez gyakoribb, mint a felnőtteknél, és majdnem ötször gyakoribb a 6–7 éves óvodások beszédéhez képest is (Horváth 2010a, b).

A gyermekek között nagyok az egyéni különbségek a kitöltött szünetek percenkénti gyakoriságában (5. ábra). Két gyermek egyáltalán nem produkált hezitálást; a maximális percenkénti gyakoriság pedig 16 volt.

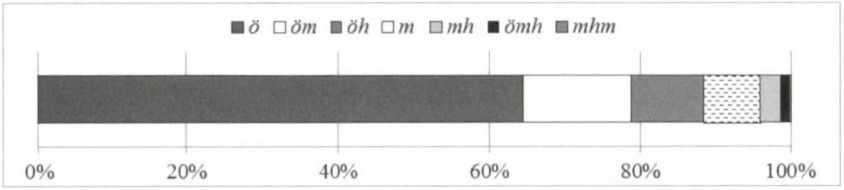
A lányok átlagosan kétszer annyi kitöltött szünetet produkáltak percenként (6,9, átl. elt. 5,5); mint a fiúk (3,1; átl. elt. 2,7); a különbség tendenciaszerű, statisztikailag nem szignifikáns. A fiúk és a lányok között ebben a mutatóban az óvodásoknál és a felnőtteknél sincs matematikailag igazolható különbség (Horváth 2010a, b).

A kitöltött szünetek sokféle realizációt mutatnak a beszédben. A leggyakoribb a semleges magánhangzó (*ö*-re emlékeztető hang), aránya 64,7% a kilencévesek korpuszában (6. ábra). A svászerű hezitálás előfordulásának már csak töredéke a hangkapcsolatból álló kitöltött szünetek aránya (14,2% ill. 9,6%). A bilabiális nazálisszerű hangot a gyermekek a kitöltött szünetek

7,5%-ában ejtették, a többi forma előfordulási aránya 5% alatti. A felnőttek beszédében szintén a svá a leggyakoribb formája a hezitálásnak, de aránya nagyobb, majdnem 80% (Horváth 2010a).



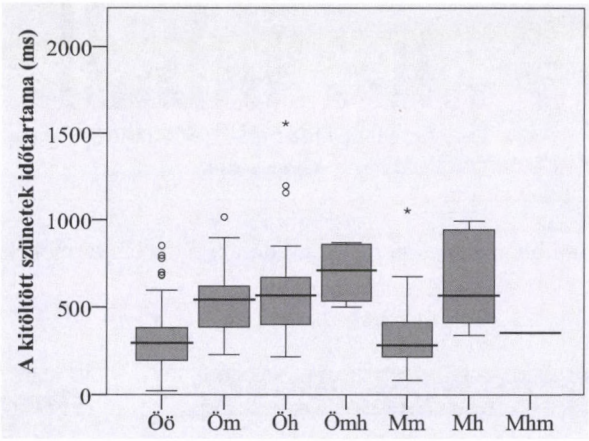
5. ábra
A kitöltött szünetek percenkénti darabszáma a gyermekeknél



6. ábra
A kitöltött szünetek realizációjának aránya a korpuszban (%)
(A semleges magánhangzót az *ö* betű jelöli az ábrán.)

A hezitálások megjelenhetnek két néma szünet között vagy egy szóhoz tapadva, szinte a szó részeként – ilyenkor az artikuláció folyamatos. Az izolált hezitálások tendenciaszerűen hosszabb időtartamban valósulnak meg (Gósy 2012). A kilencévesek összes hezitálásának 50,8%-a volt tapadó. Az egyes hezitációtípusokon belül a tapadó jelenségek aránya más-más mintázatot mutat. A leggyakoribb típus, a svá (*Ö*-vel jelölve a példában), az esetek 63%-ában a megelőző vagy követő szóhoz tapadt, például: *ÖSiófokon, ésÖ, megÖ, mentünkÖ,ígyÖ* stb. A példák azt mutatják, hogy a svá-hezitálás tartalmas és funkciószavakhoz tapadva is megjelenhet. A második leggyakoribb hezitálás, az *öm* hangkapcsolat már csak 24,5%-ban tapadt szóhoz; az *öh* viszont az esetek majdnem 60%-ában. A bilabiális nazális előfordulása már 10% alatti volt a korpuszban, és alig volt jellemző rá, hogy szóhoz tapadt volna, feltételezhetően artikulációs okok miatt. Ugyanilyen ok miatt az *ömh* hangkapcsolat mindössze négyszer fordult elő a korpuszban, és csak egyszer volt tapadó jellegű.

A kitöltött szünetek átlagos időtartama 379 ms (átl. elt. 208,6 ms). Ehhez képest az óvodások (343 ms, vö. Horváth 2010b) és a felnőttek is átlagosan rövidebb kitöltött szüneteket produkáltak (327 ms, Horváth 2010a). A kilencéveseknél a legrövidebb kitöltött szünet 21 ms időtartamban valósult meg (egy svászerű realizáció); a leghosszabb pedig 1556 ms időtartamú volt (egy *öh*-realizáció). Az egyes típusok időtartamát külön-külön is elemeztük. Az egyetlen hangból álló kitöltött szünetek rövidebb időtartamban valósulnak meg, mint a hangkapcsolatok (7. ábra).



7. ábra
A kitöltött szünetek időtartama a realizáció függvényében
(A realizációk sorrendje jobbról balra a gyakorisági sorrendet tükrözi.)

A svá időtartama átlagosan a legrövidebb, ennél 34 ms-mal hosszabbak a bilabiális nazálisok. A *mhm* hangkapcsolatot egyszer adatoltuk a korpuszban 355 ms időtartamban. 500 ms időtartamátlag felettiek az *öh* és *öm* hezitálások; a leghosszabbak pedig a *mh* és *öhm* jellegű kitöltött szünetek, az időtartamadatokat az 1. táblázat mutatja.

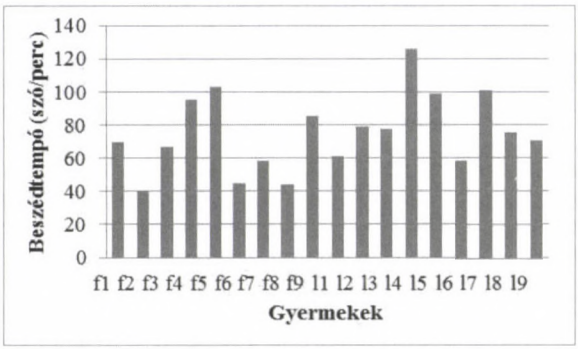
A kitöltött szünetek időtartama függ azok realizációjától; a nemparametrikus Kruskal–Wallis-próba szignifikáns különbséget igazolt az időtartamokban a hezitációtípusok között ($\chi^2 = 89{,}612$; $p < 0{,}001$; a próbát csak az elegendő elemszámú csoportokra végeztük el).

Elemeztük a gyermekek átlagos beszédtempóját, amely 75,2 szó/perc (átl. elt. 23,14 szó/perc). Ez az érték a hatéves óvodások tempójához hasonló (Neuberger 2011). Korábbi vizsgálatok szerint (Gósy 2003; Gyarmathy 2007) a felnőttek átlagos beszédtempója 140 szó/perc körüli értéket mutat, vagyis a felnőtt adatközlők kétszer több szót ejtenek egy perc alatt, mint a vizsgált gyermekek.

1. táblázat: Az egyes hezitáció-típusok időtartama (ms)

Az egyes hezitációtípusok időtartama (ms)							
	ö	*öh*	*öm*	*ömh*	*mh*	*mhm*	*m*
Átlag	308	587	532	698	642	356	341
Átlagos eltérés	144	285	175	190	260	–	213
Minimum	22	218	231	500	341	356	84
Maximum	851	1557	1015	870	992	356	1052

A lányok valamivel gyorsabban beszéltek, átlagosan 8 szóval ejtettek többet egy perc alatt (83,03 szó/perc, átl. elt. 21,70 szó/perc), mint a fiúk (67,38 szó/perc, átl. elt. 23,01 szó/perc); de a független mintás *t*-próba nem igazolt statisztikailag szignifikáns különbséget a két nem között (8. ábra).

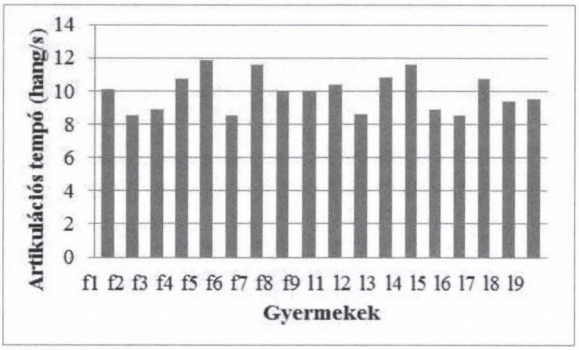


8. ábra
A beszédtempó átlagértékei az egyes beszélőknél (szó/perc)

A kilencévesek átlagos artikulációs tempója 10,03 hang/s volt (átl. elt. 3,21 hang/s). Ez lassabb, mint a felnőtteknél mért tempóérték (13,14 hang/s, vö. Gósy et al. 2011). A jelen kutatás eredménye szerint a kilencévesek átlagosan 3 hanggal ejtenek kevesebbet másodpercenként, mint a felnőttek. Ez az artikulációs tempó csoportszinten a legidősebb vizsgált beszélőkéhez, a százévesekéhez hasonló (Bóna 2012a). A Hegedűs-archívum tanúsága szerint hasonló artikulációs tempóval beszéltek a mintegy 60 évvel ezelőtt élt 10–12 éves gyermekek is (Menyhárt 2012). A nemek között nem volt szignifikáns különbség az artikulációs tempó értékeiben, a lányoknál mért átlagos érték 9,98 hang/s (átl. elt. 3,16 hang/s), a fiúké kicsit gyorsabb (10,08 hang/s, átl. elt. 3,27 hang/s).

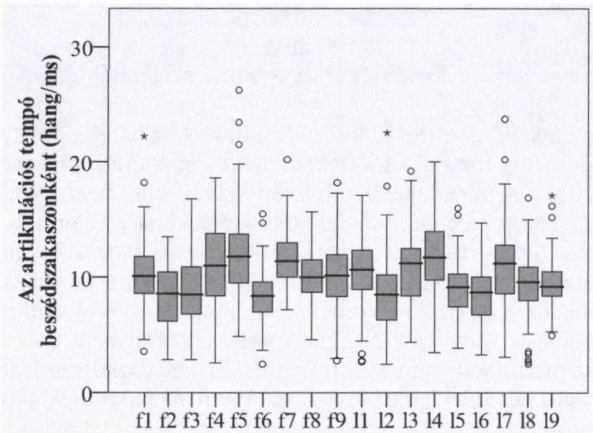
A leglassabb átlagos artikulációs tempó 8,4 hang/s volt, a leggyorsabb pedig 11,9 hang/s; mindkettőt fiúnál adatoltuk (9. ábra). Az artikulációs tempó esetében nemcsak az átlagokat, hanem a beszédszakaszonkénti értéket is meghatároztuk. A gyermekek között nagyok az egyéni különbségek; a

nemparametrikus Kruskal–Wallis-próba szignifikáns különbséget igazolt a beszédszakaszonkénti artikulációs tempóban a beszélők között ($\chi^2 = 216{,}080$; $p < 0{,}001$).



9. ábra
Az artikulációs tempó átlagértékei az egyes beszélőknél (hang/s)

Az artikulációs tempó variabilitása nemcsak a beszélők között, hanem egy narratíván belül is nagy eltérést mutat az egyes beszédszakaszokban. Több olyan beszédszakaszt is adatoltunk a korpuszban, amelyben mindössze 2,5–3 hang realizálódott másodpercenként. Az artikulációs tempó maximuma 26,2 hang/s volt egy beszédszakaszon belül (10. ábra).



10. ábra
Az artikulációs tempó értéke beszédszakaszonként

Következtetések

A jelen kutatásban kilencéves gyermekek spontán beszédének temporális jellegzetességeit vizsgáltuk. Ezek a gyermekek már harmadik éve iskolába járnak; az intézményes oktatás hatással van a nyelvtani szabályrendszer tudatosságára. Beszédük szintaktikailag összetettebb, mint az óvodásoké (Neuberger 2011), de a fejlődés ebben a korban nem zárul le a grammatika terén. Kilencéves korban még összefüggés van a grammatikai komplexitás és a megakadásjelenségek száma között, de a 13 éveseknél már nem igazolható ez az összefüggés; a bonyolultabb közlések létrehozása számukra már nem jár együtt több tervezési nehézséggel. Ez is azt mutatja, hogy a kisiskolásoknál még tart egy fejlődési szakasz a közlések grammatikai megformálásának terén (Neuberger 2011).

A grammatikai mutatók mellett a beszéd temporális szerveződése is változik iskoláskorban. A kilencévesek beszédében még jóval nagyobb arányban jelennek meg szünetek, több időt vesz igénybe számukra a gondolatok kiválogatása és a megfelelő nyelvi forma hozzárendelése, mint a felnőtteknek. A gyermekek néma szünetei átlagosan hosszabbak (944 ms), mint a középiskolásoké (760 ms, Vallent 2010) és a felnőtteké (510 ms, Gósy et al. 2011). Az adatok azt igazolják, hogy 9 éves kortól felnőttkorig szinte megfeleződik a néma szünetek átlagos időtartama, a beszélők az életkor előrehaladtával egyre rövidebb néma szüneteket tartanak, a felnőttekre jellemző átlagidőtartam időskorban is megmarad (Bóna 2012a).

A beszédszakaszok szószáma is változik az életkor előrehaladtával. A kilencéves gyermekekéhez képest a felnőttek kétszer annyi lexémát ejtenek egy szünettől szünetig terjedő egységben (Gósy et al. 2011). A különbség feltételezhető oka az, hogy a gyermekek egyrészt kisebb beszédtapasztalattal rendelkeznek, másrészt számukra még nehezebb a gondolat és a megfelelő nyelvi forma összehangolása.

A beszédtempóban ugyanezt a tendenciát láthatjuk. A kilencévesek átlagosan 75 szót, a középiskolások 123 szót, a felnőttek pedig körülbelül 140 szót ejtenek percenként (Gósy 2003; Gyarmathy 2007). Felnőttkorra tehát szinte megkétszereződik a percenként ejtett szavak száma. Ennek oka, hogy felnőttkorra a beszélők jóval nagyobb beszédtapasztalattal bírnak, emiatt gyorsabban működnek a beszédtervezési (lexikális előhívás, grammatikai, fonológiai és artikulációs) folyamatok. Az artikulációs tempó értékeinél ugyanezt a gyorsuló tendenciát látjuk 9, 15 és 18 éves között (vö. Laczkó 2009).

A vizsgált kisiskolásoknál még nagy különbségek adatolhatók a fiúk és a lányok spontán beszédében. A lányok szívesebben beszéltek; szignifikánsan hosszabb narratívákat hoztak létre, amelyek hosszabb, több szóból álló beszédszakaszokból épültek fel, néma szüneteik rövidebbek.

A kilencéveseknél előfordultak extrém hosszú, 10 másodpercet meghaladó néma szünetek is. Egy korábbi, társalgásokat is vizsgáló kutatás (Markó 2005) szerint a felnőtt beszélőknél nagyjából 5 másodperc körül van az a ha-

tár, ami fölött már zavaró a hallgatás. A gyermekek korpuszában 42 darab olyan néma szünet fordult elő, amely meghaladja ezt az értéket. A jelen vizsgálat ugyan narratívákra épül, de itt is jelen volt egy kísérletvezető, aki kérdezte a gyermekeket az adott témákról, ennek ellenére akár 10 másodpercig is csendben maradtak. Spontán beszédük tehát ebben az értelemben nem felel meg az időzítési maximának, vagyis „ne hagyd, hogy túl hosszúra nyúljon a hallgatás!" (Leech 1983). Ez azt mutatja, hogy a pragmatikai szabályok elsajátítása még nem zárul le maradéktalanul kilencéves korra. A pragmatika fejlődését igazolja a nevetés jóval kisebb aránya a felnőttek narratíváihoz viszonyítva (Neuberger 2012a). A BEA-felvételeket felnőtt beszélő rögzíti felnőtt adatközlőkkel; míg a jelen adatbázishoz a hanganyagot felnőtt készítette a gyermekekkel. A beszédpartnerek életkora és egymáshoz való viszonya pedig befolyásolja a beszéd jellegzetességeit (vö. pl. Bata–Gráczi 2009). Ebben az aszimmetrikus helyzetben a gyermekek – a pragmatikai szabályokat figyelembe véve – úgy érezhették, hogy nem illik nevetgélniük, komolyan kell venniük a feladatot. Más kutatások szerint ugyanakkor a pragmatikus tudat még nem alakult ki teljesen: egy nyelvi feladatban a tízéves kor alatti gyermekek nem kértek pontosító információt akkor sem, ha az utasítás alapján nem tudták elvégezni a feladatot. Csak 11-12 éves korra alakul ki az a képesség is, hogy a gyermek felismerje, egymásnak ellentmondó vagy érthetetlen információt tartalmaz a közlés (Pratt–Nesdale 1984).

Az anyanyelv elsajátítása tehát nem zárul le még kisiskolás korban sem, ezért nagyon fontos lenne az egyes korosztályok beszédének nagyszámú adatközlőn végzett, minél széleskörűbb vizsgálata annak érdekében, hogy a beszédfejlődés mennyiségi és minőségi változásairól minél pontosabb képet kaphassunk.

Irodalom

Adams, Anne-Marie – Gathercole, Susan E. 1995. Phonological working memory and speech production in preschool children. *Journal of Speech and Hearing Research* 38. 403–414.

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

Bata Sarolta – Gráczi Tekla Etelka 2009. A beszédpartner életkorának hatása a beszéd szupraszegmentális jellegzetességeire. In Keszler Borbála – Tátrai Szilárd (szerk.): *Diskurzus a grammatikában – grammatika a diskurzusban*. Tinta Kiadó, Budapest. 74–82.

Bóna Judit 2012a. A spontán beszéd sajátosságai idősödő, idős és matuzsálemi korban. In Markó Alexandra (szerk.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngekezdési időig*. ELTE BTK–MTA NYTI, Budapest. 100–115.

Bóna Judit 2012b. Hogyan mondanak vissza hallott szövegeket a középiskolások? *Anyanyelv-pedagógia* 2012/2. http://www.anyp.hu/cikkek.php?id=379

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.

Bortfeld, Heather – Leon, Silvia D. – Bloom, Jonathan E. – Schober, Michael F. – Brennan, Susan E. 2001. Disfluency rates in conversations: Effects of age, relationship, topic, role and gender. *Language and Speech* 44/2. 123–147.

Caruso, Anthony J. – McClowry, M. Troy – Max, Ludo. 1997. Age-related effects on speech fluency. *Seminars in Speech, Language and Hearing* 18. 171–180.

Crystal, David 2003. *A nyelv enciklopédiája.* Osiris Kiadó, Budapest.

Csákberényiné Tóth Klára – Hajdu Tünde 2011. Az iskolába lépő gyerekek szóbeli szövegalkotásának jellemzői. *Anyanyelv-pedagógia* 2011/3. http://www.anyp.hu/cikkek.php?id=332

Deme Andrea 2012. Óvodások magánhangzóinak akusztikai jellemzői. In Markó Alexandra (szerk.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngekezdési időig.* ELTE BTK–MTA NYTI, Budapest. 77–99.

DeJoy, Daniel A. – Gregory, Hugo H. 1985. The relationship between age and frequency of disfluency in preschool children. *Journal of Fluency Disorders* 10/2. 107–122.

Ferrand, Carole T. – Bloom, Ronald R. 1996. Gender differences in children's intonational patterns. *Journal of Voice* 10/3. 284–291.

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

Gósy Mária 1981. A beszédhang kialakulása a gyermeknyelvben. *Magyar Fonetikai Füzetek* 7. 67–90.

Gósy Mária 1997. A szavak hangalakjának változása a gyermeknyelvben. *Beszédkutatás '97.* 1–38.

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 2012. Sorozatmegakadások mintázata a spontán beszédben. *Beszédkutatás 2012.* 107–131.

Gósy Mária – Beke András – Horváth Viktória 2011. Temporális variabilitás a spontán beszédben. *Beszédkutatás 2011.* 5–31.

Gósy Mária – Gyarmathy Dorottya – Horváth Viktória – Gráczi Tekla Etelka – Beke András – Neuberger Tilda – Nikléczy Péter 2012. BEA: Beszélt nyelvi adatbázis. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások.* Akadémiai Kiadó, Budapest. 9–24.

Gyarmathy Dorottya 2007. Az alkohol hatása a beszédprodukcióra. *Beszédkutatás 2007.* 108–120.

Hacki, Tamas – Heitmüller, S. 1999. Development of the child's voice: premutation, mutation. *International Journal of Pediatric Otorhinolaryngology* 49. 141–144.

Horváth Viktória 2006. A spontán beszéd és a beszédfeldolgozás összefüggései gyerekeknél. *Beszédkutatás 2006.* 134–146.

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

Horváth Viktória 2010b. Hezitációs jelenségek óvodáskorban. In Gecső Tamás – Sárdi Csilla (szerk.): *Új módszerek az alkalmazott nyelvészeti kutatásban.* Segédkönyvek a nyelvészet tanulmányozásához 114. Tinta Kiadó, Budapest. 153–158.

Laczkó Mária 2009. Középiskolai tanulók beszédének temporális jellemzői. *Magyar Nyelvőr* 133. 447–467.
Laczkó Mária 2010. Megakadásjelenségek a spontán és a szónoki beszédben. *Beszédkutatás 2010*. 184–198.
Laczkó Mária 2011. Óvodások és kisiskolások spontán mondatalkotási folyamatai. *Magyar Nyelvőr* 135. 440–458.
Lee, Sungbok – Potamianos, Alexandros – Narayanan, Shrikant 1999. Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *Journal of Acoustical Society of America* 105/3. 1455–1468.
Leech, Geoffrey N. 1983. *Principles of pragmatics*. Longman, London.
Lengyel Zsolt 1981. *A gyermeknyelv*. Gondolat Kiadó, Budapest.
Loban, Walter 1976. Language development: Kindergarten through grade twelve. NCTE Committee on Research Report 18. http://www.eric.ed.gov/PDFS/ED128818.pdf
MacWhinney, Brian – Osser, Harry 1977. Verbal planning functions in children's speech. *Child Development* 48. 987–985.
Markó Alexandra 2005. A temporális szerkezet jellegzetességei eltérő kommunikációs helyzetekben. *Beszédkutatás 2005*. 63–77.
Markó Alexandra – Gráczi Tekla Etelka – Imre Angéla 2010. A diskurzusjelölők használatának fejlődése: a hümmögés formai és funkcionális sajátosságai különböző életkorokban. In Navracsics Judit (szerk.): *Nyelv, beszéd, írás. Pszicholingvisztikai tanulmányok I.* Segédkönyvek a nyelvészet tanulmányozásához 107. Tinta Könyvkiadó, Budapest. 82–92.
Markó Alexandra – Dér Csilla Ilona 2011. Diskurzusjelölők használatának életkori sajátosságai. In Navracsics Judit – Lengyel Zsolt (szerk.): *Lexikai folyamatok egy- és kétnyelvű közegben. Pszicholingvisztikai tanulmányok II.* Tinta Könyvkiadó, Budapest. 49–61.
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 2010. A beszédsebesség objektív mérési és szubjektív észlelési eredményeinek összefüggései mai és 60 évvel ezelőtti beszélőknél. *Beszédkutatás 2010*. 110–124.
Menyhárt Krisztina 2012. A beszéd temporális jellemzői 60 évvel ezelőtti gyermek beszélőknél. *Beszédkutatás 2012*. 246–259.
Nagy J. József 1978. A szófajok gyakorisági jellemzői a 8–10 éves tanulók nyelvhasználatában. *Magyar Nyelv* 74. 186–204.
Neuberger Tilda 2011. Gyermekek spontán beszédének szerkesztettsége és folyamatossága. *Beszédkutatás 2011*. 83–95.
Neuberger Tilda 2012a. Nonverbális hangjelenségek a spontán beszédben. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest. 215–235.
Neuberger Tilda 2012b. Szóelőhívás gyermekek szóasszociációiban és spontán beszédében. In Navracsics Judit – Szabó Dániel (szerk.): *A mentális folyamatok a nyelvi feldolgozásban. Pszicholingvisztikai tanulmányok III.* Tinta Könyvkiadó, Budapest. 85–94.

Nippold, Marylin. A. 2006. Language development in school-age children, adolescents, and adults. In Brown, Keith (ed.): *Encyclopedia of language and linguistics*. Elsevier, Boston. 368–373.

Pratt, Chistopher – Nesdale, Andrew R. 1984. Pragmatic awareness in children. In Tunner, William E. – Pratt, Christopher – Herrimann, Michael. L. (eds.): *Metalinguistic awareness in children. Theory, research, and implication*. Springer-Verlag, Berlin. 105–125.

Stemberger, Joseph P. 1989. Speech errors in early child language production. *Journal of Memory and Language* 28. 164–188.

Szabó Kalliopé 2008. Megakadásjelenségek nyolcévesek spontán beszédében. *Anyanyelv-pedagógia* 2008/2. http://www.anyanyelv-pedagogia.hu/cikkek.php?id=56

Templin, Mildred. 1957. *Certain language skills in children*. University of Minnesota Press, Minneapolis.

Vallent Brigitta 2010. Középiskolások spontán narratíváinak jellemzői. *Beszédkutatás 2010*. 199–210.

Váradi Viola 2010. A felolvasás és a spontán beszéd temporális sajátosságainak összehasonlítása. *Beszédkutatás 2010*. 100–109.

Vousden, Janet I. – Maylor, Elizabeth 2006. Speech errors across lifespan. *Language and Cognitive Processes* 21. 48–77.

Whiteside, Sandra P. – Dobbin, Rachel – Henry, Luisa 2003. Patterns of variability in voice onset time: a developmental study of motor speech skills in humans. *Neuroscience Letters* 374/1. 29–32.

Yairi, Ehud – Clifton, Noel F. 1972. Disfluent speech behavior of preschool children, high school seniors, and geriatric persons. *Journal of Speech and Hearing Research* 15. 714–719.

Yaruss, J. Scott – Newman, Robyn M. – Flora, Tracy 1999. Language and disfluency in nonstuttering children's conversational speech. *Journal of Fluency Disorders* 24/3. 185–207.

A tanulmány a Bolyai János Kutatási Ösztöndíj támogatásával készült.

AZ ALVEOLÁRIS ZÖNGÉTLEN EXPLOZÍVA VARIABILITÁSA

Neuberger Tilda – Gráczi Tekla Etelka

Bevezetés

A beszédhangok ejtésekor az egyes artikulációs mozgások meghatározott időtartamban történnek, az artikulációs sajátosságok pedig tükröződnek az akusztikai vetületben (Gósy 2004). A felpattanó zárhangok képzésekor a glottisz felől áramló levegőáram akadályba ütközik, amelyet nem tud áttörni, a feltorlódott levegő intraorális, illetve szupraglottális nyomása azonban felpattintja a zárat. A zár feloldása általában hirtelen, relatíve rövid zörejes légárammal történik (Stevens 1998). Ebből adódik az explozívák összetett szerkezete: az akusztikai szerkezetben megváltoznak a paraméterek, így a mássalhangzó egyes részeinek időtartama külön-külön is meghatározható: beszélhetünk (i) a zárszakasz hosszáról, (ii) a zárfeloldásról, továbbá (iii) a zönge megindulásáig tartó szakasz időtartamáról, vagyis a zöngekezdési időről (VOT, az angol „voice onset time" rövidítéséből) (Lieberman–Blumstein 1988).

A szakirodalom a múlt század ötvenes éveinek végétől nagy teret szentel a zárhangok időszerkezeti vizsgálatának. Az eredmények objektív értékekkel írják le az explozívák szerkezeti felépítését: számos kutatás született a zárszakasz hosszáról, illetve a zöngekezdési időről (pl. Lisker–Abramson 1964; Gandour–Dardarananda 1984; Cho–Ladefoged 1999; Kovács 2001; Gósy–Ringen 2009). Az akusztikai lenyomaton a zárszakasz alatt a magyar explozíváknál fojtott vagy előzönge (zöngés zárhang esetén) vagy néma fázis (zöngétlen zárhang esetén) látható. Az explozívák tipikus realizációit aerodinamikai okok, koartikulációból vagy lazább artikulációból eredő hatások megváltoztathatják, így létrejöhetnek korai vagy további felpattanások is, avagy a zár feloldása nem követhető nyomon a regisztrátumon annak gyenge intenzitása miatt (pl. Westbury–Keating 1986; Elekfi 1992; Shadle 1997; Gráczi–Kohári 2012).

A zárhangok objektív időtartamát befolyásolja a képzési hely, a hangszalagok működése, valamint a nyelvi időtartam. A fonológiai oppozíció szerinti rövid és hosszú párok zörejeleme nem mutat szignifikáns időtartam-változást, a zárszakasz jelentős megnyúlása számít a hosszúság legfontosabb jelölőjének (Kovács 2001). A zöngés mássalhangzó időtartama rendszerint rövidebb, mint a homorgán, azonos képzési módú zöngétlen párjáé (Magdics 1966; Docherty 1992; Olaszy 2006; Lousada et al. 2010; Gráczi 2011). Az időviszonyok nem függetlenek a mássalhangzó fonetikai pozíciójától, valamint a

közvetlen beszédhangkontextustól (Kovács 2000). A zárhang környezetében lévő magánhangzó minősége, időtartama például hatással lehet a zöngekezdési idő értékére (Kassai 1979). Gósy (2010) a zöngétlen alveoláris explozívát spontán beszédben elemezte az időtartam, a zár felnyitása utáni zöngétlen rész időtartama (a követő magánhangzó 2. formánsáig) szempontjából. Három közlésbeli helyzetet vetett össze, a szókezdő, szó belseji intervokális és a szóvégi helyzetet. A szókezdő és a szó belseji helyzet lényegében nem tért el egymástól, a szóvégi pozícióban azonban jelentősen hosszabb időtartamokat és nagyobb szóródást adatolt.

Különbség mutatható ki a hangidőtartamokban az életkor, a nem, illetve a beszédtípus függvényében, például az olvasott vagy ismételt beszéd, illetve az izolált szavak, mondatok ejtése és a spontán beszéd között (Ryalls et al 1997; Gósy 2000a, 2000b; Torre–Barlow 2009; Bóna 2011).

Az explozívák időszerkezetére a felsorolt tényezőkön kívül a beszélők egyéni artikulációs stratégiái, illetve anatómiai eltérései (pl. a szájüreg térfogata) is hatnak.

A jelen kutatás fő kérdései, hogy (i) mennyiben befolyásolja a fonetikai helyzet és a kontextus a zárhangok akusztikai szerkezetét, illetve (ii) az egyéni ejtés milyen jellegzetes eltéréseket eredményez a vizsgált explozívák időtartamában. Feltételeztük, hogy (i) a mássalhangzó teljes, illetve egyes részeinek időtartamára hatással lesz a hangkörnyezet, valamint hogy (ii) a beszélő sajátos artikulációja hangkörnyezettől függetlenül hatást gyakorol az időtartamokra.

Az elemzéshez a [t] fonémát választottuk – relatíve nagy gyakorisága miatt (pl. Szende 1973; Gósy 2004). Noha felpattanó zárhangok minden eddig elemzett nyelvben előfordulnak, közülük leggyakoribbak a zöngétlen és nem aspirált mássalhangzók (Henton et al. 1992). Egy nyelven belül a különböző típusú beszédhangok előfordulási gyakorisága eltérő, ami meghatározza az adott nyelv sajátos hangzását. A magyar mássalhangzókat tekintve a képzési mód függvényében a felpattanó zárhangok előfordulási gyakorisága a legmagasabb (37%-os), a képzés helye szerint pedig az alveolárisok vannak túlsúlyban (50,6%); spontán beszédben a [t] hang jelenik meg leggyakrabban a mássalhangzók közül (6,6%-os előfordulás) (Gósy 2004).

Kísérleti személyek, anyag, módszer

Az egyéni sajátosságok elemzésére egy olvasott beszéden alapuló vizsgálatot terveztünk, amelyben egy beszédhang beszélőnkénti közel 60 előfordulását elemeztük különböző fonetikai helyzetekben és kontextusokban.

Tíz (5 férfi, 5 nő) adatközlőt választottunk ki a BEA spontánbeszéd-adatbázis (Gósy 2008) beszélői közül. Életkoruk 20–29 év közötti (átlagéletkor: 23,7 év), mind budapesti, magyar anyanyelvű személy. Az adatközlők feladata 25 egyszerű és összetett mondat felolvasása volt (például *A tulipánágyások gyomlálásával ütötte el az időt. Nem kötött biztosítást, ezért kisebb*

vagyonba került a kórházi ellátás). Ebben az anyagban azokat a /t/ fonémákat elemeztük, amelyek megvalósulása az adott hangkörnyezetben [t]-ként volt várható, amelyekben legalább az egyik szomszédos fonéma magánhangzó volt, a másik szegmens pedig vagy egy mássalhangzó (kettős mássalhangzó-kapcsolat), vagy szünet (abszolút szókezdő és -végi helyzet), avagy magánhangzó (intervokális szókezdő, szóvégi vagy szó belseji helyzet (1. táblázat). Néhány esetben természetesen a mondatok többféle tagolhatósága miatt a szókezdő vagy szóvégi /t/ előfordulások előtt az aktuális meghangosítás függvényében változott az elemzett fonéma hangkörnyezete. Összesen 590 darab fonéma realizációját elemeztünk.

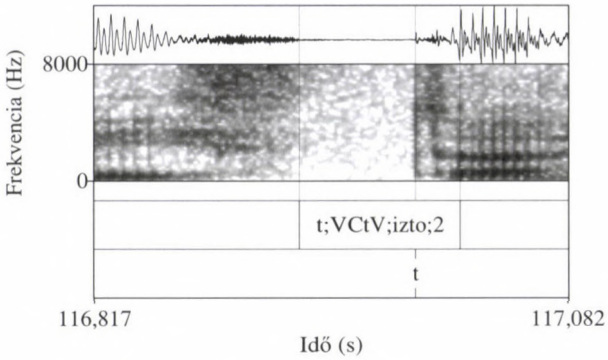
1. táblázat: Az explozívák előfordulása a fonetikai helyzet, illetve a kontextus alapján

Helyzet	Kontextus	Darabszám a szövegben	Megvalósított darabszám
||_V	*tö*	1	12
V#_V	*a#te (2), e#tú, a#tu, a#tú, e#te, i#tú*	7	79
V_V	*itá, ütö, ete (3), (au)tó, etá, íté, otá, uta, éte, ató, itö, ötö, átá, ota, ató*	17	188
V_#V	*et#a, et#e (2), et#é, it#a*	5	45
V_||	*őt, ét, at, át*	4	39
V_CV	*ötni, étvé, etje, étle*	4	48
VC_V	*esti, élte, eszte, egta, esté, űrtu, ista, elté, olta, izto (2), ulta, entő (2), érte, ente, esztü*	17	179

Az elemzés során a mássalhangzó időtartamát és belső szerkezetét vetettük össze. A rögzített hanganyagokat a Praat 5.3 (Boersma–Weenink 2011) szoftver segítségével címkéztük és elemeztük. Kézi annotálást alkalmaztunk a spektrogram és az oszcillogram mintázata és meghallgatás alapján. A vizsgált mássalhangzó határait (1. ábra) a megelőző, illetve a követő magánhangzó második formánsa alapján jelöltük. Mássalhangzós környezet esetében a mássalhangzóra jellemző akusztikai lenyomat alapján (rés kezdete/vége, zár kezdete/feloldás vége, második formáns kezdete/vége) határoztuk meg a hanghatárokat. A zár feloldásának kezdetét több felpattanás esetében az oszcillogramon és a spektrogramon elsőként megjelenő vonalazódásnál jelöltük. Adatoltuk az esetlegesen nem felpattanó zárhangként megvalósult előfordulások képzési módját, a fonetikai kontextust, illetve a zárszakasz alatt esetlegesen megjelenő réses áramlást és az egyéb tényezőket.

A mássalhangzó teljes időtartamát, a zárfeloldás időtartamát, e két időtartam arányát, a megvalósult képzési módot, a feloldás egyes jellemzőit és a zárszakasz alatti esetleges réses áramlás megjelenését elemeztük a kontextus

és a beszélők függvényében. A statisztikai elemzéseket az SPSS 13.0 verzió-számú szoftver segítségével végeztük.



1. ábra.
A címkézés bemutatása a *biztosítást* szóban

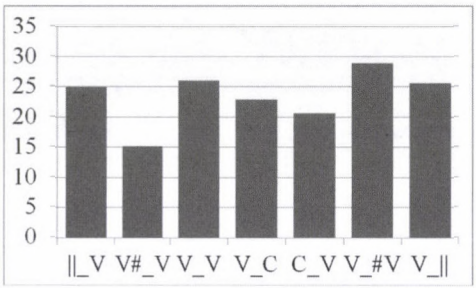
Eredmények

A vizsgált mássalhangzók nagy része zárhangként realizálódott, vagyis a spektrogramon és az oszcillogramon a zárszakasz látható volt. A nem mássalhangzó-kapcsolatban mért adatok minden esetben így jelentek meg, a zár feloldása pedig ugyancsak adatolható volt. 76%-uknál egyszeri, 24%-uk esetében pedig többszöri felnyílást találtunk. A többszöri feloldás gyakorisága 25–29% között alakult az egyes szóbeli magánhangzós pozíciókban – egy kivétellel, az intervokális szókezdő (V#_V) helyzetben, ahol az előfordulásuk csak 15,2% volt (2. ábra).

A mássalhangzó-kapcsolatok első tagjaként elemzett fonémák realizációi ugyancsak minden esetben zárhangként jelentek meg, a zár feloldása azonban az esetek 2,1%-ában nem volt adatolható, 23,0%-uk esetében pedig többszöri felpattanás jelent meg. Mindkét alveoláris szonoráns előtti helyzetben 26,3%-ban, a /v/ előtt csak 10,0%-ban fordult elő a zár többszöri feloldása.

Azokban a realizációkban, ahol a /t/ a mássalhangzó-kapcsolat második tagja volt, az esetek 16,2%-ában réshangként jelent meg. Néhány esetben a réshang és a zár alatti spirantikus szivárgás elkülönítése nem egyértelmű. A 3. ábrán egy-egy tipikus példát mutatunk be. A bal oldali ábra esetében a réshangtól csak kismértékben tér el a vizsgált fonéma realizációjának szerkezete, míg a jobb oldali ábra esetében annak kezdeti szakaszában nagyon alacsony intenzitású réses elem jelenik meg, majd a zárfeloldásra jellemzően hirtelenebb spirantikus zörej látható. Az előbbi megvalósulásokat réshangként, az utóbbiakat spirantikus szivárgással megvalósult zárhangként elemeztük. A két kategória közötti határ ugyanakkor nem húzható meg egyértelmű-

en, mivel egyrészt a felvételek nem fejmikrofonnal készültek, így az intenzitásviszonyok csak tájékoztató jellegűek lehetnek; másrészt maga a jelenség is fokozatosnak tekinthető. A spirantikus szivárgás a nem tökéletes zárképzés eredménye vagy a nagy szájüregi nyomás következménye.



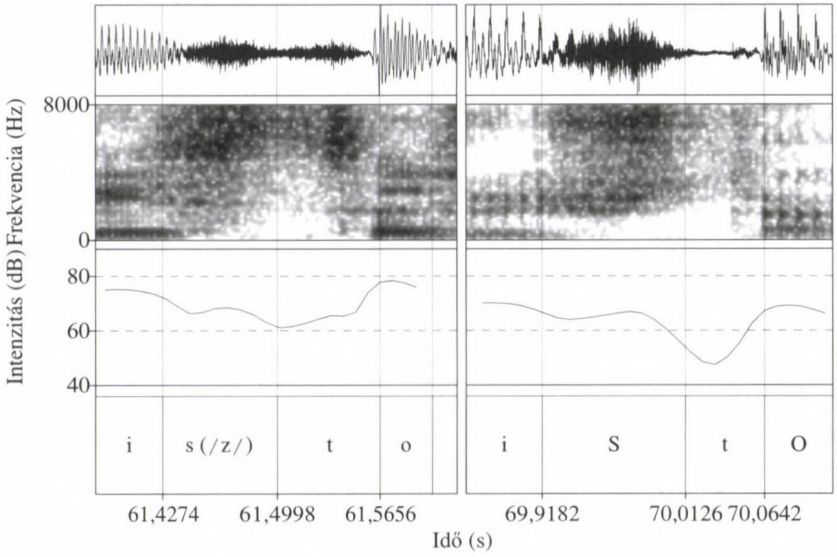
2. ábra.
A zár többszöri felnyílásának gyakorisága (%) a fonetikai helyzet függvényében

A réses szivárgás az összes C_V helyzetben elemzett hang 8,9%-át, csak a zárképzéssel megvalósult eseteket számítva 100%-nak pedig azok 10,7%-át érintette. A /t/ elsősorban réshang utáni helyzetben valósult meg gyakrabban réshangként (36,8%), míg szonoránsok (/n, l, r/) után ritkán (2,0%). A veláris zöngés felpattanó után 20,0%-ban, ez azonban két ejtést jelent az összesen tíz elemzettből. A spirantikus szivárgás ugyancsak a réshangok utáni helyzetben volt jellemző (40,5%), míg a további helyzetekben összesen egy előfordulást adatoltunk (/l/ után). A zárhang-realizációk között további 14,0%-ban jelent meg a zár határozott feloldása előtt réses elem.

A zárhangként megvalósult beszédhangok esetében gyakrabban fordult elő C_V pozícióban a zár többszöri felnyílása, mint más helyzetekben. A réshangok utáni pozícióban 20,9%, a /l/ környezetében 26,5%, a /r/ kontextusában 20,0%, a /n/-t követően 6,7%, a zöngés veláris explozívával alkotott mássalhangzó-kapcsolatban pedig 37,5% volt ennek gyakorisága.

Az egyes beszélők ejtésében eltérő arányban fordultak elő az egyes realizációs módok. A magánhangzós és C_V helyzetekben elemeztük a beszélők közötti variabilitást, mivel a V_C pozícióban csak 5-5 adatot mértünk egy kísérleti személytől. Az elemzett mássalhangzó-kapcsolatokban a zárhang-realizációk 0–33,3%-ában jelent meg többszöri felnyílás, míg az összes magánhangzós környezetben azok 5,3–55,6%-ában. Ez utóbbi csoportban két beszélő ejtésében a 47,2% és az 55,6%-os értékek kiugró adatnak tekinthetőek, azaz az ő esetükben a többi adatközlőnél jóval gyakrabban jelent meg ez a típusú megvalósulás. A kétféle helyzetben kapott adatok nem mutatnak összefüggést (Pearson-féle korreláció: $r = 0{,}437$, $p = 0{,}206$). A csak szó belseji

intervokális (V_V) helyzetet (0–63,2%) és a C_V pozíciót összehasonlítva sem található összefüggés (Pearson-féle korreláció: $r = 0{,}475$, $p = 0{,}165$). A C_V helyzetben adatolt réshang-realizációk száma az egyes beszélők ejtésében 5,6% és 27,8% közötti, a spirantikus szivárgást mutatók pedig 0% és 36,4% közötti gyakoriságot mutattak. A két jelenség között nem adatoltunk összefüggést (Pearson-féle korreláció: $r = 0{,}293$, $p = 0{,}412$).



3. ábra
Bal oldalon: a /t/ réshang-realizációja (a *biztosítást* szóban),
jobb oldalon: spirantikus szivárgás/áramlás (az *űrturista* szóban)

Jellemzően a férfiak ejtésében volt gyakoribb a réses megvalósulás (három beszélő esetében jelent meg 27,8%-ban réshang, míg a másik két férfi esetében 0% és 11,1%-ban), a nők kiejtésében pedig kisebb volt a beszélők közötti eltérés (5,6–22,2%). A spirantikus szivárgás megjelenésének gyakorisága kevésbé tért el a nemek között, de a két legmagasabb adatot a férfiak ejtésében mértük.

Elemeztük, hogy a mássalhangzó teljes időtartamában, illetve a zárfeloldás időtartamában milyen egyéni ejtési sajátosságok tükröződnek, illetve ezek mennyiben függenek a hangkörnyezettől.

A mássalhangzó-előfordulások átlagos időtartama az egyes beszélőknél 80 és 100 ms között szóródott (2. táblázat). A beszélők közötti szignifikáns eltérések csoportszinten megmutatkoznak az egyváltozós ANOVA szerint: $F(9{,}577) = 3{,}723$; $p < 0{,}001$. A tíz adatközlő közül az alábbiak mássalhang-

zó-időtartamaiban mutathatók ki szignifikáns eltérés ($p < 0,05$): n2 – f1, f2, f4, f5; n3 – f1.

2. táblázat: A mássalhangzó-időtartam, a feloldás időtartama és a feloldás-időtartam aránya beszélőnként (átlag és szórás)

Beszélő	Teljes időtartam (ms)	A feloldás időtartama (ms)	A feloldási és a teljes időtartam aránya (%)
n1	91±30	27±12	31,3±11,0
n2	100±23	21±10	21,9±11,2
n3	96±22	26±12	27,6±11,8
n4	91±23	26±9	30,1±9,9
n5	86±22	27±11	32,7±11,4
f1	80±21	28±13	34,9±11,4
f2	85±38	25±13	31,0±13,2
f3	92±24	26±11	29,3±13,5
f4	82±21	25±9	31,7±11,8
f5	83±25	31±16	37,2±12,9

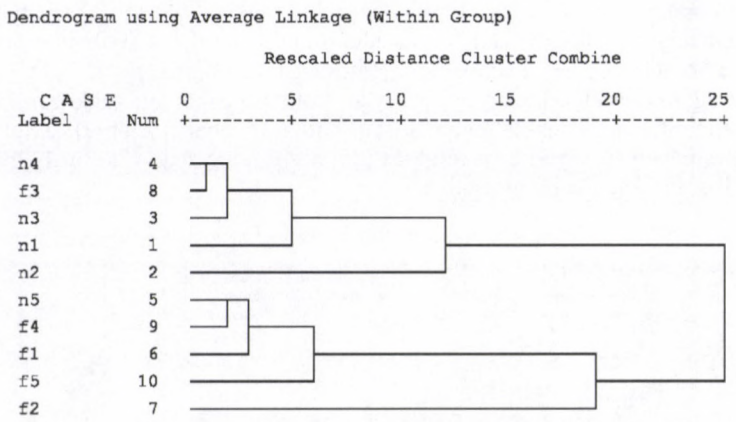
A zárfeloldás átlagos időtartama 21,4 és 31,3 ms közötti volt a tíz beszélőnél. Szignifikáns eltérést találtunk csoportszinten [egyváltozós ANOVA: $F(9,566) = 2,511$; $p = 0,008$)], illetve a Tukey-féle post hoc teszt alapján az n2 és az f5 adatközlők adatai között ($p < 0,001$).

A feloldás aránya átlagosan 21,9 és 37,2% közötti volt az egyes beszélőknél. Ebben a tekintetben az egyének között szignifikáns az eltérés [egyváltozós ANOVA: $F(9,554) = 7,021$; $p < 0,001$)]. A tíz adatközlő közül az alábbiak mássalhangzó-időtartamaiban mutatható ki szignifikáns eltérés ($p < 0,05$): n2 – többi személy (n3 kivételével); f5 – n3, n4, f3.

Szignifikáns különbséget találtunk a nemek között az explozíva teljes időtartamában és a feloldás arányában [$F(1,577) = 15,960$; $p < 0,001$ és $F(1,554) = 16,801$; $p < 0,001$). A férfiak átlagosan rövidebb időtartamban ejtették a [t] hangot (84,5±27 ms), mint a nők (93,2±24,8 ms). A zárfeloldás időtartama közel azonos volt a két nem felolvasásaiban (átlagosan 26 ms, illetve 28 ms), de a férfiaknál nagyobb arányt tett ki a mássalhangzó időtartamából (32,9±12,8%), mint a nőknél (28,7±11,6%). A hierarchikus klaszteranalízis az időtartam alapján a tíz beszélőt közel egyértelműen tudta csoportosítani a nemek tekintetében (4. ábra), egyetlen női és egyetlen férfi adatközlőt sorolt a másik nem csoportjába.

A [t]-realizációk átlagos időtartamát, feloldásuk időtartamát, illetve a feloldás arányát a fonetikai helyzetek szerint a 3. táblázat mutatja. Szignifikáns eltérést találtunk a [t] időtartamában a konszonáns fonetikai környezete szerint az egyváltozós ANOVA alapján: $F(5,577) = 50,709$; $p < 0,001$, illetőleg az egyes helyzetekben megjelent [t] hangok időtartamában, kivéve a C_V,

V_C – V_#V és a V_C – V_V helyzetek között (Tukey-féle post hoc teszt: $p < 0,05$). A mássalhangzók a legrövidebb időtartamban a C_V hangkörnyezetben valósultak meg (73,8±17,4 ms), a leghosszabban pedig V_|| abszolút szövégi helyzetben (127,1±29,1 ms). Utóbbi esetében a feloldás időtartama (41,3±21 ms) is szignifikánsan hosszabb volt, mint az összes többi helyzetben megjelent mássalhangzó feloldása. A C_V helyzetben kapott adatok esetében összevetettük a zár- és a réshangként megvalósult /t/-k időtartamait. Az előbbiek 76±17 ms, az utóbbiak pedig 63±13 ms átlagos időtartamban jelentek meg. Ez a Mann–Whitney-próba szerint szignifikáns eltérés ($Z = -3,844$; $p < 0,001$).



4. ábra.
A hierarchikus klaszteranalízis eredménye

3. táblázat: A mássalhangzó-időtartam, a feloldás időtartama és a feloldás-időtartam aránya a [t] fonetikai helyzete szerint (átlag és szórás)

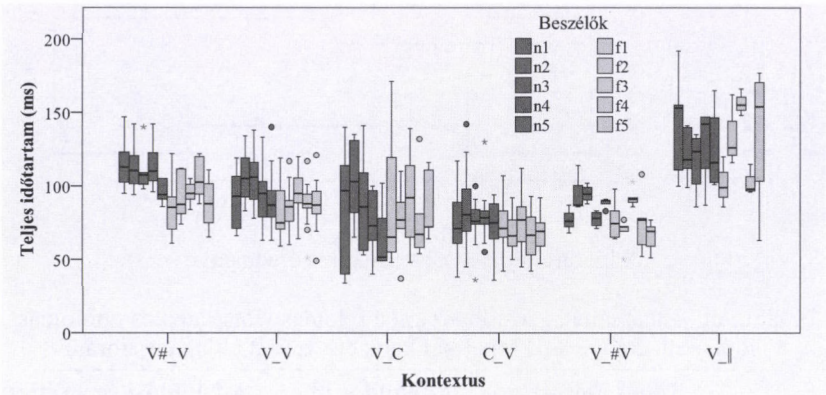
Helyzet	Teljes időtartam (ms)	A feloldás időtartama (ms)	A feloldási és a teljes időtartam aránya (%)
||_V	–	21±5	–
V#_V	103±30	22±6	22,8±7,2
V_V	92±17	24±8	26,8±9,3
V_#V	80±14	25±10	31,8±11,4
V_||	127±29	41±21	31,7±12,5
C_V	73±17	26±9	36,8±13,3
V_C	84±33	33±20	38,2±14,3

A feloldás időtartamában és a zárhang teljes időtartamához képest számított arányában is kimutatható a szignifikáns eltérés a helyzetek szerint [előb-

bi: $F(5,566) = 16,728$; $p < 0,001$), utóbbi: $F(5,554) = 26,334$; $p < 0,001$)]. A feloldás időtartama a mássalhangzó teljes időtartamához képest a V#_V helyzet esetében a legkisebb arányú (22,8%), míg a V_C helyzetben előforduló [t] hangoknál a legnagyobb arányú (38,2%). Az eredmények azt mutatják, hogy a mássalhangzó-környezetben lévő zárhangoknál a feloldás nagyobb arányú (36,8, illetve 38,2%), mint a magánhangzó-környezettel rendelkezőknél (22,8–31,8% közötti).

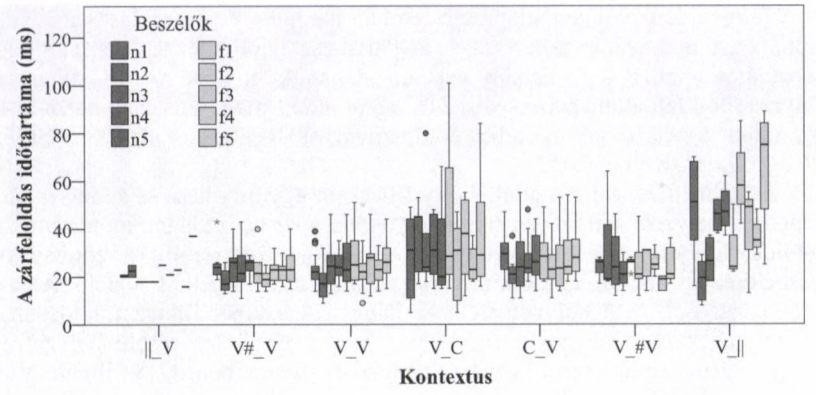
A beszélő egyéni ejtése és a konszonáns fonetikai helyzete tehát hatással van a mássalhangzó időtartamára, de a kettőjük interakciójában nem mutatható ki szignifikáns hatás az időtartamokra (MANOVA). A mássalhangzó-időtartam teljes varianciájából 33,4%-ot [korrigált R^2, $F(59,554) = 5,702$; $p < 0,001$] magyaráznak az általunk vizsgált független változók, amelyből a helyzet nagyobb befolyással bír az időtartamokra ($\eta^2 = 0,319$; $p < 0,001$), mint a beszélők egyéni artikulációs sajátosságai ($\eta^2 = 0,061$; $p < 0,001$).

Az egyes beszélők között nagy egyéni különbségeket találtunk mind a három fonetikai jegy tekintetében: a teljes időtartamban (5. ábra), a feloldás időtartamában (6. ábra) és a feloldás-időtartam arányában (7. ábra) egyazon fonetikai pozíción belül vizsgálva.

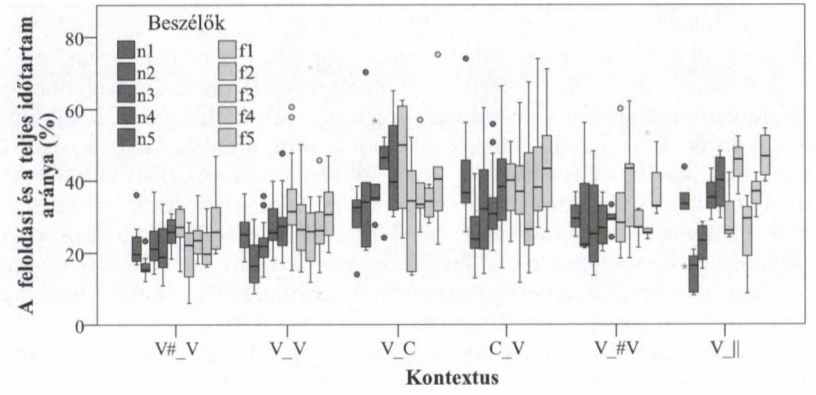


5. ábra
A /t/-realizációk teljes időtartama a vizsgált fonetikai helyzetekben beszélőnként

A legnagyobb egyéni eltérések a mássalhangzó teljes időtartamában a V_#V fonetikai helyzetben megjelent /t/-realizációk esetében rajzolódtak ki. Az egyazon helyzetben megvalósított [t] hangok esetében a V_C helyzetben szóródtak a legnagyobb mértékben az időtartam-értékek az egyes beszélőknél.



6. ábra
A /t/-realizációk zárfeloldásának időtartama a vizsgált fonetikai helyzetekben beszélőnként



7. ábra
A feloldás-időtartam aránya a vizsgált fonetikai helyzetekben beszélőnként

Következtetések

A jelen kutatásban az alveoláris zöngétlen felpattanó zárhang időviszonyait elemeztük tíz beszélő olvasott beszédében. A vizsgált fonetikai jegyek között a mássalhangzó teljes időtartama, a feloldás időtartama, illetve a feloldás időtartamának a teljes időtartamhoz viszonyított aránya szerepeltek. Tanulmányoztuk a /t/ realizációjának módját, a felpattanások számát, valamint időtartam-értékeket különböző fonetikai helyzetekben, különös tekintettel az egyéni artikulációs sajátosságokra.

A felolvasásokban nem alaprealizációban megjelent előfordulásokat is adatoltunk. A tipikus zárszakasszal és felpattanással jellemezhető zárhang-előfordulások mellett a /t/ néhány esetben réshangként képezve valósult meg. Egynél több felpattanás az esetek 24%-ában jelent meg, ami nagy hasonlóságot mutat a szakirodalom eddigi eredményeivel (Keating et al. 1980; Gráczi 2011; Gráczi–Kohári 2012).

Az eredmények azt mutatták, hogy a beszélő egyéni ejtése és a konszonáns fonetikai helyzete szignifikánsan befolyásolja a [t] beszédhang időtartamát, a feloldás időtartamát és arányát. A beszélők közötti variabilitás azonban kevésbé érezteti hatását az időtartamértékekre, mint a fonetikai pozíció. Az abszolút szóvégi, V_|| helyzetben lévő hang időtartama, illetve feloldásának időtartama szignifikánsan hosszabb volt, mint az összes többi helyzetben megjelent mássalhangzóé. A mássalhangzó-környezetben (C_V, illetve V_C helyzetben) előforduló [t] zárhangok időtartama átlagosan rövidebb, mint a magánhangzó-környezetben előfordulóké. Előbbiek esetében azonban a feloldás időtartama a mássalhangzó teljes időtartamához képest nagyobb arányt tesz ki, mint a magánhangzós helyzetekben megjelenők esetében.

A nemek között szignifikáns eltéréseket találtunk az explozíva időtartamát és a zárfeloldás időtartamának arányát illetően. A beszélők csoportokba osztása során (a /t/ realizációinak időviszonyai alapján) is a nemek szerinti együvé tartozás rajzolódott ki. A férfiak rövidebb időtartamban, ám nagyobb feloldásidőtartam-aránnyal valósították meg a vizsgált zárhangokat, valamint a réses realizációk is jellemzőbbek voltak rájuk, mint a nőkre. Nemzetközi (főként angol) kutatások (pl. Koenig 2002; Whiteside et al. 2004; Oh 2011) is kimutattak nemek közötti különbségeket a zöngétlen explozívák időszerkezetének bizonyos komponensében (a VOT értékében). Ezekre a különbségekre magyarázatául szolgálhatnak a férfiak és nők közötti anatómiai-fiziológiai különbségek, eltérések a beszédtempóban, az artikulációs pontosságban vagy ejtési tudatosságban, esetleg stilisztikai-szociolingvisztikai faktorok.

Az egyéni ejtési sajátosságok vizsgálatának az alkalmazott fonetika több területén is nagy jelentősége lehet, mint például a beszédhibák fonetikai elemzésében, az akcentusvizsgálatban vagy a beszéd alapján történő személyazonosításban (Gósy 2004).

Irodalom

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.)

Bóna Judit 2011. A [p, t, k] mássalhangzók zöngekezdési ideje idősek és fiatalok spontán beszédében és felolvasásában. *Beszédkutatás 2011*. 61–72.

Cho, Taehong – Ladefoged, Peter. 1999. Variations and universals in VOT: evidence from 18 languages. *Journal of Phonetics* 27. 207–229.

Docherty, Gerard J. 1992. *The timing of voicing in British English obstruents*. Foris Publications, Berlin.

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

Gandour, Jack – Dardarananda, Rochana 1984. Voice onset time in aphasia: Thai II. Production. *Brain and Language* 23. 177–205.

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

Gósy Mária 2000b. The voice onset time of the Hungarian voiceless plosives in words and in spontaneous speech. *International Journal of Speech Technology* 3–4. 155–164.

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

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

Gósy Mária 2010. Phonetic variation in Hungarian /t/. *Studia Slavica Hungarica* 55. 255–261.

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

Grászi Tekla Etelka 2011. Intervokális explozívák a zöngésségi oppozíció függvényében. *Beszédkutatás 2011*. 46–60.

Grászi Tekla Etelka – Kohári Anna 2012. A zöngekezdési idő egy módszertani kérdés függvényében. In Markó Alexandra (szerk.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngekezdési időig*. ELTE Bölcsészettudományi Kar–MTA Nyelvtudományi Intézet, Budapest, 228–248.

Henton, Caroline – Ladefoged, Peter – Maddieson, Ian. 1992. Stops in the world's languages. *Phonetica* 49. 65–101.

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

Keating, Patricia – Westbury, John R. – Stevens, Kenneth N. 1980. Mechanisms of stop-consonant release for different places of articulation. Előadás: *The Spring meeting of the Acoustical Society of America* 1980.

Koenig, Laura L. 2000. Laryngeal factors in voiceless consonant production in men, women, and 5-year olds. *Journal of Speech, Language and Hearing Research* 43. 1211–1228.

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

Kovács Magdolna 2001. A zöngétlen zárhangok és affrikáták időszerkezete. *Beszédkutatás 2001*. 46–60.

Lieberman, Philip – Blumstein, Sheila E. 1988. *Speech physiology, speech perception, and acoustic phonetics*. Cambridge University Press, Cambridge.

Lisker, Leigh – Abramson, Arthur S. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20. 384–422.

Lousada, Marisa – Jesus, Luis M. T. – Hall, Andreia 2010. Temporal acoustic correlates of the voicing contrast in European Portuguese stops. *Journal of the International Phonetic Association* 40/3. 261–276.

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

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

Oh, Eunjin 2011. Effects of speaker gender on voice onset time in Korean stops. *Journal of Phonetics* 39. 59–67.

Ryalls, John – Cliche, Annie – Fortier-Blanc, Kulie – Coulombe, Isabelle – Prud'Hommeaux, Annie 1997. Voice-onset time in younger and older French-speaking Canadians. *Clinical Linguistics and Phonetics* 11/3. 205–212.

Shadle, Christine H. 1997. The aerodynamics of speech. In Hardcastle, William J. – Laver, John (eds.): *The handbook of phonetic sciences*. Blackwell, Oxford, 33–64.

Stevens, Kenneth N. 1998. *Acoustic phonetics*. The MIT Press. Cambridge, Massachusetts–London.

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

Torre III, Peter – Barlow, Jessica A. 2009. Age-related changes in acoustic characteristics of adult speech. *Journal of Communication Disorders* 42/5. 324–333.

Westbury, John R. – Keating, Patricia A. 1986. On the naturalness of stop consonant voicing. *Journal of Linguistics* 22. 145–166.

Whiteside, Sandra P. – Henry, Luisa – Dobbin, Rachel. 2004. Sex differences in voice onset time: A developmental study of phonetic context effects in British English. *Journal of the Acoustical Society of America* 116/2. 1179–1183.

A PROGRESSZÍV AFÁZIA AKUSZTIKAI-FONETIKAI VONATKOZÁSAI (ESETTANULMÁNY)

Gósy Mária – Gyarmathy Dorottya

Bevezetés

A beszédképzés és a beszédmegértés zavara számos betegséghez kapcsolódóan jelenik meg, adott esetben vezető tünet lehet. Az organikus zavarok, főként, ha az agyműködéssel kapcsolatosak, különféle beszédprodukciós és beszédpercepciós problémához vezetnek (pl. Bánréti 1999; Obler–Gjerlow 1999). A jól ismert kórképek gyakran a szokásostól eltérő formában jelennek meg, amelyek a diagnózist, illetve a differenciáldiagnosztikát nehezítik (Szirmai 2002). Az atipikus beszédjelenségek mintázatának elemzése segít a ritkább vagy szokatlanabb formában előforduló kórképek azonosításában. A neurolingvisztikai, a pszicholingvisztikai, illetve a fonetikai megközelítések együttesen fontos kiegészítői lehetnek a korszerű orvosi diagnosztikáknak.

Az afáziás betegek nemegyszer nehezen kategorizálható tünetegyüttest produkálnak. Az 1970-es évek végétől egy amerikai neurológiai klinikán feltűnt az orvosoknak, hogy olyan esetek fordultak elő, amelyek nem voltak osztályozhatók az afázia ismert fajtái szerint, de a tünetek sem demenciát, sem egyéb hasonló betegséget nem igazoltak (l. Mesulam 2007). A bal agyfélteke frontális és temporális lebenyét érintő progresszív betegség első leírását Pick tanulmányában találták meg több mint 120 évvel ezelőtt (1892). Mintegy 90 évvel később hat hasonló eset alapján írták le az afázia egy ritka típusát, amelynek a neve a kezdetekben „lassan előrehaladó, demencia nélküli afázia" lett (*Slowly Progressive Aphasia Without Generalized Dementia*, l. Mesulam 1982). A nyelvvel kapcsolatos tünetek sokfélék voltak, akadtak fluens és nem fluens betegek; voltak, akik szóértési nehézségről panaszkodtak, mások nem. A jellegzetes tünetek nem egyeztek meg a stroke-ra jellemzőkkel. A sokféle tünetegyüttes miatt a kórképet 2007-től elsődleges progresszív afáziaként azonosították (*primary progressive aphasia*), és három típusát különítettek el: (i) agrammatikus, nem fluens, (ii) szemantikus és (iii) időszakosan nem fluens (erre a típusra Mesulam angol eredeti terminusa *logopenic*, ami neologizmus az angol nyelvben, l. 2007, S9). Ezekre a típusokra a következő jellemzőket állapították meg. Az agrammatikus, fluens típus esetében a szintaxis károsodik, de a szóértés megőrzött; a szemantikusnak nevezett típusban a szóértés nem működik, de a szintaxis érintetlen; a harmadik típusra jellemző a relatíve ép szintaxis és szóértés, de nehézséget jelent a lexikális

hozzáférés (a szótalálási nehézséget jelző szünetek gyakorisága). Az emlékezeti működések nehezítettsége kisebb-nagyobb mértékben valamennyi típusban előfordult, főként a harmadikra volt jellemző. A betegség előrehaladtával a beteg problémái fokozódhatnak, akár súlyos emlékezeti zavarok, agnózia, sőt személyiségbeli változások is kialakulhatnak.

Az elmúlt 25 évben sok száz tanulmány foglalkozott az afáziának ezzel a speciális típusával, azonban a relatíve ritka előfordulás miatt a klinikai diagnosztizálás még jelenleg is komoly nehézségekbe ütközik. A kórkép hatékony jóslásához nagy mennyiségű klinikai adat, jellemzők, képalkotó eljárások, illetve genetikai eredmények szükségesek. Mára már bebizonyosodott, hogy a betegséget valóban a bal félteke frontális és/vagy temporális területének degeneratív folyamatai okozzák, általában 65 éves kor előtt alakul ki, a beteg nyelv- és a beszédhasználata fokozatosan romlik. A tapasztalatok szerint az első mintegy két évben a kommunikáció bizonyos nehézségein túl semmilyen más kognitív vagy viselkedést érintő zavarral nem jár (Kertesz–Munoz 1997). A betegség általában szótalálási nehézségekkel, enyhe beszédmegértési zavarral, a grammatikai szabályok alkalmazásának problémájával kezdődik, továbbá artikulációs zavarok, sőt dadogás is felléphet (Mesulam 1982; Duffy–Petersen 1992). A jellemző tünetek rendszerint igen lassan alakulnak ki, a diagnózist akár évekkel megelőzően. Noha az eddigi vizsgálatok főként a spontán beszéd sajátosságait és a képmegnevezéseket érintették, megjegyzik, hogy több év folyamatában romlik a beszéd ismétlésének képessége (Grossman–Ash 2004).

Gomo-Tempini és munkatársai megkísérelték az elsődleges progresszív afázia mint ritka neurológiai szindróma pontosabb osztályozását (2011). Kutatásukban a nyelv és a beszéd jellemzőit a fent említett három típusban tovább specifikálták, amelyeket különféle klinikai, illetve genetikai adatok alapján igazoltak. A beszédjellemzők az egyes típusokban az alábbiak. A **szemantikus** típus jellemzői: a megnevezés nehezítettsége, izolált szavak jelentésének értési problémája, illetve három vagy több az alábbiak közül: tárgyak és/vagy személyek nehézkes azonosítása, felszíni diszlexia, az ismétlési készség megőrzöttsége, a beszéd motoros kivitelezése hibátlan. A **„logopenic"** típus jellemzői: az izolált szavak előhívásának nehézsége, kifejezések és mondatok ismétlésének nehezítettsége, három vagy több az alábbiak közül: ejtéshibák, megőrzött motoros beszédkivitelezés, az izolált szavak megértése nem sérül, tárgyfelismerés jó, az agrammatizmus hiánya. A **nem fluens, agrammatikus** típus jellemzői: grammatikai egyszerűsítés és grammatikai hibák megléte, nehezített beszéd, gyakori megállások, ejtéshibák, kettő vagy több a következők közül: a szintaktikai szerkezetek megértésének sérülése, a tartalmas szavak megértése és a tárgyak felismerése megőrzött.

Tekintettel az elsődleges progresszív afázia kórképének változékonyságára, újabb és újabb adatok alapján igyekszenek azt tovább pontosítani. A nonfluens típusban jellemzőnek ítélik a fonémikus parafáziákat, a dysarthriát, va-

lamint a diszprozódiát is (Grossmann–Ash 2003). A beszédtempó csökkenését és nagyobb hezitációs szünetek megjelenését találták további betegeknél (Thompson et al. 1997). A hosszabb szavak ejtésében gyakoriak az artikulációs nehézségek. Úgy tűnik, hogy a betegség első időszakában a betegek fluensek, a beszéd folyamatossága nem érintett, később azonban egyre kevésbé képesek folyamatosan beszélni (Mesulam 2001). Minthogy az elsődleges progresszív afázia a nyelvi funkciók működését érinti negatívan, így előjelezhető következményei vannak az írott nyelvre is. A hallási-verbális, valamint a szemantikus memória működésének korlátozottsága, a vizuális-percepciós folyamatok bizonytalan működése, valamint a mondatfeldolgozás nehezítettsége mind-mind az olvasás romlásához vezet (Grossman–Ash 2004). Feltűnő az angolban az irreguláris ejtésű szavak olvasásának zavara, valamint a betűhang megfeleltetés ejtési szabályainak alkalmazási képtelensége (Noble et al. 2000). Úgy gondolják, hogy mindez a szavak vizuális formájának modalitásspecifikus deficitjével van kapcsolatban.

Elsődleges progresszív afáziával diagnosztizált magyar anyanyelvű betegről nem találtunk publikációt. A ritka kórképek azonosítását éppen előfordulásuk alacsony száma nehezíti, ezért az esettanulmányoknak nagy a jelentőségük. Kutatásunkban egy olyan magyar anyanyelvű beteg mondatismétléseinek és hangos olvasásának sajátosságait mutatjuk be, akinek nyelvi és beszédzavarai számos klinikai vizsgálatot követően két éve fennállnak anélkül, hogy egyértelmű diagnózisa lenne. Vizsgálatsorozatunk célja a beteg beszédének akusztikai-fonetikai elemzése, és az objektív adatok értelmezése alapján a kórkép lehetséges azonosítása. Hipotézisünk az volt, hogy adatközlőnk betegsége az elsődleges progresszív afázia lehet, azon belül is a *logopenic*nek nevezett típusba sorolható. Ismereteink szerint ilyen beszédzavart mutató beteg beszédének részletes akusztikai-fonetikai vizsgálatát még nem végezték el. A mondatismétlés és a hangos olvasás elemzését azért választottuk, mert ebben a két helyzetben egyszerűsödik a rejtett beszédtervezési mechanizmus (szemben a spontán beszéddel), így feltételezhetően könnyebb az ok-okozati összefüggések megtalálása. Úgy gondoljuk, hogy eredményeink hozzájárulnak adatközlőnk egyértelmű diagnosztizálásához. Az elemzési eredmények segíthetnek továbbá a nyelvspecifikus terápia kialakításában.

Kísérleti személy, anyag, módszer

Adatközlőnk 57 éves, diplomás nőbeteg (a továbbiakban Z.), akinek anamnézise szerint beszéde vizsgálatainkat megelőzően két évvel kezdett megváltozni. Először angol nyelven érzékelte a problémát, nem jutottak eszébe szavak, és nehezen formálta az angol beszédhangokat. Nem sokkal ezt követően anyanyelvén is jelentkeztek a tünetek: nem jutottak eszébe, illetőleg torlódtak a szavak, kiejtése bizonytalanná vált, sokat hibázott, hangja rekedtes lett, megnyilatkozásai nélkülözték a beszéddallamot, beszédtempója lassult. Be-

szédében folyamatos romlást vesz észre, írásban, olvasásban nem tapasztal zavart. Általános egészségi állapotára vonatkozó panasza nincsen.

Z. 2009-től számos orvosi, neuropszichológiai és logopédiai vizsgálaton esett át, ezek egyike sem adott megfelelő magyarázatot állapotára. A 2010. júniusi diagnózisa szerint a diszprozódia hátterében szorongás áll. 2011-ben és 2012-ben végzett újabb vizsgálatok demenciát, anómiát nem igazoltak, beszédében az absztrakció, a kategorizáció megtartott, a szövegértés, az ismétlés, az olvasás és az írás nem mutatott eltérést a normáltól; a pszichés funkciók, a memória, a gondolkodás és a számterjedelem ép, IQ-értéke 119. Diagnózisai között szerepelt a motoros afázia, amelyet a további vizsgálatok eredményei alapján megkérdőjeleztek, majd elvetettek. A legfrissebb klinikai diagnózis dysarthriát (a beszédmozgások koordinációs kivitelezésének zavara), és anarthriát (a szavak tagolt kiejtésének hiánya, a beszédmozgásokért felelős motoros képesség elvesztése) diagnosztizáltak.

Adatközlőnk azzal a céllal érkezett intézetünkbe, hogy a klinikai diagnózis kiegészítéseként elkészítsük (2011 őszére már) jól hallható beszédzavarának akusztikai-fonetikai elemzését. Negyedévente, összesen három alkalommal rögzítettük adatközlőnkkel a BEA-protokollt (Gósy 2012), amely különböző típusú spontán beszédet, társalgást, felolvasást és mondatismétlést tartalmaz. A jelen kutatáshoz az utóbbi két részt elemeztük. Az adatközlőnek (i) 25 változatos grammatikai szerkezettel, szórenddel és koartikulációs szabályokkal rendelkező egyszerű és összetett mondatot kellett megismételnie, és (ii) egy 234 szóból álló tudománynépszerűsítő cikket felolvasnia. Az akusztikai elemzésekhez a hanganyagot a Praat szoftverrel (Boersma–Weenink 2010) annotáltuk. Vizsgáltuk az alaphangmagasságot, a zöngeminőséget (a jittert, a shimmert, a jel/zaj viszonyt, a glottalizáció mértékét), illetőleg a beszéd- és artikulációs tempót, továbbá a megakadásjelenségek előfordulását. Az alaphangmagasságot automatikusan mértük ki úgy, hogy a programmal beszédszakaszonként (két néma szünet által határolt szakasz) minden 10 ms-ban kiírattuk az F_0-átlagot, folyamatos auditív és vizuális ellenőrzés mellett. A kapott értékek közül kiszűrtük a mérési hibából adódó kiugró szélső értékeket, majd meghatároztuk az adott szakasz teljes időtartamára számolható alaphangmagasság-átlagot. A jitter és a shimmer, valamint a jel/zaj viszony meghatározásához hasonló elven működő automatikus mérést alkalmaztunk. A glottalizáció mértékét szótagszinten határoztuk meg, mondatonként, illetve beszédszakaszonként úgy, hogy a glottalizált szótagok számát elosztottuk az adott szakasz teljes szótagszámával. Az adatok statisztikai elemzését egytényezős varianciaelemzéssel (ANOVA), Tukey-féle post-hoc teszttel, illetve a nem-parametrikus Kruskal–Wallis-teszttel és kereszttábla-elemzéssel végeztük 95%-os konfidenciaszinten az SPSS 13.0-as szoftver segítségével.

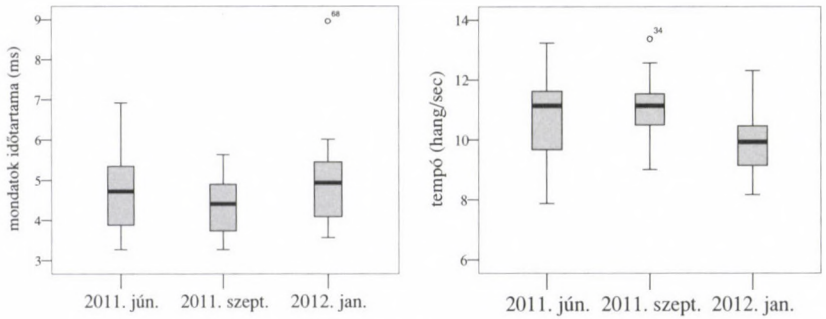
Eredmények

Mondatismétlés

Mondatok ismétléséhez a beszédprodukciós mechanizmus fonológiai és fonetikai tervezési szintjének és a munkamemóriának a megfelelő működése, valamint a mentális lexikon elvárt aktiválása szükséges. Ugyanezen beszédtervezési folyamatok működtetésére van szükség hangos olvasáskor, azokat a vizuális dekódolás és a graféma-fonéma megfeleltetés előzi meg.

Az egyes akusztikai-fonetikai jellemzők különböző mértékben változtak a három felvétel alkalmával. A statisztikai elemzéseket normalitásvizsgálat előzte meg (Shapiro–Wilk-teszt), amelynek eredményei szerint az elemzett paraméterek közül az alaphangmagasság átlagai és maximumai, a shimmer értékei, a jel/zaj viszony és a mondatok tempója volt normál eloszlású.

Az első felvételkor az ismételt mondatok átlagidőtartama 4,6 s, a második felvételnél 4,4 s, a harmadiknál pedig 4,9 s (1. ábra); az időtartamátlagok nem mutattak szignifikáns különbséget. Az időtartamokat elsősorban a mondat szavainak, szótagjainak, illetve beszédhangjainak száma határozta meg, a szintaxisnak, valamint a jelentésnek nem volt hatása. A tendencia az, hogy az első felvétel ismétléseihez képest a másodiknál csökkennek az időtartamok, a harmadiknál pedig növekednek.

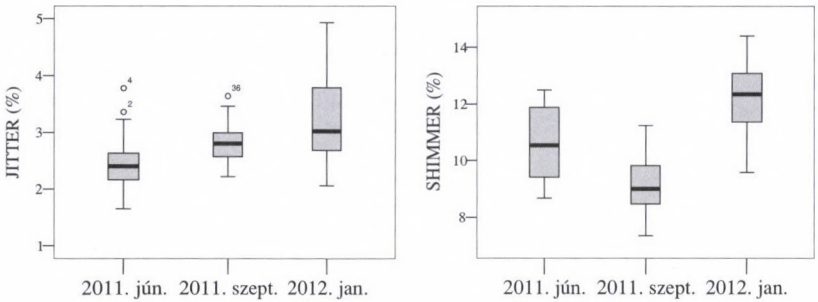


1. ábra
Az egyes mondatok ismétlésének időtartamai (bal oldalon) és a beszédtempók (jobb oldalon) a három felvétel esetében

A legrövidebben és a leghosszabban ejtett mondatok között nagyok a különbségek, a minimumértékek 3 másodperc körüliek, a maximumértékek 6,9 s az első, 5,4 s a második és 8,9 s a harmadik felvétel esetében. A beszédtempó adatait az 1. ábra jobb oldali képe szemlélteti. Ennek eredményei szignifikáns különbséget igazoltak a felvétel függvényében [egytényezős ANOVA: $F(2) = 6{,}023$, $p = 0{,}040$]. Az első és a második felvétel között a Tukey post hoc teszt nem mutatott szignifikáns eltérést, az átlagok 10,8 hang/s (átl. elt.: 1,4 hang/s) az első és 11,1 hang/s (átl. elt.: 1 hang/s) a második esetben. A harmadik felvé-

telkor a beszédtempó azonban nagyobb mértékben csökken, 9,9 hang/s az átlag (átl. elt.: 1 hang/s). Szó/percre átszámolva, ez azt jelenti, hogy percenként 9 szóval ejtett kevesebbet a beszélő az utolsó alkalommal a megelőzőhöz képest, ami jellegzetes lassulást igazol.

A hangszalagműködés sajátosságait négy paraméter, a jitter, a shimmer, a jel/zaj viszony (Gelfer–Fendel 1995; Gósy 2004; Wertzner et al. 2005; Brockmann et al. 2008) és a glottalizáció (Slifka 2006; Markó 2012) elemzésével jellemeztük. A jitter értékét százalékban adják meg, ami arra utal, hogy milyen a hangszalagrezgések frekvenciaingadozásának mértéke. Minél alacsonyabb ez az érték, annál jobb a zöngeminőség. A hangszalagrezgések amplitúdóingadozásának mértéke a shimmer (%-ban kifejezve), minél alacsonyabb az érték, a zöngehang annál tisztább csengésű. A zönge és a levegőáramlási zörej arányát jellemzi a jel/zaj viszony (dB-ben kifejezve). Minél nagyobb a jel/zaj viszony értéke, annál tisztább a primér hang, vagyis a zöngehang annál kevesebb zörejelemet (ill. súrlódó levegőt) tartalmaz. Adatközlőnk zöngeműködése változik a felvételi alkalmak függvényében; az első és a harmadik felvétel között egyértelmű romlás tapasztalható. A shimmer a második felvétel esetében javulni látszik (az átlagérték 10,58%-ról 9,11%-ra csökken; az átlagos eltérések: 1,22% és 0,98%), tehát az amplitúdóingadozás pozitívan változik, de a harmadik felvételkor nagymértékű növekedés látható, amelynek átlaga 10,09%, az átlagos eltérés értéke 1,35% (2. ábra).

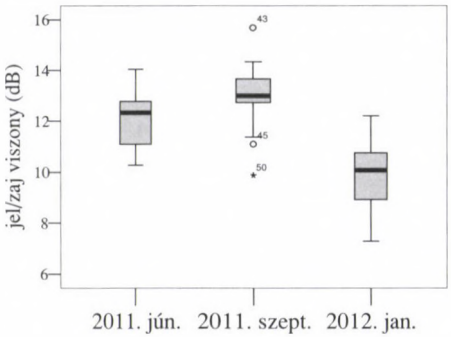


2. ábra
A jitter és a shimmer értékeinek alakulása a felvételek során

A jitter átlagértéke az első felvételkor 2,44% (átl. eltérés: 0,49%), a másodiknál 2,8% (átl. elt.: 0,36%), harmadiknál pedig 3,29% (átl. elt.: 0,77%). Az adatok nem normál eloszlásúak, ezért a statisztikai elemzéseket nemparametrikus teszttel végeztük, ezek eredménye szignifikáns különbséget igazolt a felvételek függvényében [$\chi^2(2) = 22{,}031$, $p = 0{,}001$]. A shimmer és a jel/zaj viszony mért adatai normál eloszlásúak, az egytényezős ANOVA mindkét esetben ugyancsak szignifikáns különbséget mutatott a felvételek függvényé-

ben [shimmer: $F(2, 74) = 38{,}688$, $p = 0{,}001$; jel/zaj viszony: $F(2, 74) = 46{,}587$, $p = 0{,}001$]. A Tukey post hoc tesztek minden felvétel között szignifikáns eltérést igazoltak.

A jel/zaj viszony értékei kismértékben javulnak a második felvételkor; az átlagértékek 12,1 dB (átl. elt.: 1,1 dB), illetve 13 dB (átl. elt.: 1,1 dB). A harmadik felvételnél azonban nagymértékű romlást tapasztaltunk abban az értelemben, hogy a zörejelemek aránya erősen megnövekedett, így a jel/zaj viszony átlagértéke jelentősen csökkent, 9,8 dB-re (átl. elt.: 1,3 dB) (3. ábra).

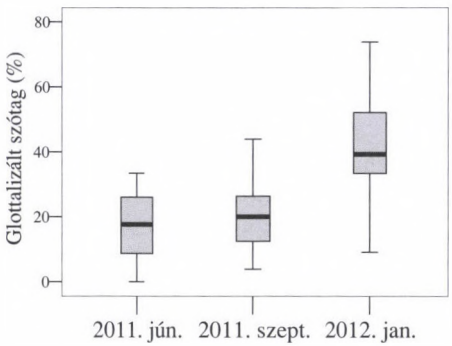


3. ábra
A jel/zaj viszony értékei a három felvétel mondatismétléseiben

A minimum- és maximumértékek a zörejelemek növekedésére utalnak. Az első felvételkor a legkevésbé és a legnagyobb mértékben zörejes mondat között 3,8 dB-t mértünk, a második felvételkor 5,8 dB-t, míg a harmadik felvételkor 4,9 dB-t. Mindhárom felvételnél jelentősek az eltérések az egyes mondatok ejtésében a paraméterek szórásában. Mind a jitter, mind a shimmer maximum- és a minimumértékeinek különbsége a harmadik felvételnél volt a legnagyobb (2,87% az előbbinél és 4,8 dB az utóbbinál). Hozzátesszük, hogy mindkét paraméter esetében már a kiinduló értékek is lényegesen rosszabbak voltak, mint amit ép zöngeképzés esetén tapasztalunk (l. Gósy 2004; Brockmann et al. 2008). A zöngeműködést jellemző paraméterek közül legnagyobb mértékben az amplitúdóingadozás és a jel/zaj viszony értékei magyarázzák a változást, esetünkben a romlást, mintegy 50%-ban (a parciális η^2 értékei: 51,5% és 56,4%).

Glottalizációkor a zönge képzése – szándékos vagy a beszélő akaratától független következményként – eltér a megszokottól, vagyis irreguláris rezgéseket tartalmaz. A hallgató ezt rendszerint rekedtes hangként észleli. Az irreguláris zönge meglehetősen átfogó meghatározás, amely a primér hang többféle típusát tartalmazza (Markó 2012). Komplex jelenség, amely elsődlegesen a hangszalagok sajátos működési eredményét jelenti, de amelynek létrejötté-

hez az alaphangmagasság változása és egyéb tényezők is hozzájárulhatnak. A glottalizáció nagymértékben függ a beszélőtől (Slifka 2006). Idősebb nők glottalizált szótagjainak aránya (az összes ejtett szótaghoz viszonyítva) közel 30% spontán beszédben és néhány százalékkal kevesebb felolvasásban (Markó 2012). A mondatismétlések esetében kisebb arányú glottalizáció várható el. Z. ismétléseiben növekvő mértékben fordultak elő glottalizált beszédrészletek, leggyakrabban a harmadik felvételkor (több mint 10%-kal meghaladta a spontán beszédre közölt átlagos arányt). Ez adatközlőnk beszédének erőteljesen rekedtes színezetet adott. A statisztikai elemzés szignifikáns különbséget igazolt a felvételek függvényében [$\chi^2(2) = 31{,}664$, $p = 0{,}001$]. A glottalizált szótagok előfordulásának átlaga 3,52 db az első, 4,32 db a második és 9,2 db a harmadik felvételben. Az arányokat tekintve ezek a darabszámok 17,61%-nak (átl. elt. 9,22%), 20,52%-nak (átl. elt. 10,95%), illetve 41,16%-nak (átl. elt. 17,43%) felelnek meg (4. ábra). A minimum- és maximumértékek különbsége az első felvételkor 33,3%, ekkor még akadt olyan ismételt mondat, amelyben egyáltalán nem tapasztaltunk glottalizációt. A második felvételkor minden mondatban hallható volt, a minimum- és maximumértékek különbsége 40,15%, a harmadik felvételkor a legkevesebb előfordulás az előző felvétel minimumértékének a háromszorosa volt, a különbség pedig 64,82%.



4. ábra
Glottalizált szótagok aránya a három felvétel mondatismétléseiben

A mondathossz, a szószám, illetve a tartalom nem mutat semmilyen összefüggést a glottalizálás megjelenésével. Két mondat tért el jelentősen e tekintetben, és mindössze egyetlen olyan adódott, amelynél a glottalizált szótagok száma hasonló volt. *A tegnap esti repülőszerencsétlenséget senki sem élte túl* mondat (8 szó, 48 beszédhang) az első felvétel sokszor, a második alig tartalmazott glottalizált szótagot, míg *Az ügyfeleknek kompromisszumot kellett kötniük* (5 szó, 37 beszédhang) mondatban az első felvételkor alig, a harma-

diknál gyakran hangzott glottalizálás. *A hegyi mentőknek sűrű ködde*l *kellett megküzdeniük* mondatban (7 szó, 40 beszédhang) mind az első, mind a harmadik felvételben gyakori volt a glottalizálás. Ugyanannyi szót és beszédhangot tartalmazó mondatok az egyik felvételkor alig, a másik alkalommal pedig igen sok glottalizálást tartalmaztak, például *A magyar űrturista kalandjait az egész országban figyelték* (8 szó, 47 beszédhang). A 10 és 11 szót tartalmazó mondatokban egyik felvételben sem hangzott jelentős mértékű glottalizálás, és csupán a harmadik felvételnél tapasztaltuk, hogy gyakoribbá válik az ismétlendő mondatok számának függvényében.

A korrelációelemzés közepesen erős összefüggést igazolt a glottalizáció és a jitter ($r = 0{,}67$), a shimmer ($r = 0{,}55$), a jel/zaj viszony ($r = 0{,}56$), valamint a shimmer és a jel/zaj viszony ($r = -0{,}69$) között; továbbá a jitter és a shimmer ($r = 0{,}54$), valamint a jitter és a jel/zaj viszony között ($r = -0{,}48$).

A mért adatok a beszéddallam beszűkülését, a mondatok ismétlésének erősen monoton jellegét igazolják (1. táblázat). Egy-egy mondat ismétlésekor az alaphangmagasság frekvenciatartománya 40-50 Hz körüli; gyakran ez is csak annak köszönhető, hogy a mondat első szótagja magasabb F_0-értéken realizálódik a többinél, de a továbbiakban alig mutatkozik eltérés.

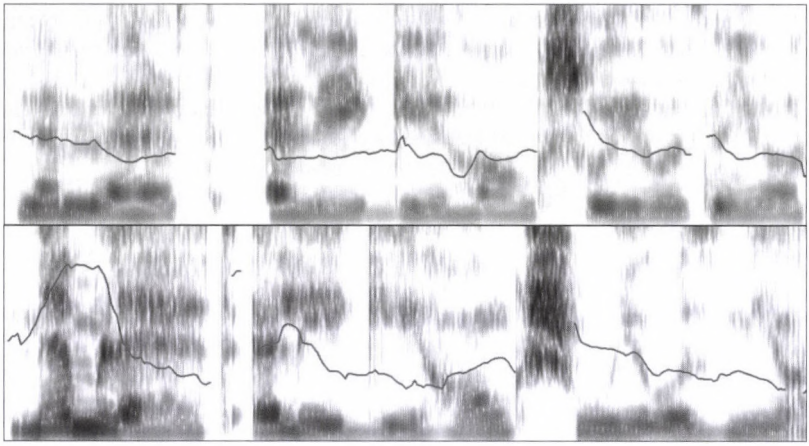
1. táblázat: Az alaphangmagasság jellemzői a mondatismétlésekben

Felvételek	Az alaphangmagasság akusztikai jellemzői (Hz)		
	Átlag	Átlagos eltérés	F_0-tartomány
Első	176	6	51
Második	171	5	45
Harmadik	176	6	37

Az 5. ábra egyetlen mondat hangszínképét és a rárajzoltatott dallamvonalat szemlélteti. Ugyanannak a mondatnak az ejtéséről készült regisztrátum az interjúvezető produkciójában is látható, amely szemlélteti a beszéddallam különbségeit a két beszélő között. Az alaphangmagasság változása szignifikánsan különbözik a felvételek sorszámától függően [egytényezős ANOVA $F(2, 74) = 6{,}119$, $p = 0{,}004$], bár a Tukey post hoc teszt szerint az első és a harmadik felvétel között nincs szignifikáns különbség. A monoton beszéddallam mindhárom felvételre jellemző.

Elemeztük a szünetek és az ejtési hibák sajátosságait. A szünetek jelezhetik az emlékezeti működésekhez szükséges időket, és így az ismétlési nehézség mutatói is. A szünetek darabszáma felvételenként növekszik, az átlag az első alkalommal 1,88 db (átl. elt.: 0,83 db) volt, a másodiknál 2,04 db (átl. elt.: 1,09 db), a harmadiknál pedig 2,68 db (átl. elt.: 1,24 db); a különbség statisztikailag szignifikáns [$\chi^2(2) = 6{,}057$, $p = 0{,}048$]. A szünetek maximális előfordulása is növekedett felvételenként; az elsőnél 3 db, a másodiknál 5 db, a harmadiknál 6 db volt mondatonként. Az első felvételben továbbá összesen

hét kitöltött szünetet (hezitálást) adatoltunk, a másodikban tizenegyet, a harmadikban pedig tízet.

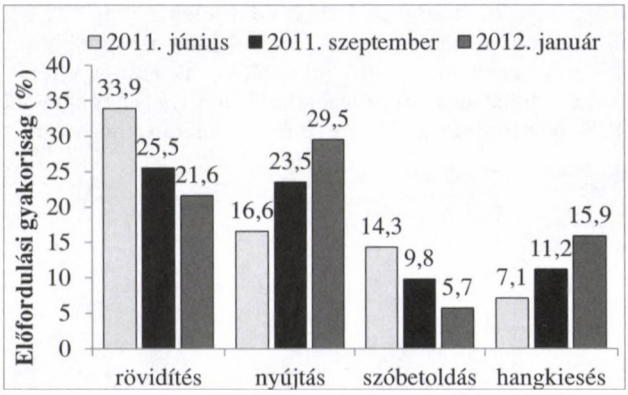


5. ábra
A *Nem lehet eléggé hangsúlyozni a túlzásba vitt napozás veszélyeit* mondat az adatközlő ejtésében (felső kép) és az interjúvezető ejtésében (alsó kép)

Adatközlőnk mondatismétlései sajátos megakadásokat tartalmaztak, a létrejöttük oka valószínűleg más, mint a mindennapi spontán közlésekben előfordulóké (Gósy 2005; Gyarmathy 2010). Ezért a továbbiakban összefoglalóan „ejtéshibáknak" nevezzük őket, bár a terminust sajátosan értelmezzük a mondatismétlés műveletére. Az ejtéshibák egy része bizonytalanságra utalt, más részük hibának volt tekinthető, amelyek a beszédhangok képzési konfigurációjának megváltozását is jelentették. A leggyakoribbak a nyújtások, a rövidítések, a szóbetoldások és a hangkiesések voltak (6. ábra), ezek eredményezték a hibák mintegy 70%-át (felvételenként 71,9%, 70% és 72,7% az összes hiba arányában).

Egy vagy néhány alkalommal adatoltuk a következő hibákat: téves kezdés (*speti* /szünet/ *specialitása*), hangbetoldás (*vulkánkitörés* helyett *vulkántkitörés*), perszeveráció (*buszbaleset* helyett *buszbalaset*), anticipáció (*túlzásba* helyett *túlzasba*), szünet a szóban (*special* /szünet/ *itása*), szókihagyás (egy-egy mondatvégi, illetve mondat eleji szó ismétlésének képtelensége), zöngésítés (pl. *betegzség betegség* helyett), zöngétlenítés (*galériából* helyett *kaléria-ából*). A szócserék (pl. *vidéki galériából* helyett *külföldi galériából*), avagy egyes grammatikai változtatások, például toldalékos szó ejtése névutó helyett (*akció során* helyett *akciónál*) feltehetően abból (is) adódnak, hogy a beszélő a kiejtésre fordítja minden figyelmét, így nem mindig rögzül pontosan a hallott mondat. A kontextushatással nem magyarázható beszédhangcserék való-

színűleg nem nyelvbotlások (*kirándulni* helyett *kirándulnyi* vagy *tilos* helyett *tilas*), hanem a beszédprobléma megjelenési formái.



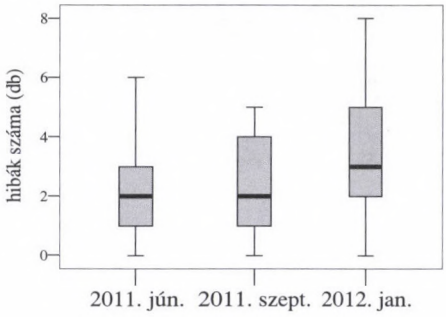
6. ábra
Az „ejtéshibák" négy típusának arányai felvételenként

A koartikuláció elmaradása (pl. *sérültjeinek* vagy *buszbaleset* esetében) adatközlőnk iskolázottságát és foglalkozását tekintve ugyancsak a beszédprobléma tüneteként értelmezhető. Az artikulációs gesztusok szabályozásának átmeneti blokkolódása időnként értelmetlen hangsorok ejtését eredményezte, például az *alakultak* szó ismétlésekor az alábbi hangsorokat produkálta: *alkala* (szünet) ***kaa****gut* (a félkövérített betűkettőzés a nyújtást jelzi).

A hibák sokasága akcentushatást kölcsönöz adatközlőnk kiejtésének. Ez különösen szembetűnő a három alsóbb nyelvállású magánhangzó nem normatív ejtésekor. Az összes hiba száma az első felvételben 55 volt, a másodikban igen kis mértékben csökkent (51 db), a harmadikban ismét megnövekedett (88 db), lásd a 7. ábrát.

A mondatokban előforduló összes szót tekintve, a hibák aránya 26,9%, 25,0% és 43,1% felvételenként. Az időre vetítve, adatközlőnk az első felvételkor 29,2 db hibát ejtett percenként, a második alkalommal 27,6-ot, a harmadik felvétel ismétléseiben pedig percenként 42,5 darabot. A statisztikai elemzés szignifikáns összefüggést igazolt a felvételek és a hibák száma között [$\chi^2(2) = 6{,}296$, $p = 0{,}043$]. Keresztтábla-elemzéssel tovább vizsgáltuk a hibaszám változásait a felvételek során. A hibátlan mondatismétlések rendkívül ritkák, legnagyobb arányban a második felvételkor voltak tapasztalhatók, az első és a harmadik felvétel között nem volt különbség. Egyetlen hiba az első és a második felvételre volt jellemző azonos arányban, a harmadik felvételben alig volt ilyen. Két hiba az első és a harmadik felvételben hasonlóan, három hiba valamennyi felvételben azonos arányban jelentkezett. Négy hiba

a második felvételben fordult elő nagymértékben, öt hiba döntően a harmadik felvételre volt jellemző. Hat-hat hibát az első és a harmadik felvételben adatoltunk, hét és nyolc hiba csupán az utolsó alkalommal fordult elő. Az első felvételben legnagyobb arányban 1 és 2 hiba jelenik meg (28% és 32%), a második alkalommal a hibátlan és az 1 és a 4 hiba a nagyobb arányú (20%, 28% és 24%). A harmadik felvétel hibaszámainak aránya eltérő mintázatot mutat. A 2 és 5 hibát tartalmazó mondatok aránya a legnagyobb (24% és 20%), 12-12%-ban fordulnak elő az 1, 3 és 7 hibát tartalmazó ismétlések.



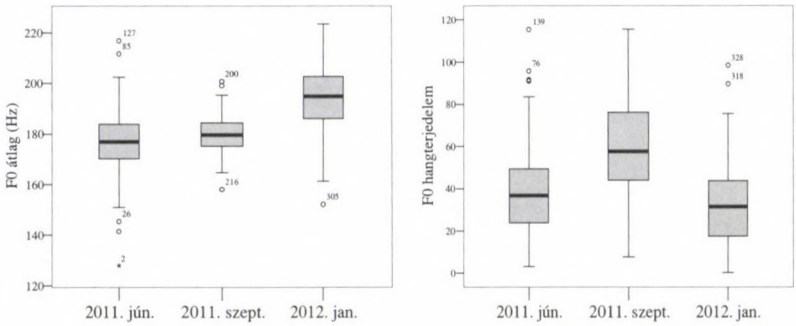
7. ábra
A hibák száma a mondatokban felvételenként

A négy leggyakoribb megakadásjelenség előfordulása jellegzetes eltéréseket mutat az egyes felvételekben. Az első mondatismétlés-sorozatra a beszédhangok feltűnően rövidebb ejtése volt jellemző (*vádlot* a *vádlott* helyett vagy *mikorra* helyett *mikora*, avagy *igy* az *így* helyett, a legalsó nyelvállású magánhangzó rövid időtartamban), a hangkiesés (pl. *szerencsétlenü* a *szerencsétlenül* helyett) elenyésző mértékben jelentkezett. A második felvételkor a rövidítés és a nyújtás (pl. ***sszilveszterkor*** vagy ***estíí*** az *esti* helyett) hasonló arányban volt tapasztalható, legkisebb arányban a szóbetoldások (főként a *hogy* kötőszó) szerepeltek. A harmadik felvételben a nyújtást adatoltuk a legnagyobb arányban, a szóbetoldás ekkor elenyésző mértékű volt. Ezek az eltérések eléggé szabályosnak mutatkoznak: ahogy az adatközlő beszédprodukciója romlik, úgy a rövidítést gyakoriságban felváltja a nyújtás, utalva a fokozott időnyerési szükségletre, valamint a hangkiesés, bár a beszédhangoknak a kívántnál rövidebb ejtése továbbra is jellemzően előfordul. A hibamintázat az időzítés egyre nagyobb mértékű felbomlását támasztja alá. Ugyanezen mondatok hibátlan ismétlésének átlaga fiatal beszélőknél 84%-os, az idősek (68–80 évesek) átlaga pedig 66%-os volt (Neuberger 2010). Adatközlőnk hibátlan ismétléseinek átlaga az első és a harmadik alkalommal 8-8%-os, a második alkalommal 20%-os volt.

Felolvasás

A statisztikai elemzéseket megelőző normalitásvizsgálat alapján a felolvasásban vizsgált akusztikai paraméterek közül az alaphangmagasság átlagai, minimumai és maximumai, a szünetek hossza, a jitter és a shimmer értékei, valamint a jel/zaj viszony bizonyultak normál eloszlásúnak.

A beszélő alaphangmagassága a három felvétel során eltérően alakult, a felolvasásnál egyértelmű F_0-növekedés volt kimutatható. Az egyes felvételek közötti eltérést a statisztikai elemzések is megerősítették: $F(2, 457) = 111{,}926$; $p = 0{,}001$ (8. ábra). Az első alkalommal Z. átlagos alaphangmagassága 177 Hz volt (átl. elt.: 12 Hz), a második felvételen 180 Hz (átl. elt.: 7 Hz), míg a harmadiknál 195 Hz (átl. elt.: 12 Hz). A Tukey post hoc teszt szerint az első és a második felvétel között nincs szignifikáns különbség.

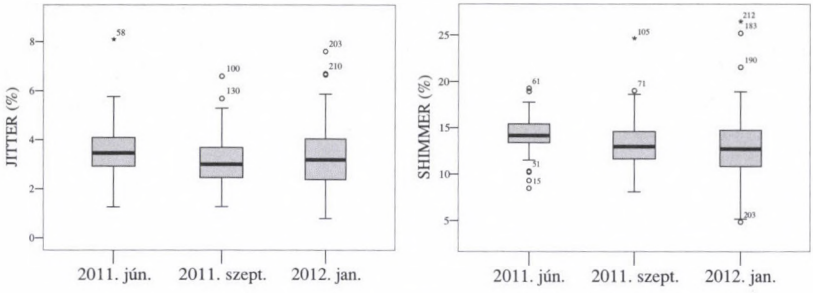


8. ábra
Az alaphangmagasság átlagának és frekvenciatartományának változása a felolvasásban felvételenként

Az F_0 maximum- és minimumértékeinek különbsége jól mutatja a beszéddallam szűkülését, monotonabbá válását. Az F_0 frekvenciatartománya az első felolvasáskor átlagosan 38 Hz volt (átl. elt.: 19 Hz), a második alkalommal kevésbé monoton, 58 Hz (átl. elt.: 25 Hz), míg a harmadiknál mindössze 32 Hz (átlagos eltérés 18 Hz). Az F_0 hangterjedelme a felvételek sorszáma szerint szignifikánsan különbözik [egytényezős ANOVA: $F(2, 457) = 57{,}264$, $p = 0{,}001$, vö. 8. ábra], a Tukey post hoc teszt pedig igazolta az egyes csoportok egymáshoz viszonyított szignifikáns eltérését (első és második, ill. második és harmadik felvétel: $p = 0{,}001$; első és harmadik felvétel: $p = 0{,}017$).

A hangszalagműködés sajátosságait jellemző négy paramétert (jitter, shimmer, jel/zaj viszony, glottalizáció) a felolvasásban is részletesen elemeztük. A jitter, azaz a hangszalagrezgések frekvenciaingadozásának mértéke az első felvételben 3,58% (átl. elt.: 1,13%), a második felvételben 3,19% (átl. elt.: 1,01%), a harmadik felvételben 3,37% (átl. elt.: 1,31%). Az elsőhöz képest a

második alkalommal kiegyenlítettebb volt a zöngeminőség, a harmadik alkalomkor azonban romlás volt tapasztalható (9. ábra). A felvételek közötti különbség matematikailag nem igazolható. A hangszalagrezgés amplitúdóingadozásának mértéke, azaz a shimmer a felolvasásokban javulást mutat. Az első felvételen 14,26% volt az átlagérték (átl. elt.: 1,96%), a másodikon 13,34% (átl. elt.: 2,41%), a harmadikon pedig 13,07% (átl. elt.: 3,45%).



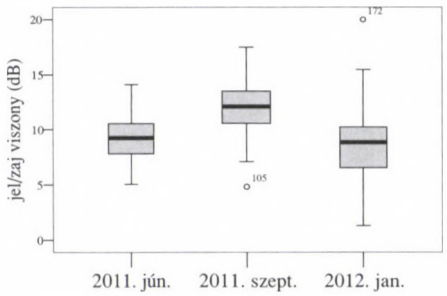
9. ábra
A jitter és a shimmer értékeinek alakulása a felolvasásban

Az egytényezős ANOVA igazolta, hogy a három felvétel szignifikánsan különbözik egymástól [$F(2, 219) = 3{,}596$; $p = 0{,}029$]. A Tukey post hoc teszt azonban csak az első és a harmadik felvétel shimmerértékei között igazolt szignifikáns különbséget ($p = 0{,}026$). A legnagyobb szórástartományt mindkét elemzett paraméter esetében a harmadik felvételnél adatoltuk: jitter 6,83%, shimmer 21,61%.

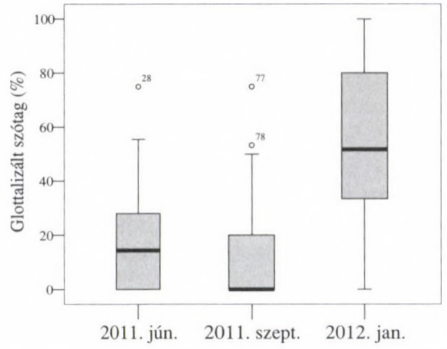
Az olvasás zörejkomponenseinek aránya a második felvételen az elsőhöz képest kismértékű javulást mutat – az előbbi átlagértéke 11,8 (átl. elt.: 2,3 dB), az utóbbié 9,3 dB (átlagos eltérés 2,1 dB). A harmadik felvételen a zörejelemek aránya jelentősen megnőtt, a jel/zaj viszony átlagértéke csökkent, 8,5 dB (átl. elt.: 2,9 dB) (10. ábra). A zörejösszetevők növekedését jelzi az is, hogy növekszik a minimum- és maximumértékek által behatárolt szórástartomány; az első felvételen 9,1 dB, a másodikon 12,7 dB, a harmadikon pedig 18,67 dB. Az egytényezős ANOVA az egyes felvételek függvényében szignifikáns különbséget igazolt [$F(2, 219) = 34{,}831$; $p = 0{,}001$]. A Tukey post hoc teszt szerint az első és a második, illetve a második és a harmadik felvétel szignifikánsan eltér egymástól ($p < 0{,}001$).

A glottalizált ejtés a tipikus beszélőkre is jellemző, főleg szakaszvégen (Markó 2012). Adatközlőnk beszéde azonban jól hallhatóan erősen rekedtes, így feltételeztük, hogy nála nagyobb arányban fordulnak elő glottalizált beszédrészletek, mint az átlagos beszélőknél. A felolvasásban az első felvételen a glottalizált szótagok darabszámának átlaga 1,87 db, a másodikon 1,24, a harmadikon 3,88. A glottalizált szótagok aránya az összes szótaghoz viszo-

nyítva az első alkalommal 17,52% volt (átl. elt.: 18,45%), a második felvételen kismértékű a csökkenés, az átlagérték 13,32% (átl. elt.: 20,55%). A harmadik felvételen az arány jelentősen megnövekedik; az összes szótag 54,67%-a irreguláris zöngével valósul meg (átl. elt.: 31,75%) (11. ábra). A Kruskal–Wallis-próba szignifikáns különbséget igazolt a glottalizált szótagok számában a felvételek függvényében ($\chi^2(2) = 83{,}172$, $p < 0{,}001$).



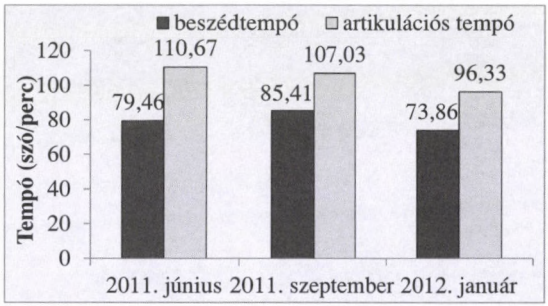
10. ábra
A jel/zaj viszony mértéke a felolvasásban felvételenként



11. ábra
A glottalizált szótagok aránya a felolvasásban felvételenként

A szórástartomány növekszik. Noha mindhárom felvételen előfordult olyan beszédszakasz, amely egyáltalán nem tartalmazott irreguláris zöngét, ezek aránya felvételenként csökkent (az első felvételen még 39,13%, a harmadikon csupán 10%). Az első két felvételnél adatolt maximumérték 75% volt, míg a harmadik felvételen már előfordult olyan beszédszakasz is (az összes beszédszakasz 16,66%-a), amelyben az összes szótag glottalizált ejtéssel realizálódott.

A tudománynépszerűsítő szöveg felolvasása mindhárom alkalommal közel 3 percet vett igénybe. Az első alkalommal 3 percig és 4 másodpercig tartott a felolvasás, másodszorra 2 perc 49 másodpercig, a harmadik felvétel 3 perc 20 másodpercig tartott. Ez a tendencia jelzi a beszélő fokozódó kivitelezési nehézségeit; a beszéd- és artikulációs tempót elemezve ez a kép tovább árnyalható (12. ábra).



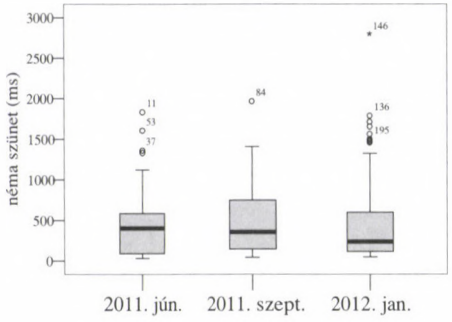
12. ábra
A beszéd- és az artikulációs tempó alakulása a felolvasásban felvételenként

A felolvasás esetében mindkét tempóértéket szó/percben adjuk meg a jobb szemléltethetőség érdekében. A beszédtempó (adott idő alatt ejtett beszédjelek és néma szünetek) értékei csökkenést igazolnak a felvételek sorszámának függvényében, és az artikulációs tempó (beszédjelek képzésének sebessége) értékei is alátámasztják a feltűnő lassulást.

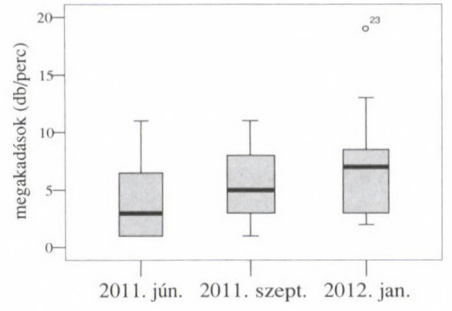
A beszéd folyamatosságát alapvetően meghatározza az a tény, hogy a beszélő milyen gyakran tart szünetet, és azok milyen hosszúságban realizálódnak. Az első felvételben 65 db néma szünetet adatoltunk, a másodikban 62 dbot, míg a harmadikban összesen 90-et. Az átlagos szünetidőtartam mindhárom esetben közelített az 500 ms-hoz, ami relatíve hosszúnak tekinthető. Hírolvasásokban a néma szünetek értékei átlagosan 150–200 ms közé estek, noha itt a bemondók gyakorlott beszélők voltak (Rákli 2009). Az első alkalommal az átlagos szünetidőtartam 466 ms (átl. elt.: 415 ms), a másodiknál 483 ms (átl. elt.: 434,5 ms), míg a harmadik esetben 470 ms (átl. elt.: 539 ms) volt (13 ábra).

A szünetidőtartamok növekedését jól szemléltetik a néma szünetek minimum- és maximumértékei is. A legrövidebb szünet az első felvételen 33 ms, a leghosszabb 1833 ms volt, a második esetben a minimumérték 43 ms, a maximum 1965 ms volt, a harmadik alkalommal 49 ms és 2800 ms között szóródtak az értékek. A három felvétel szünetértékei ugyan nem különböznek egymástól szignifikánsan, tendenciaszerűen azonban azt mutatják, hogy az adatközlő egyre gyakrabban és egyre hosszabb néma szüneteket tart. A beszédben előforduló megakadásjelenségek gyakran megszakítják Z. felolvasá-

sát, ezáltal tovább csökkentik a folyamatosságot. Itt is sajátosan értelmeztük a spontán beszédben megszokott megakadásjelenség-fogalmat, és ejtéshibákról beszélünk. Az első felvételben 47 hibát adatoltunk, a másodikban 61-et, a harmadikban már 78 darabot. A felvételek időtartamához viszonyítva ez azt jelenti, hogy az első alkalommal az adatközlő percenként 15,31 hibát produkált, a második alkalommal 21,71 darabot, az utolsó felvételen percenként 23,32 hiba fordult elő (14. ábra). A statisztikai elemzések nem igazoltak szignifikáns különbséget.



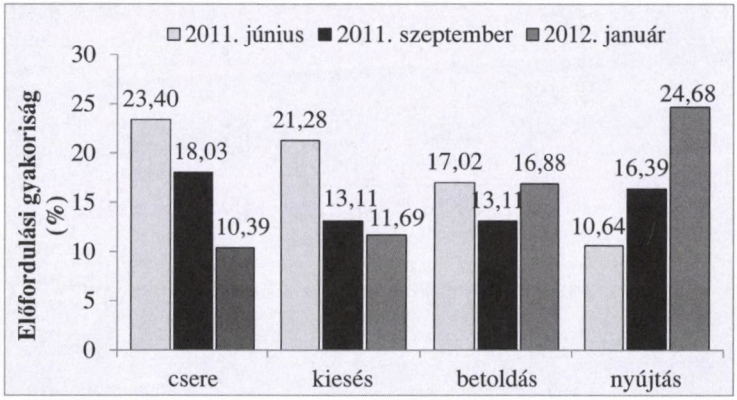
13. ábra
A néma szünetek hossza a felolvasásban felvételenként



14. ábra
A hibák percenkénti előfordulása a felolvasásban felvételenként

A hibázások egy része a beszélő bizonytalanságára vezethető vissza, míg másik részük valódi hiba, amely az artikulációs konfiguráció megváltozását eredményezi. A leggyakoribb jelenségek mindhárom felvételen a csere, a kiesés, a betoldás és a nyújtás voltak, előfordulási gyakoriságuk azonban felvételenként eltér (15. ábra). Az első két felvételen a leggyakoribb megakadás a csere volt (*a szakhatóságok elmúlt években **d**égzett **p**izsgálatai*, helyesen:

végzett vizsgálatai). Leggyakrabban egyetlen képzési jegy változik meg. Beszédét jellemzik az olyan magánhangzócserék, amikor a két hang írásképe tér el egymástól az ékezetben (pl. *a–á*, *e–é*, *u–ü*), például *immünrendszeri* az *immunrendszeri* helyett, *szezonalis* a *szezonális* helyett. A kiesés (pl. *növényvédő szerekbő* vagy *visszamara* – hiányzó szóvégi [t:] – *vegyi anyagokat*) az első felvételen szintén az egyik leggyakoribb hibának számít, ezt a későbbiekben felváltja a betoldás (pl. *az elmúlt égvekben*; *piacokbon kapható zöldségek*). A nyújtás az első alkalommal az összes hibajelenség 10,64%-át adta, a harmadik felvételen már ez a leggyakoribb jelenség. A beszélő felolvasás közbeni hibázásainak változásaiból az látszik, hogy az artikuláció egyre nehezítettebbé válik, az ejtés időzítése egyre jobban felbomlik.



15. ábra
Az „ejtéshibák" leggyakoribb négy típusának arányai felolvasásban

A kétféle beszédtípus adatainak összevetése

Adataink több tekintetben ugyanazon, illetve hasonló jellegzetességeket mutatják Z. beszédében, de akad eltérő eredmény is; a beszédtípus tehát bizonyos tekintetben hatással van a kivitelezésre. Adatközlőnk beszédtempója mind a mondatismétlésben, mind az olvasásban csökkenő értékeket mutat a harmadik alkalommal, lassul. Ennek két, egymással összefüggő magyarázata is lehet. Egyrészt az artikulációs nehezítettség fokozott összpontosítást kíván a beszélőtől, ehhez pedig több időre van szüksége. Másrészt többé-kevésbé tudatosan igyekszik lassabban beszélni, hogy a kiejtésben a célkonfigurációt elérje. A jitter és a jel/zaj viszony értékeinek növekedése a frekvenciaingadozás nagyobb mértékére utal, ez a mondatok ismétlésekor fokozatosan roszszabbodik, az olvasáskor nincsen lényeges eltérés a különböző időpontokban rögzített felvételekben. Ez utóbbi a beszédtípus eltérő tervezési folyamataival magyarázható. Az amplitúdóingadozás ellentétes tendenciát mutat a kétféle

beszédtípusban. A legrosszabb eredményt mondatismétléskor a harmadik alkalommal, felolvasásban pedig az első alkalommal tapasztaltuk.

A zörej részesedése (jel/zaj viszony) mindkét típusban a harmadik alkalommal a legnagyobb. A glottalizálás fokozatosan növekszik a mondatismétlésekben, míg a második felolvasáskor ebben némi csökkenés tapasztalható. Az alaphangmagasság átlagértékei növekedést mutatnak a felolvasásban, a mondatismétlésben nem mutatkozott lényeges különbség. Az F_0 növekedése olvasáskor feltehetően egyfajta tudatos kompenzációs stratégia a mélyebb alaphang és a szűkebb F_0-tartomány ellensúlyozására. A frekvenciatartomány mindkét beszédtípusban igen szűk, a mondatismétlésekben fokozatos a csökkenés, felolvasáskor a második alkalmat kivéve ugyancsak a szűkülés a jellemző. A megakadásjelenségek előfordulása fokozatosan növekszik a felolvasásban, míg a mondatismétléseknél kis csökkenés tapasztalható a második alkalommal és erőteljes növekedés a harmadik felvételkor. A tendencia tehát beszédtípustól függetlenül azonos. Az „ejtéshibák" előfordulása azonban a mondatismétlésekben percenként csaknem kétszer annyi, mint a felolvasásban.

Összegzésképpen az mondható, hogy Z. beszédsebessége lassú, a glottalizált szótagok aránya relatíve nagy, a hangszalagműködés paraméterei jelentős eltéréseket mutatnak a normál viszonyoktól, a rekedtes benyomás többféle tényező együttműködésének eredménye. Az alaphangmagasság kismértékben magasodik, a jól hallható diszprozódia oka a szűk frekvenciatartomány. A kiejtés pontossága nagymértékben csökkent, a szükséges artikulációs konfigurációk jó része nem felel meg a normatív ejtésnek, gyakoriságuk sokszorosan meghaladja a normál beszédre jellemző értékeket. Az egyes paraméterek tendenciaszerűen romlást mutatnak az első és a harmadik felvétel adatait összevetve. A nyolc elemzett tényező közül mondatismétléskor ötben volt kismértékű átmeneti javulás a második alkalommal, a felolvasásban pedig hatban (bár statisztikailag ezek a különbségek nem voltak igazolhatók).

Következtetések

Esetismertetésünkben egy nehezen diagnosztizálható, a felszínen sajátos beszédzavart mutató kórkép fonetikai leírását kíséreltük meg a beteg mondatismétlései és hangos olvasása alapján. Ugyanazon akusztikai-fonetikai tényezőket elemeztük a kétféle beszédanyagban egymást mintegy háromhavonta követő három felvétel során. Az eredményeink mind a zöngeműködés, mind az artikuláció tekintetében jellegzetes eltéréseket mutatnak a normál beszédképzéshez képest. Ezek a sajátosságok a dysarthriát támasztják alá. Adatközlőnk képtelen a beszédprodukciós mechanizmus alsóbb szintjeinek kontrollált működtetésére, függetlenül a beszédtípustól.

Feltételezhető, hogy a vizsgálataink idején egyértelműen (még) nem kimutatható organikus ok következményeként sérült a nyelvi modul, a fonológiai és a fonetikai tervezés sokszor véletlenszerű, zavart, bizonytalan működésű. A mondatismétlésekben tapasztalt gyenge teljesítmény a verbális memória

csökkent működésével, a korlátozott kapacitású munkamemóriával is összekapcsolható (Németh 2001). A vizsgált beszédfolyamatok általános lassulása jellemző Z. beszédprodukciójára.

A mondatismétlés és a felolvasás elemzési eredményei alapján adatközlőnknél az elsődleges progresszív afázia *logopenic*nek nevezett típusát (Mesulam 2001) valószínűsítjük. A mintegy két évig lappangó, illetve lassan progrediáló beszédzavar a kezdetekben semmilyen organikus okkal nem volt alátámasztható. A tünetek gyakoriságának növekedése, illetve a beszéd romlásának ezt követő felgyorsulása, a fonetikai elemzések tényszerű adatai egyértelműen igazolják ezt a feltételezett diagnózist. Adatközlőnk beszédmegértése, beleértve a szóértést is, hibátlan, beszédprodukciója agrammatikus jelenségeket nem tartalmaz, folyamatosan beszél. Az artikulációs jelenségek értelmezése teszi lehetővé a differenciáldiagnózist. Ehhez a spontán beszéd tervezett elemzése jelentős segítséget fog nyújtani. Mindezek alapján megtervezhető lesz a célzott terápia adatközlőnk számára.

Irodalom

Bánréti Zoltán (szerk.) 1999. *Nyelvi struktúrák és az agy. Neurolingvisztikai tanulmányok.* Corvina, Budapest.

Bonner, Michael F. – Ash, Sharon – Grossman, Murray 2010. The new classification of primary progressive aphasia into semantic, logopenic, or nonfluent/agrammatic variants. *Current Neurology and Neuroscience Reports* 10. 484–490.

Brockmann, Meike – Storck, Claudio – Carding, Paul N. – Drinnan, Michael J. 2008. Voice loudness and gender effects on jitter and shimmer in healthy adults. *Journal of Speech, Language, and Hearing Research* 51. 1152–1160.

Duffy, Joseph R. – Petersen, Ronald C. 1992. Primary progressive aphasia. *Aphasiology* 6. 1–15.

Gelfer Marylou P. – Fendel, Dawn M. 1995. Comparisons of jitter, shimmer, and signal-to-noise ratio from directly digitized versus taped voice samples. *Voice* 9. 378–382.

Gliebus, Gediminas G. 2010. Primary progressive aphasia: clinical, imaging, and neuropathological findings. *American Journal of Alzheimer's Disease and Other Demencias* 25. 125–127.

Gorno-Tempini, Maria Louisa – Hillis, Argye Elizabetj – Weintraub, Sandra – Kertesz, Andrew – Mendez, Mario – Cappa, Stefano Francesco – Ogar, Jennifer M. – Rohrer Jonathan D. – Black, Steven – Boeve, Bradley F. – Manes, Facundo – Dronkers, Nina F. – Vandenberghe, Rik – Rascovsky, Katya – Patterson, Karalyn – Miller, Bruce L. – Knopman. David S, – Hodges, John R. – Mesulam, Marcel – Grossman, Murray 2011. Classification of primary progressive aphasia and its variants. *Neurology* 15. 1006–1014.

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

Grossman, Murray – Ash, Sharon 2004. Primary progressive aphasia: a review. *Neurocase* 10. 3–18.

Gyarmathy Dorottya 2010. Megakadásjelenségek: A beszélő és a hallgató stratégiái. In Navracsics Judit (szerk.): *Nyelv, beszéd, írás. Pszicholingvisztikai tanulmányok I.* Tinta Kiadó, Budapest. 74–82.

Kertesz, Andrew – Munoz, David G. 1997. Primary progressive aphasia. *Clinical Neuroscience* 4. 95–102.

Markó Alexandra 2012. Az irreguláris zönge funkciói és gyakorisága olvasott és spontán beszédben. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások.* Akadémiai Kiadó, Budapest. 25–42.

Mesulam, M.-Marsel 1982. Slowly progressive aphasia without generalized dementia. *Annals of Neurology* 11. 592–598.

Mesulam, M.-Marsel 2001. Primary progressive aphasia. *Annals of Neurology* 49. 425–432.

Mesulam, M.-Marsel 2007. Primary progressive aphasia – A 25-year retrospective. *Alzheimer Disease and Associated Disorders* 21. 8–11.

Mesulam, M.-Marsel 2003. Primary progressive aphasia – a language-based dementia. *The New England Journal of Medicine* 349. 1535–1542.

Neuberger Tilda 2010. Mondatészlelési sajátosságok fiatal- és időskorban. In Gecső Tamás – Sárdi Csilla (szerk.): *Új módszerek az alkalmazott nyelvészeti kutatásban.* Kodolányi János Főiskola–Tinta Könyvkiadó, Székesfehérvár–Budapest. 220–225.

Németh Dezső 2001. A munkamemória szerepe a mondatmegértésben. In Pléh Csaba – Lukács Ágnes (szerk.): *A magyar morfológia pszicholingvisztikája.* Osiris Kiadó, Budapest. 83–118.

Noble, Kimberly G. – Glosser, Guila – Grossman, Murray 2000. Reading in dementia. *Brain and Language* 74. 48–69.

Obler, Loraine K. – Gjerlow, Kris. 1999. *Language and the brain.* Cambridge University Press, Cambridge.

Pick, Arnold 1892. Über die Beziehungen der senilen Hirnatrophie zur Aphasie. *Prager Medizinische Wochenschrift* 17. 165–167.

Rákli Veronika 2009. *A beszédprodukció és a beszédpercepció összefüggései: az elhangzó hírek feldolgozása.* Doktori disszertáció. ELTE, Budapest.

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

Szirmai Imre 2002. A kognitív zavarok neurológiai alapjai. In Tariska Péter (szerk.): *Kortünet vagy kórtünet? Mentális zavarok az időskorban.* Medicina Könyvkiadó, Budapest. 87–114.

Thompson, Cynthia K. – Ballard, Kathleen J. – Tait, Mary E. – Weintraub, Sandra – Mesulam, Marsel 1997. Patterns of language decline in non-fluent primary progressive aphasia. *Aphasiology* 11. 297–331.

Wertzner, Haydée F. – Schreiber, Solange –Amaro, Luciana 2005. Analysis of fundamental frequency, jitter, shimmer and vocal intensity in children with phonological disorders. *Brazilian Journal of Otorhynolaryngology* 71. 582–588.

A ZÖNGÉSSÉGI KONTRASZT ELSAJÁTÍTOTTSÁGÁNAK MINTÁZATA ATIPIKUS NYELVFEJLŐDÉS ESETÉN

S. Tar Éva

Bevezetés

A zöngés/zöngétlen kontraszt produkcióban való megjelenítése komoly nehézséget jelent atipikus nyelvi fejlődésű gyermekeknek, beszédükben a neutralizáció tovább fellelhető, mint tipikus fejlődés esetén (pl. Leonard 2000; Hua–Dodd 2006). A zöngés/zöngétlen explozívák képzéséhez a fonológiai zöngésségi oppozíció elsajátításán túl számos artikulációs, laringális és respirációs mozzanat összehangolt működésére van szükség, mely az egyes beszédszervek működése, illetőleg a különböző beszédszervek közti aktivitás koordinálásának megfelelő szintű fejlettségét feltételezi. A beszédszervi működések (pl. a velofaringális kapu zárása a lágyszájpad megemelésével, a szubglottális nyomás fenntartása folyamatos kilégzéssel, a hangszalagok zárása/nyitása, a szájüregi zár létrehozása/megszüntetése) többsége fellelhető a létrehozott hang akusztikai jellemzőiben. A szótag eleji explozívák esetében az egyik legtöbbet vizsgált paraméter a zöngekezdési idő (VOT), vagyis a szájüregi zár felpattanását jelző zörej és a zönge kezdete közti időintervallum (Lisker–Abramson 1964). A VOT értékéből a gégeműködés feletti kontroll megszerzésére (Koenig 2000), illetőleg a toldalékcső és a gégefő tevékenységének, vagyis az artikulációs (beszédhang-minőséget létrehozó) és fonációs (zöngeképzési) szakaszokban zajló izommozgásoknak idői összerendezettségére lehet következtetni (Abramson 1977).

Mivel a zöngésségi kontraszt fonetikai megvalósítása tekintetében különböző nyelvek eltérő megoldásokat választanak (pl. Lisker–Abramson 1964, 1967; Cho–Ladefoged 1999), elsajátításuk mintázata is ennek megfelelően különbözően alakulhat, annak gyorsaságát, illetőleg a hibázás természetét illetően egyaránt. A zöngekezdési idő alapján különbség mutatkozik az ún. előzöngét, illetőleg aspirációs kontrasztot használó nyelvek közt (Kager et al. 2007). Az aspiráló nyelvek közé tartozó angol nyelvben például a zöngétlen explozívák hosszú, pozitív VOT-vel jellemezhetők – ezek elsajátításában Macken és Barton (1980a) három fejlődési szakaszt különít el. Az első szakaszban a gyermek főként zéró vagy pozitív, rövid (< 20 ms) VOT-t produkál, majd a második szakaszban kezdi elkülöníteni a célfonéma zöngésségi értékének megfelelően a beszédhangokat, mégpedig úgy, hogy a zöngés szegmentum pozíciójában zéróhoz közeli VOT-t realizál, a zöngétlen célhan-

got pedig hosszabb, pozitív VOT-vel ejti. Ebben a fejlődési szakaszban azonban a zöngétlen célhang VOT értéke (< 40ms) nem éri el az angol zöngétlenére jellemző hosszúságot, a realizált beszédhang ezért az észlelésben még zöngésként jelenik meg. A harmadik szakaszban, kb. 18–29 hónapos korra (Lowenstein–Nittrouer 2008) azonban a gyermek már a felnőtt számára is észlelhető módon képzi a zöngétlen aspirált zárhangot. A pozitív, hosszú VOT-vel jellemezhető hangok kétéves kor körüli megjelenése további bizonyítást nyert a kantoni nyelvet elsajátító gyermekek vizsgálatával is (Clumeck et al. 1981).

Az aspirációs kontraszttal jellemezhető nyelvekkel szemben például a spanyol, a holland vagy a francia előzöngével realizálja a zöngés explozívákat. Macken és Barton (1980b) a mexikói spanyol, Allen (1985) a francia nyelvre vonatkozóan azt találta, hogy az előzöngével képzett orális zárhang az aspiráltnál dokumentálttól később, 4-5 éves korban jelenik meg a beszédben. A fonációs típusok fentiekben ismertetett elsajátítási sorrendjét támasztja alá Gandour és munkatársai (1986) tanulmánya is, amely a háromtagú oppozíciót használó thai nyelvre vonatkozóan szolgáltat adatokat. Az eredmények alapján a rövid, pozitív VOT, a hosszú, pozitív VOT és az előzönge közti hármas differenciálást a thai gyermekek kb. ötéves korra sajátítják el, amikor is a hosszú, pozitív VOT-vel jellemezhetőt követően utolsóként megjelenik beszédükben az előzöngével képzett zöngés explozíva is.

Az előzöngét elsajátító gyermekek vizsgálatakor továbbá az is kiderült, hogy a már elsajátított fonológiai oppozíciót a gyermekek különböző nyelvspecifikus stratégia (pl. spirantizáció, előnazalizálás, nazalizálás, illetőleg svábetoldás) segítségével jelenítik meg beszédükben addig is, míg a felnőtt nyelvre jellemző fonetikai viselkedést el nem sajátítják (pl. Macken–Barton 1980b; Allen 1985; Kuijpers 1993; Kong et al. 2007).

Az előzöngét, illetőleg aspirációs kontrasztot használó nyelvek fentiekben bemutatott elsajátítási sorrendje természetes módon az észlelt hibázás mintázatában is visszatükröződik. Míg az előzöngés nyelvet elsajátító gyermek szó eleji helyzetben jellemzően zöngétleníti a zöngés explozívákat, az aspirációs kontrasztot elsajátító ugyanebben a fonetikai helyzetben zöngéssel helyettesíti a zöngétlen célszegmentumot (Kager et al. 2007).

A különböző zöngésségi realizációk elsajátítási rendjét Kewley-Port és Preston (1974) aerodinamikai sajátosságokkal magyarázza. Aspirált zöngétlen explozívák produkciójakor a hangszalagok nyitott állását a szájüregi zár felpattanása után bizonyos ideig fenn kell tartani ahhoz, hogy az azonosításhoz szükséges hallható aspirációs zörej létrejöjjön. Előzönge esetén pedig az aerodinamikai sajátosságból adódó artikulációs nehézség abból adódik, hogy az explozíva zárszakasza alatt növekvő szájüregi levegő nyomása akadályozza a zönge fennmaradását, szókezdő helyzetben (a hangszalagok tág lélegző állása miatt) annak megindítását is. Az előzönge hosszabb idejű fenntartása a szájüreg méretének növelése révén lehetséges, mely – a passzívan bekövet-

kező változásokon túl – kompenzációs artikulációs mozzanatok beépítését igényli (pl. Ohala 1983), komplexebbé téve ezáltal a képzés módját. Összességében tehát elmondható, hogy az aspirált zöngétlen, illetőleg az előzöngével képzett explozívák elsajátítása, szemben a rövid, pozitív VOT-vel jellemezhető hangokéval, a laringális és szupralaringális működés pontosabb idői koordinálását feltételezi, és ezzel párhuzamosan komplexebb neuromuszkuláris kontrollt igényel.

Az észlelési adatok megerősítése céljából (vö. pl. Kent 1996, Farmer 1997), és mert az elsajátított fonológiai kontraszt fonetikai megvalósítására való törekvés gyakran csak szubfonémikus szinten tárható fel, számos VOT-elemzés született atipikus nyelvi adatok feldolgozásával is. Angol nyelvű specifikus nyelvi károsodást mutató gyermekek beszédmintájának elemzése szignifikánsan eltérő VOT-értékeket eredményezett a zöngétlennek észlelt zöngés célszegmentumok és azok zöngétlen párjai tekintetében, utalva arra, hogy a gyermekek fonológiai szinten a zöngésségi oppozíciót már elsajátították (vö. Leonard 2000). Perceptuálisan fel nem ismert, ugyanakkor akusztikailag elkülönülő hangzóprodukcióról számolnak be Aozou és munkatársai (2000) fluens afáziásokra vonatkozó kutatásokat összegezve, a zöngekezdési idő értékeiben fellelhető variabilitást ugyanakkor szerzett apraxia, illetőleg dysarthria jeleként tartják számon. A fonológiai fejlődés késését/zavarát mutató gyermekek szókezdő explozíváinak akusztikai elemzése szintén arra mutatott rá, hogy az auditoros alapú fonológiai vizsgálat során feltárt hasonló hibázások mögött nagyon is különböző produktív tudás állhat (pl. Maxwell–Weismer 1982; Gierut–Dinnsen 1986). Az észlelésben még nem azonosítható, azonban akusztikailag, a zöngekezdési idő értékeiben már szignifikánsan elkülönülő zöngés-zöngétlen explozívarealizáció [az irodalomban *fedett* (*covert*) *kontraszt*nak nevezett jelenség] klinikai szempontból is fontos fejlődési mutató, egyrészt az oki háttér természetének azonosítása (Scobbie et al. 2000), másrészt a prognózis megfogalmazása szempontjából. Tyler, Edwards és Saxman (1990) például fonológiai zavart mutató gyermekek terápiára való reagálását vizsgálva azt tapasztalta, hogy szubfonémikus szintű szembeállítást produkáló gyermekeknél a terápiában célzott kontraszt gyorsabban generalizálódott, mint azoknál, akik nem rendelkeztek ezzel a tudással.

A magyar nyelvben a zöngésségi oppozíció az explozívákat tekintve fonetikailag az előzöngével jellemezhető zöngés és a zöngétlen nem aspirált szembeállítással realizálódik (Gósy 2004). Gósy és Ringen (2009) felnőtt adatközlőktől szólista felolvasása révén nyert adatokat, és a szó eleji fonetikai helyzetben álló /p, t/ realizációkra átlagosan 9,7 és 16 ms VOT-értéket adatolt, míg a /b, d/ realizációk esetében –94,6 és –95,1 ms-ot; a zöngés célfonémák mindegyike előzöngével realizálódott. A zöngétlen bilabiális és alveoláris VOT-értéke szignifikánsan különbözött egymástól. Az átlagos eltérés, valamint a minimum- és a maximumértékek közti különbség a zöngés hangzók esetében volt magasabb, a képzési hely hatása a szóródási mutatókra

nem volt jelentős. Gósy (2000, 2001) intervokális helyzetben álló zöngétlen zárhangok VOT-értékeinek vizsgálata során ugyanakkor azt tapasztalta, hogy a hangok akusztikai sajátosságai eltérően alakulnak gondozott és spontán beszédben, a VOT értékét beszédstílustól függően befolyásolja a követő magánhangzó minősége, valamint az explozíva képzési helye. A /p/ és /t/ realizációk VOT-értéke nem különbözött gondozott beszédben, viszont jelentős volt a követő magánhangzó hatása, a minőségét meghatározó képzési tulajdonságok mindegyike vonatkozásában.

Bóna (2011) a zöngétlen mássalhangzók zöngekezdési idejét vizsgálva további olyan faktorra mutatott rá, mely hatással van a hangzók akusztikai sajátosságaira. Idősek és fiatalok kétféle beszédstílusból származó adatait elemezve azt találta, hogy a természetes idősödés következtében az elöl képzett hangzók VOT-értéke nő, a különféle beszédhelyzetek nem befolyásolják az eredményt. A szerző a jelenséget az artikulációs tempó lassúbbodásával, illetőleg a beszédszervek működését érintő strukturális és fiziológiai változásokkal magyarázza.

Gráczi (2012) az explozívarealizációk sajátosságait logatomokat tartalmazó mondatok felolvasásában és spontán beszédben elemezte olyan fonetikai jegyek mentén is, mint például a teljes mássalhangzó időtartamához viszonyított zöngétlen rész aránya és ezek összefüggése a képzési hellyel. A mondatfelolvasás adatai alapján abszolút szó eleji helyzetben 1. feltételezhetően a zönge létrehozását könnyítendő, előfordul magánhangzós ejtésindítás; 2. a zöngésségi párok tagjai teljes mértékben elkülöníthetők a képzés alatt megjelenő zönge/zöngétlenség arányával; 3. a zöngés mássalhangzók teljes időtartamához viszonyított zöngés részek aránya nem különbözik jelentősen a szó belseji intervokális helyzettől; 4. az alveoláris esetében a zöngearány a bilabiálisénál valamivel alacsonyabb átlagos értéket mutat.

További fonetikai jegyként Gráczi és Kohári (2012) a zárfelpattanás kvantitatív sajátosságait vizsgálta zöngétlen realizációkban, rámutatva arra, hogy – laboratóriumi beszéd esetén – a felpattanás elmaradása ritkán, többszöri felpattanása gyakran (az esetek közel felében) előfordult, utóbbi ezen kívül gyakrabban alveoláris, mint bilabiális képzés esetén. A felpattanások mennyiségét a környező magánhangzó minősége statisztikailag igazolható módon csak a bilabiális esetében befolyásolta; a felpattanások mennyisége és képzési hely szerinti eloszlása a beszélők ejtési sajátosságainak is függvénye. A kettős, illetőleg többszörös felpattanás intenzitásbeli jellemzőit tekintve a szerzők azt találták, hogy a legintenzívebb felpattanás a legtöbbször az elsővel esett egybe.

Magyar gyermeknyelvi kutatásokból tudjuk, hogy bár az elöl képzett (bilabiális, alveoláris) explozívák zöngésség szerinti elkülönítése a fonológiai fejlődés korai fejleménye, jelentős gyermekek közti eltérés tapasztalható a szóhatár pozícióban, különösen szóvégi helyzetben álló zöngés hangok korrekt produkcióját illetően (Meggyes 1971; Gósy 1984, 1998; Lőrik 1982, S. Tar

2006). A zöngés explozívák szóhatár pozícióban való korrekt ejtése nagy kihívást jelent atipikus nyelvi fejlődésű gyermek számára is. Elsődleges nyelvi zavart mutató, átlagosan 6;7 életkorú gyermekek beszédprodukciójának elemzése feltárta, hogy 60%-uk zöngétlenít szóhatár pozícióra előírt obstruenst, az érintett elemek az esetek többségében zöngés explozívák, a hibázás gyakorisága – a tipikus fejlődésben tapasztaltaktól eltérően – mindkét szóhatár pozícióban hasonlóan alakul (S. Tar 2008).

Atipikus nyelvfejlődés esetén végbemenő zöngésségi kontraszt elsajátításáról a magyar nyelvre vonatkozóan ez idáig kevés adat áll rendelkezésre. A jelen tanulmány célja, hogy elsődleges nyelvi zavarral diagnosztizált gyermekek zöngésségre vonatkozó produkciós tudását észlelésalapú és akusztikai elemzés segítségével feltérképezze. A vizsgálat az explozívák közül a bilabiális és alveoláris párokra /p, b, t, d/ terjed ki, tekintve, hogy a gyermeknyelvi adatok szerint ezek azok a zárhangok, melyek hibás realizációja (az esetek döntő többségében) a zöngésségi kontraszton kívül egyéb jegyet nem érint (szemben például a velárisával, ahol a képzéshelyre vonatkozó kontraszt is érintett lehet). A zöngésségi kontraszt elsajátítottságát a zöngés célfonémák egyénenkénti korrekt ejtéseinek száma, illetőleg a teljes mintára vonatkozó zöngekezdési idő értékei alapján vizsgáltuk.

Anyag, módszer, vizsgálati személyek

Az adatgyűjtés egy szélesebb körű vizsgálat részeként, képmegnevezés segítségével történt. Egy, a szegmentális fonológiai profil feltárására összeállított, egymorfémás főneveket ábrázoló képanyag 19 iteme volt a jelen vizsgálat anyaga. A képek által megjelenített szóalak a célfonémát szó eleji CV szekvenciában, fonémánként legalább négy különböző fonetikai környezetben tartalmazta, melyek közül egyben a követő magánhangzó felső nyelvállású volt. A szótagszám 1–3 között váltakozott.

A gyermek produkcióját (összesen 285 szót) lejegyeztük, ezenkívül Sony ICD–MS525 típusú diktafonnal rögzítettük. A vizsgálatra egyénileg, a gyermek óvodájában, egy csendes helyiségben került sor.

A vizsgálati anyagon kétféle elemzést végeztünk. Az észlelésalapú vizsgálatban a fonematikus átírással lejegyzett anyagot gyermekenként elemeztük, az értékelés alapja a fonémarealizáció helyessége (a célfonémának megfelelő realizáció), illetőleg a hibázás jellege volt. A perceptuálisan nem egyértelmű realizációt inkorrektnek értékeltük. Az inkorrekt produkciót a szerint kategorizáltuk, hogy zöngésségi értéket érintett-e a hibázás, vagy egyéb más tulajdonságot.

Az akusztikai vizsgálat során a digitálisan rögzített hanganyagot a Praat 5.3. szoftverrel (Boersma–Weenink 2011) elemeztük. Az adatolást a hangszínkép és az oszcillogram alapján manuálisan végeztük. A zöngekezdési idő meghatározásához egyrészt a zárszakasz feloldásának kezdetét, másrészt az abszolút zöngeindulást jelöltük. Többszörös felpattanás esetén az elsőt figye-

lembe véve mértük az időtartamot. Negatív VOT meghatározásánál az időintervallum a zönge megindulásától a(z) (első) felpattanás kezdetéig tartott, beleértve az esetlegesen előforduló néma szakaszt (a részben zöngétlenedett adatokat) is. Összesen 265 VOT-értéket mértünk. A nem elemzett szóalakok azok voltak, melyekben a célfonéma nem zárhangként realizálódott, vagy frázisban hangzott el, zajos volt a felvétel, illetőleg nazális betoldással vagy előnazalizációval képződött a zönge. Az akusztikai elemzés a korrektség megítélésére, illetőleg a zöngekezdési idő meghatározására tért ki. A korrektség megítélésének alapja ebben az esetben a realizáció VOT-értékének pozitív, illetőleg negatív tulajdonsága volt.

Az egyénenkénti korrekt ejtések számán alapuló elsajátítottságmintázatot három kategória mentén írtuk le: 1. nincs várt ejtés, 2. legalább két várt ejtés, 3. mindegyik realizáció korrekt. Elsajátítottnak vettünk egy fonémát, ha a gyermek legalább kétszer korrekten (az elemzési szempont alapján vártnak megfelelően) realizálta azt. A korrekt ejtés gyakoriságának megítéléséhez a perceptuális és akusztikai elemzés eredményeit összevetettük.

A statisztikai elemzéshez az SPSS 17.0. szoftvert használtuk.

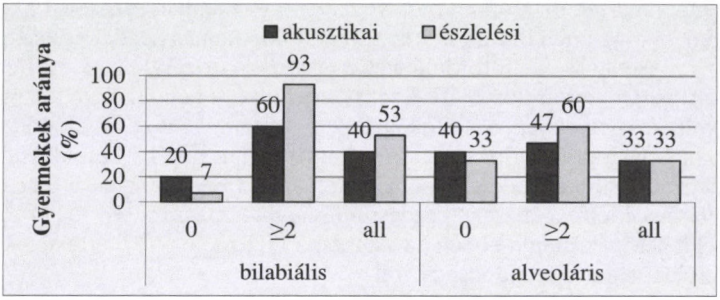
A vizsgálatban 15 5;6–7;7 éves korú, a nyelvfejlődés elsődleges zavarával diagnosztizált gyermek, 11 fiú és 4 leány vett részt. Átlagéletkoruk 6;7 év volt. Intelligenciakvóciensük az átlagos övezetbe esett. Az atipikus beszédfejlődés hátterében organikus, funkcionális vagy pszichoszociális okot, illetőleg a környezeti beszédingerek deficitjét nem lehetett kimutatni.

Eredmények

A korrekt ejtések száma alapján vett zöngésségi kontraszt-elsajátítottság mintázata

A vizsgált fonémarealizációk észlelési adatainak és akusztikai elemzésének összevetése alapján a bilabiálisra vonatkozóan 88%-os, az alveolárisra 93%-os az egyezés. Mindkét esetben a zöngésnek észlelt adatok bizonyulnak tévesnek.

A két vizsgálati eljárásból nyert eredmények közti nagyfokú egyezés ellenére az akusztikai elemzés eredményeként változnak az elsajátítottság mintázata leírásához választott kategóriaértékek (1. ábra). A legalább két korrekt realizáción alapuló korrekt ejtés aránya csökken, az elsajátítottság bilabiálisra vonatkozóan 60%, alveolárisra 47%. Célszótól független korrekt ejtés a képzési hely tekintetében megközelítőleg hasonló arányban jellemzi a gyermekeket, a bilabiálist 40%-uk, az alveolárist 33%-uk ejti megfelelően. Azon gyermekek aránya, akik mindegyik célszóban zöngétlenül valósítják meg a zöngés célfonémát, az észlelésalapú elemzés eredményeihez képest nő, alveolárisra vonatkozóan a gyermekek közel fele (40%) sikertelen a korrekt produkcióban, szemben a bilabiálisra vonatkozó 20%-os értékkel.



1. ábra
A vizsgált gyermekek százalékos aránya az észlelési és akusztikai elemzés alapján korrektnek értékelt zöngés mássalhangzók száma alapján

A gyermeki hibázások az esetek többségében a zöngésséget érintik, a bilabiális esetében a nem célfonémának megfelelő realizáció 90%-a, az alveoláris 96%-a a zöngésségi pár zöngétlen tagja.

A zöngekezdési idő értékei

A VOT-elemzés időtartamra vonatkozó eredménye az 1. táblázatban látható. A táblázatban a vizsgált gyermekekre vonatkozó átlagos VOT-érték, valamint két szóródási mutató (átlagos eltérés, terjedelem) mellett – a célértékkel való könnyebb összevethetőség érdekében – Gósy és Ringen (2009) felnőtt beszédmintáiban mért, szó eleji pozícióra kapott VOT-értékei is szerepelnek.

1. táblázat: A célfonémák korrektség szerinti realizációinak VOT-értékei és szóródási mutatói (Az *n* értéke az adott fonetikai kategóriában elemzett előfordulások számát jelöli)

	Időtartam (ms)					
	Átlag		Átlagos eltérés		Terjedelem	
	Gyermek	Felnőtt	Gyermek	Felnőtt	Gyermek	Felnőtt
[p]	21,7 (n = 69)	9,7	16,5	5,4	106	30,8
[t]	24,5 (n = 74)	16	16,7	6,7	99	38,2
Zöngétlenen ejtett bilabiális (/b/)	32,9 (n = 29)	–	44,2	–	190	–
Zöngétlenen ejtett alveoláris (/d/)	24,8 (n = 32)	–	11,7	–	44	–
[b]	–91,2 (n = 34)	–94,6	34,3	28,1	140	178,4
[d]	–84,7 (n = 27)	–95,1	54,8	29,8	170	152,7

A **célfonémának megfelelő realizációk** átlagos VOT-értékei a felnőtt nyelvben mért eloszlási adatokat tükrözik, időtartamuk azonban a zöngétlenekre nézve jelentősen hosszabb (különbség bilabiálisra: 12 ms; alveolárisra: 8,5 ms), a zöngések esetében (arányait tekintve) mérsékelten rövidebb (különbség bilabiálisra: 3,4 ms; alveolárisra: 10,4 ms). A szóródási mutatók a zöngés bilabiálisra vonatkozó terjedelemértéktől eltekintve magasabbak a gyermekek beszédmintájában, mint a felnőttekében, különösen figyelemre méltó a zöngétlenek terjedelemmutatója tekintetében adatolt eltérés (bilabiálisra: 75,2 ms; alveolárisra: 60,8 ms).

A **zöngés célfonémák zöngétlen realizációi** átlagos VOT-értéke a rövid, pozitív VOT kategóriába esik. Kategórián belül a zöngétlen korrekt és zöngés inkorrekt realizációk közt a bilabiálisra vonatkozó átlagos értékek közt jelentős eltérés tapasztalható (bilabiálisra: 11,2 ms; alveolárisra: 0,3 ms). A zöngétlenen ejtett bilabiális realizációk szóródási értékei a bilabiális zöngétlenéhez viszonyítva lényegesen magasabbak (zöngétlenen ejtett *b* és *p* kategóriák közti különbség: átlagos eltérés: 27,7 ms; terjedelem: 84 ms). Az inkorrekt bilabiális szóródási értékével szemben a zöngétlenen ejtett alveoláris realizációk a vizsgált fonéma megvalósulások legstabilabbjának bizonyulnak (átl. elt.: 11,7 ms; terjedelem: 44 ms).

A zöngétlen realizációk terjedelemértékei azon kívül, hogy jelentős mértékű variabilitást jeleznek, arra is rámutatnak, hogy a realizációk közt akad olyan, melynek VOT-értéke nem illeszkedik a magyar nyelvre jellemző fonetikai zöngésségi kategóriák egyikéhez sem (ilyen pl. az „erősen aspirált" kategória, vö. Cho és Ladefoged 1999).

A gyermeki mintában mért **zöngétlen realizációk** átlagos VOT-értékeire megállapítható, hogy – tekintet nélkül a célfonéma zöngésségi értékére – a képzési hely nem befolyásolja szignifikánsan az átlag zöngekezdési időt; valamint nem mutatkozik statisztikailag szignifikáns eltérés a zöngétlen és a zöngétlenen ejtett zöngés célfonémák realizációi közt sem (egytényezős ANOVA $F(3, 200) = 1{,}76$, $p = 0{,}16$).

A **zöngétlen realizációk célszavankénti VOT-értékeit** a 2. és 3. táblázat foglalja össze. A korrekt zöngétlen hangok átlagos VOT-értékei közt – a legrövidebb és leghosszabb idők közti jelentős különbség ellenére (bilabiálisra: 14,2 ms; alveolárisra: 14,0 ms) – nincs statisztikailag igazolható eltérés (egytényezős ANOVA bilabiálisra: $F(4, 64) = 2{,}07$, $p = 0{,}09$; alveolárisra: $F(4, 69) = 1{,}51$, $p = 0{,}21$). A zöngekezdési időt nem befolyásolja egyértelműen a követő magánhangzó nyelvállásfoka. A VOT-értékek sorrendje a felső állású felé haladva: 16,3 és 17 ms; 19,6 és 30,5 ms; 25 ms a bilabiálisra, valamint 23,9 ms; 18,9–25 ms; 32,9 ms az alveolárisra vonatkozóan. Az irodalomban dokumentált sorrendtől (magasabb nyelvállású követő magánhangzó hosszabb zöngekezdési időt eredményez (pl. Lisker–Abramson 1967; magyar intervokálisra Gósy 2000, 2001) való eltérést a /poha:r/, illetőleg a /ta:ʃkɔ/ szavak vártnál magasabb VOT-értékei okozzák.

2. táblázat: A zöngétlen fonémák realizációinak célszavankénti VOT-értékei

Célszó	Időtartam (ms)	
	Átlag	Átlagos eltérés
/poha:r/	30,5	25,3
/piʒɔmɔ/	25,4	22,6
/pøc:/	19,6	7,6
/pɔputʃ/	17,0	4,9
/pɛrɛts/	16,3	8,0
/ty:/	32,9	22,3
/tɛhe:n/	25,0	22,0
/ta:ʃkɔ/	23,9	15,6
/tɛmplom/	21,6	8,6
/tɛlɛfon/	18,9	7,3

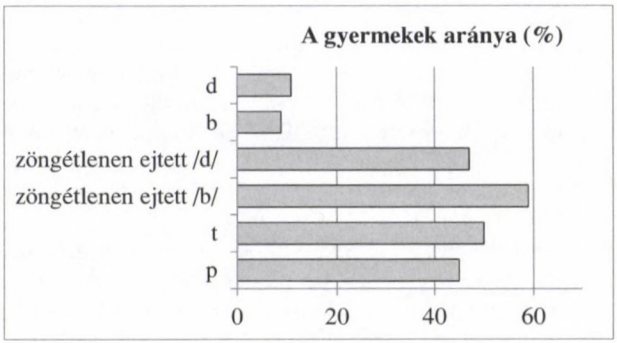
3. táblázat: A zöngés fonémák zöngétlen realizációinak célszavankénti VOT-értékei

Célszó	Időtartam (ms)	
	Átlag	Átlagos eltérés
/bus/	68,5	89,1
/bodʑɔ/	46,3	65,0
/boho:ts/	24,7	8,6
/bɔgoj/	21,4	9,7
/be:jɛg/	17,7	3,8
/doboz/	29,3	11,6
/diɲ:ɛ/	26,7	13,1
/domino:/	21,2	9,8
/dɔra:ʒ/	20,6	12,2

Az inkorrekt zöngétlen megvalósulások VOT-értékeire a minta kis elemszáma miatt (célszavanként átlagosan 7 realizáció) nem végeztünk statisztikai számításokat, az adatok azonban ennek hiányában is informatívnak bizonyulnak. A leghosszabb és legrövidebb átlagos VOT-érték közti különbség bilabiális képzési helyen számottevő (bilabiális: 50,8 ms; alveoláris: 8,7 ms). Ebben a kategóriában két célszóban is (/bus/, /bodʑ:ɔ/) kiugróan magas zöngekezdési időt adatoltunk, azonban az átlagos eltérés is ezekre a szavakra vonatkozóan a legmagasabb. A zöngétlenen ejtett bilabiális realizációk VOT-értékeinek fentiekben ismertetett jelentős szóródása tehát magyarázható a néhány célszóban adatolt szélső értékekkel.

Az átlag zöngekezdési idők közti relatív eltérés ebben az esetben sem követi egyértelműen a követő magánhangzó nyelvállásfoka alapján várt sorrendet (felső nyelvállású felé haladva az értékek bilabiálisra: 21,4 ms; 17,7–46,3 ms; 68,5 ms; alveolárisra 20,6 ms; 21,2–29,3 ms és 26,7 ms).

Az elemzés során tapasztalt, **egyéb fonetikai jegyek** tekintetében vett sajátosság az egynél több felpattanást mutató fonémamegvalósulások magas, összességében az elemzett szavak 40%-ában adatolt aránya. A jelenség elsődlegesen a zöngétlen realizációkat jellemzi, függetlenül a célfoném zöngésségi értékétől, illetőleg képzési helyétől (2. ábra). A felpattanások intenzitását illetően a mintázat jellemzően úgy alakul, hogy egy gyenge felpattanást követ egy vagy több erősebb intenzitású felpattanás.



2. ábra
Egynél több felpattanást mutató célfoném-realizációk kategóriánkénti százalékos aránya

A zöngés realizációk zárszakasza az esetek döntő többségében teljes zöngével realizálódik. Néma szakaszt a zöngés bilabiális 2%-a, az alveoláris 11%-a foglal magában. A zöngétlen szakasz a zárképzés végén, a felpattanás előtt valósul meg, időtartama egy kivételtől (60 ms) eltekintve 10 ms körüli.

Következtetések

A jelen tanulmány célja, hogy a zöngésségi kontraszt elsajátítottságmintázatát 5;6 évesnél idősebb (6;7 éves átlagéletkorú) elsődleges nyelvi zavart mutató gyermekek képmegnevezéssel kiváltott beszédprodukcióját elemezve feltárja. A kontraszt meglétét szó eleji CV helyzetben levő bilabiális és alveoláris képzési helyű explozívákban észlelés alapján történő és akusztikai elemzés során nyert adatokkal vizsgálta.

A rögzített adatok 8%-án nem történt akusztikai elemzés, a kizárás oka többnyire a zöngés célfonémák nazalizált vagy nazális betoldásával történő megvalósítása volt. A jelenség ugyanakkor tovább vizsgálandó, mivel az orr-

üreg felé vezető út megnyitását az előzönge létrehozásakor fennálló aerodinamikai nehézsége leküzdése egyik stratégiájaként tartja számon az irodalom (pl. Ohala 2011), és előzöngés nyelvet elsajátító gyermekek is gyakran élnek ezzel a felnőttétől eltérő zöngeképzési stratégiával (Allen 1985, Kong et al. 2007).

Az akusztikai elemzés eredménye, valamint az észlelési adatok közt nagyfokú egyezés tapasztalható, a két elemzési mód eredménye közti eltérés 10% körüli, és a korrektnek észlelt zöngés realizációkat érint. A zöngésnek észlelt, ugyanakkor a regisztrátumon pozitív VOT-vel jellemezhető realizációk VOT-értéke nem tér el a zöngétlennek észlelt realizációk zöngekezdési idejétől, a téves azonosítás hátterében a zöngekezdési időn kívüli tényezők állhatnak.

A korrektnek észlelt itemek száma alapján értelmezett elsajátítottság mintázata az akusztikai elemzés eredményeként módosul, csökken azon gyermekek száma, akikre a kritériumfeltétel teljesül, és nő azoké, akik egy célszóban sem realizálnak zöngés fonémát. A vizsgálat ezen eredménye megerősíti az az irányú törekvéseket, melyek a klinikum területén is hasznosnak és szükségesnek tartják az objektív mérési eljárások alkalmazását (Kent 1996; Ball–Code 1997; Auzou et al. 2000; Howard–Heselwood 2011), illetőleg a beszéd finomabb eltéréseit is megragadó fonetikai átírást (Hewlett–Waters 2004; Munson et al. 2010; McLeod–Goldstein 2012).

Az egyénenkénti elemzés eredményeit összegezve elmondható, hogy a vizsgált gyermekek nagy részének problémát jelent a szó eleji zöngésségi kontraszt produkcióban való megjelenítése, bár a probléma természete különböző lehet. A feltételezhető okok alapján három, az elsajátítás különböző szakaszában álló fejlődési csoportot lehet elkülöníteni:

1. Az egyénenként legalább két korrekt ejtést alapul véve a gyermekek 60%-a a bilabiális, 47%-a az alveoláris helyre elsajátította a fonológiai zöngésségi kontrasztot – a fejlettség feltárására irányuló kutatásokban szokásosan alkalmazott 75%-os csoportszintű kritériumszintet figyelembe véve (vö. pl. Smit et al. 1990) tehát egyik zöngés szegmentum sem tekinthető elsajátítottnak –; kétharmaduk fonetikai környezettől függetlenül realizálja a zöngés explozívákat.

2. A nyelvi kontrasztot elsajátító gyermekek fennmaradó egyharmada alkotja a második csoportot, ők azok, akik csak bizonyos szavakban ejtik a zöngés explozívát, számukra a meglévő fonológiai tudás környezetfüggetlen kivitelezése jelent kihívást.

3. A harmadik csoportba tartozó gyermekek nem rendelkeznek a kivitelezés feltételéül szolgáló fonológiai tudással sem, a vizsgált gyermekek egyötöde képzési helytől függetlenül, további egyötödük alveoláris képzési helyen nem ejt zöngés explozívát.

A VOT-értékek elemzése a gyermekek fonetikai viselkedésének további sajátosságát tárta fel. A célfonémának megfelelően realizált beszédhangok zöngekezdési értékeinek jellemzői (a felnőtteknél tapasztalthoz képest hosz-

szabb pozitív, illetőleg rövidebb negatív VOT-érték, valamint a nagyobb variabilitás) arra engednek következtetni, hogy bár a gyermekek egy része elsajátította az artikulációs és laringális gesztusok nyelvspecifikus koordinálását lehetővé tevő tudást, mégpedig oly mértékben, hogy az az észlelésben is kifejeződik, a zöngésségi kontraszt kivitelezéséhez szükséges koordinációs ügyesség még nem érte el a felnőttre jellemző mértéket. A motoros kontroll fejlődő voltára (pl. a bilabiális hangok felnőttitől eltérő képzése vö. Green et al. 2000) utalhat az a vizsgálati eredmény is, mely szerint a felnőttek zöngétlen explozívái VOT-értékének a képzési hely függvényében történő változása (Gósy–Ringen 2009) a vizsgált gyermekcsoport átlagos VOT-értékeiben nem lelhető fel.

A zöngés célfonémák zöngétlenül realizált beszédhangjainak zöngekezdési ideje és a zöngésségi pár zöngétlen tagjára vonatkozó értékek eloszlása nem utal szubfonémikus szintű kontraszt megvalósítására. A VOT értékeinek eloszlási mintázata tehát megerősíti azt a fenti megállapítást, mely szerint a vizsgálatban szereplő gyermekek 20%-a a bilabiálisra sem, 40%-a az alveolárisra nézve nem sajátította el a fonológiai zöngésségi kontrasztot szó eleji fonetikai helyzetben. Esetükben a tipikus fejlődéshez képest jelentős, több éves az elmaradás.

A vizsgált gyermekek fonetikai viselkedésük alapján a Scobbie és munkatársai (2000) által javasolt, négy szakaszból álló elsajátítási modell első, illetőleg harmadik szakaszába sorolhatók. Az első csoportba tartozók beszédmintájában nincs észlelhető különbség a zöngés és zöngétlen célfonémák realizációi közt, a harmadik csoport tagjai pedig ún. éretlen kontrasztot produkálnak, tehát a nyelvre jellemző akusztikai kulcsot a felnőttitől eltérő értékkel használják. Miután azonban a zöngekezdési idővel történő elkülönítés nem az egyetlen módja a zöngésségi kontraszt megvalósításának, nem zárható ki, hogy az első csoportba tartozó gyermekek vagy legalábbis egy részük, a felnőttétől eltérő stratégiát választott a zöngés fonéma realizálásához (lásd pl. a jelen vizsgálatban nem elemzett, a zárszakaszban nazális zöngét tartalmazó hangzók szerepét).

A VOT-értékekben tapasztalt, felnőttekénél jelentősebb mértékű variabilitás a gyermeki produkció sajátossága. A gyermek fonetikai viselkedésének e tulajdonsága utalhat arra, hogy az explozívák zöngekezdési idő alapján történő szembeállításának kivitelezését lehetővé tevő beszédszervi működések idői összerendezésének neurális irányítása még folyamatban van. Az oralmotoros reflex fejlődésének átmeneti időszaka (Smith 2010), illetőleg a hangszalagok, valamint a légzőrendszer életkorhoz köthető strukturális változása, az ezek következtében előálló aerodinamikai tényezők, továbbá a gége beszéd céljából való működtetése önmagában is (Koenig 2000) eredményezhet gyermekenkénti eltéréseket a zöngekezdési idő értékeiben. A nagymértékű variabilitás ugyanakkor motoros tervezési zavar jele is lehet (Auzou et al. 2000). Annak eldöntése, mennyiben tekinthető életkori sajátosságnak, illető-

leg mennyiben a zavart fejlődés jelének a vizsgálat adataiban tapasztalt variabilitás, további kutatásokat kíván. A komplexebb neuromuszkuláris kontrollt igénylő, előzöngével képzett hangok VOT-értékei közt fellelhető variabilitás magyarázható a fentiekben említett életkori tényezővel. A zöngétlen hangok esetében adatolt magas szóródási értékek értelmezéséhez azonban tipikus fejlődésről számot adó adatokra lenne szükség annak megítéléséhez, hogy vajon ebben az esetben is a motoros kontroll fejlődésének adott életszakaszra jellemző variabilitásával van-e dolgunk.

A zöngekezdési idő értékét a fonetikai környezet – bár nem statisztikailag igazolt mértékben – befolyásolja, azonban a felnőtt beszédben tapasztalttól eltérően néhány esetben nem a követő magánhangzó nyelvállásfoka a leginkább ható tényező. A tipikus fejlődésről számot adó és célzottan a fonetikai környezet (követő magánhangzó egyéb minőségi jellemzői, illetőleg a teljes szó fonetikai tulajdonsága, szótagszám) hatását feltáró vizsgálatokra van szükség annak eldöntéséhez, hogy életkori jelenségnek vagy az atipikus fejlődés jelének tekinthető-e a fonetikai környezet jelen vizsgálatban tapasztalt mértékű és természetű befolyásoló hatása.

Az elemzett beszédminta felnőttitől való eltérését mutatja a pozitív VOT realizálódásokban, különösen a bilabiálisokéban adatolt többszörös felpattanások magas száma is, valamint az intenzitás mintázata (vö. Gráczi–Kohári 2012). E jelenség esetén szintén felmerül a kérdés, hogy a jelenség a respirációs, laringális és artikulációs működések kontrollálásának életkori sajátosságából adódik, vagy esetleg időzítési problémára utal, illetőleg hogy egyáltalán módszertanilag indokolt-e a felpattanások kvantitatív jellemzőit a háttérben zajló motoros folyamatok jeleként vizsgálni, értelmezni.

Összefoglalás

A jelen vizsgálat eredményeit összegezve elmondható, hogy a mintában szereplő atipikus nyelvi fejlődésű gyermekek zöngésségi kontrasztot érintő produkciós tudása még nem teljes, jóllehet az elsajátítottság foka/természete tekintetében jelentős gyermekek közti eltérés tapasztalható. A kizárólag perceptuális elemzésre támaszkodó vizsgálat néhány esetben téves következtetésekhez vezetett volna a hiányosság fonológiai, illetőleg fonetikai jellegének elkülönítésében. A nyelvi kontraszt elsajátítottságának megítéléséhez a VOT értéke szükséges, de nem elégséges mutató, valamint sok esetben további adatnyerési módszerre (pl. egy adott item többszöri ismétlésére), illetőleg a tipikus fejlődésről számot adó viszonyítási pontra lenne szükség a háttérben álló okok természetének, a mintázat atipikus voltának pontosabb megítéléséhez.

Irodalom

Abramson, Arthur S. 1977. Laryngeal timing in consonant distinctions. *Phonetica* 34. 295–303.

Allen, George D. 1985. How the young French child avoids the pre-voicing problem for word-initial voiced stops. *Journal of Child Language* 12. 37–46.

Auzou, Pascal – Özsancak, Canan – Morris, Richard J. – Jan, Mary – Eustache, Francis – Hannequin, Didier 2000. Voice onset time in aphasia, apraxia of speech and dysarthria: a review. *Clinical Linguistics and Phonetics* 14. 131–150.

Ball, Martin J. – Code, Chris (eds.) 1997. *Instrumental clinical phonetics*. Whurr, London.

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 2011. A [p, t, k] mássalhangzók zöngekezdési ideje idősek és fiatalok spontán beszédében és felolvasásában. *Beszédkutatás 2011*. 61–73.

Cho, Taehong – Ladefoged, Peter 1999. Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics* 27. 207–229.

Clumeck, Harold – Barton, David – Macken, Marlys A. – Huntington, Dorothy A. 1981. The aspiration contrast in Cantonese word-initial stops: data from children and adults. *Journal of Chinese Linguistics* 9. 210–224.

Gandour, Jack – Petty, Soranee H. – Dardarananda, Rochana – Dechongkit, Sumalee – Mukngoen, Sunee 1986. The acquisition of the voicing contrast in Thai: A study of voice onset time in word-initial stop consonants. *Journal of Child Language* 13. 561–572.

Farmer, Alvirda 1997. Spectrography. In Ball, Martin J. – Code, Chris (eds.): *Instrumental clinical phonetics*. Whurr, London, 22–64.

Gierut, Judith A. – Dinnsen, Daniel 1986. On word-initial voicing: Converging sources of evidence in phonologically disordered speech. *Language and Speech* 29. 97–114.

Green, Jordan R. – Moore, Christopher A. – Higashikawa, Masahiko – Steewe, Roger W. 2010. The physiologic development of speech motor control: Lip and jaw coordination. *Journal of Speech, Language, and Hearing Research* 43. 239–255.

Gósy Mária 1984. *Hangtani és szótani vizsgálatok hároméves gyermekek nyelvében*. Nyelvtudományi Értekezések 119. Akadémiai Kiadó, Budapest.

Gósy Mária 1998. A szavak hangalakjának változása a gyermeknyelvben. *Beszédkutatás '98*. 1–39.

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

Gósy, Mária 2001. The voice onset time of the Hungarian voiceless plosives in words and in spontaneous speech. *International Journal of Speech Technology* 3-4. 155–164.

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

Gósy, Mária – Ringen, Catherine O. 2009. *Everything you always wanted to know about VOT in Hungarian*. Előadás az ICSH 2009 kongresszuson, Debrecenben, 2009 szeptember 1. http://icsh9.unideb.hu/pph/handout/Ringen_Gosy_handout.pdf (A letöltés ideje: 2012. április 6.)

Gráczi Tekla Etelka 2011. Intervokális explozívák a zöngésségi oppozíció függvényében. *Beszédkutatás 2011*. 46–60.

Gráczi Tekla Etelka 2012. *Zörejhangok akusztikai fonetikai vizsgálata a zöngésségi oppozíció függvényében*. PhD-értekezés. ELTE, Budapest.

Gráczi Tekla Etelka – Kohári Anna 2012. A zöngekezdési idő egy módszertani kérdés függvényében. In Markó Alexandra (szerk.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngekezdési időig*. ELTE BTK–MTA NYTI, Budapest, 228–248.

Hewlett, Nigel – Waters, Daphne 2004. Gradient change int the acquisition of phonology. *Clinical Linguistics and Phonetics* 18/6–8. 523–533.

Howard, Sarah – Heselwood, Barry 2011. Instrumental and perceptual phonetic analysis: The case for two-tier transcriptions. *Clinical Linguistics and Phonetics* 25/11–12. 940–948.

Hua, Zhu – Dodd, Barbara (eds.) 2006. *Phonological development and disorders in children: A Multilingual Perspective*. Multilingual Matters. Clevedon, Buffalo, Toronto.

Kager, René – van der Feest, Suzanne – Fikkert, Paula – Kerkhoff, Annemarie – Zamuner, Tania S. 2007. Representations of [voice]: Evidence from Acquisition. In van der Torre, Erik J. – Weijer, Jeroen van de (eds.): *Voicing in Dutch: (De)voicing – phonology, phonetics, and psycholinguistics*. Benjamins, Amsterdam.

Kewley-Port, Diane – Preston, Malcolm S. 1974. Early apical stop production: A voice onset time analysis. *Journal of Phonetics* 2. 195–210.

Kent, Raymond R. 1996. Hearing and believing: Some limits to the auditory-perceptual assessment of speech and voice disorders. *American Journal of Speech-Language Pathology* 5/3. 7–23.

Koenig, Laura L. 2000. Laryngeal factors in voiceless consonant production in men, women, and 5-year-olds. *Journal of Speech, Language and Hearing Research* 43/5. 1211–28.

Kuijpers, Cecile T. L. 1993. Temporal aspects of the voiced-voiceless distinction in speech development of young Dutch children. *Journal of Phonetics* 21. 313–327.

Kong, Eunjong – Beckman, Mary E. – Edwards, Jan 2007. Fine-grained phonetics and acquisition of Greek voiced stops. In Trouvain, Jürgen – Barry, William J. (eds.): *Proceedings of the XVIth ICPhS*. Saarbruecken. 865–868.

Leonard, Laurence B. 2000. *Children with specific language impairment*. The MIT Press, Cambridge–London.

Lisker, Leigh – Abramson, Arthur S. 1964. A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20. 384–422.

Lisker, Leigh – Abramson, Arthur S. 1967. Some effects of context on voice onset time in English stops. *Language and Speech* 10. 1–28.

Lowenstein, Joanna H. – Nittrouer, Susan 2008. Patterns of acquisition of native voice onset time in English-learning children. *Journal of the Acoustical Society of America* 124/2. 1180–1191.

Lőrik József 1982. Állami gondozott óvodások szegmentumállományának fejlettsége. *Gyógypedagógiai Szemle* 10/2. 100–112.

Macken, Marlys A. – Barton, David 1980a. The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language* 7. 41–74.

Macken, Marlys A. – Barton, David 1980b. The acquisition of the voicing contrast in Spanish: a phonetic and phonological study of word-initial stop consonants. *Journal of Child Language* 7. 433–458.

McLeod, Sharynne – Goldstein, Brian (eds.) 2012. *Multilingual aspects of speech sound disorders in children. Communication disorders across languages 6.* Multilingual Matters, Bristol, New York, Ontario.

Maxwell, Edith M. – Weismer, Gary 1982. The contribution of phonological, acoustic and perceptual techniques to the characterization of a misarticulating child's voice contrast for stops. *Applied Psycholinguistics* 3. 29–43.

Munson, Benjamin – Edwards, Jan – Schellinger, Sarah – Beckman, Mary E. – Meyer, Marie K. 2010. Deconstructing phonetic transcription: Covert contrast, perceptual bias, and an extraterrestrial view of Vox Humana. *Clinical Linguistics and Phonetics* 24/4–5. 245–260.

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In MacNeilage, Peter F. (ed.): *The production of speech.* Springer-Verlag, New York, 189–216.

Ohala, John J. 2011. Accommodation to the aerodynamic voicing constraint and its phonological relevance. In Lee, Wai-Sum – Zee, Eric (eds.): *Proceedings of the XVIIth ICPhS.* Hong Kong. 64–67.

Scobbie, James M. – Gibbon, Fiona – Hardcastle, William J. – Fletcher, Paul 2000. Covert contrast as a stage in the acquisition of phonetics and phonology. In Broe, Michael B. – Pierrehumbert, Janet B. (eds): *Papers in Laboratory phonology V: Acquisition and the lexicon.* Cambridge University Press, London, 192–203.

Sebestyénné Tar Éva 2006. *A 3-6 éves kori fonológiai fejlődés kronológiai mintázata a magyarban.* Open Art, Budapest.

Sebestyénné Tar Éva 2008. *Az atipikus nyelvi fejlődés szegmentális fonológiai szintjének elemzése.* PhD-értekezés. PTE, Pécs.

S. Meggyes Klára 1971. *Egy kétéves gyermek nyelvi rendszere.* Nyelvtudományi Értekezések 73. Akadémiai Kiadó, Budapest.

Smit, Ann B. – Hand, Linda – Freilinger, Joseph – Bernthal, John E. – Bird, Ann 1990. The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders* 55. 779–798.

Smith, Anne 2010. Development of neural control of orofacial movements for speech. In Hardcastle, William J. – Laver, John – Gibbon, Fiona E. (Eds.): *Handbook of phonetic sciences.* Blackwell, Oxford.

Tyler, Ann A. – Edwards, Mary L. – Saxman, John H. 1990. Acoustic validation of phonological knowledge and its relationship to treatment. *Journal of Speech and Hearing Disorders* 55. 251–261.

KORASZÜLÖTT GYERMEKEK NYELVHASZNÁLATI JELLEMZŐIRŐL

Krepsz Valéria

Bevezetés

Beszélgetés során megnyilatkozásainkra számos tényező van hatással, így az adott beszédhelyzet, a hely, az idő, a társalgás célja, a beszédpartner és a beszélő saját jellemzői (például neme, kora, testi és lelki állapota stb.). A sikeres kommunikáció feltétele a hallgató és a beszélő között a beszéd produkciójának és percepciójának összehangolása, amelyben fontos szerepet játszik az egyéb beszédet kísérő jelenségek mellett a beszédtempó is. A beszédsebesség olyan jellemző, amelyet az egyéni adottságok, illetve számos változó és nem változó tényező befolyásol, így az életkor, a foglalkozás, a nem vagy az adott beszédhelyzet, téma (Gósy 2004; Markó 2005; Bata–Gráczi 2009; Bóna 2009). A jelenség alapja az, hogy milyen gyorsak vagy lassúak az artikulációs működéseink (Gósy 2004). A beszédtempó nemcsak az időzítést érinti, hanem fonológiai és artikulációs szinten is jelen van a beszédfolyamatokban, emellett pedig hatással van a szemantikai és a grammatikai feldolgozásra is (Bóna 2007).

A gyermek (társas) kommunikációs képessége fokozatosan fejlődik a beérkező ingerek hatására. A tipikus gyermekkori kognitív fejlődést, így a kommunikációs készséget már akár az első napokban megszerzett tapasztalatok is befolyásolják (Szanati 2011). Az esetleges, már a születéskor fellépő biológiai károsodások később értelmi, strukturális, funkcionális, esetleg neurológiai sérülést eredményezhetnek, amelyek hatással lehetnek a beszéd produkciós és percepciós oldalára is (Szanati–Nagy 2006). A koraszülöttség problémája mind magyar, mind pedig nemzetközi viszonylatban jelen van, az újszülött populáció 10-11%-át érinti (Estefánné Varga 2003; Aszmann 2006). A WHO definíciója alapján koraszülöttnek tekintenek minden 2500 g-nál kisebb élve született gyermeket, akinek gesztációs ideje a 37 hetet (259 napot) nem haladta meg, vagy akin az éretlenség kétségtelen jelei, például bőrszín, megállapíthatók (Wohlmuth 1986). Bár a koraszülöttség és a kognitív fejlődés, így a nyelvi terület fejlettsége, illetve esetleges elmaradása közvetlen kapcsolatára máig nincsenek cáfolhatatlan bizonyítékok, néhány tanulmány rámutatott, hogy a koraszülött gyermekek nyelvi fejlődése időre született, azonos korú társaikéhoz viszonyítva megkésett (pl. DeHitsch–Jansky 1965; Bailey–Wolery 1989; McCormick–Scheifelbusch 1990). Ezzel szemben más kutatások arra az eredményre jutottak, hogy a nyelvi és kommunikációs elmaradás nem gyakoribb a koraszülött gyermekeknél, mint az időre és normál súllyal szüle-

tett társaiknál, valamint az időben diagnosztizált eltérések 2-3 éves korra fokozatosan megszűnnek (pl. Greenberg–Crnic 1988; Szanati–Nagy 2006). A fejlődés üteme, az esetleges elmaradás szintje több tényezőtől is függ. Koraszülött gyermekek vizsgálata során fontos figyelembe venni a koraszülöttség okát, illetve a fejlődést befolyásoló egyéb biológiai és környezeti tényezőket is (Szanati 2011).

Az elmúlt néhány évtizedben számos kutatás vizsgálta a koraszülöttség és a kognitív fejlődés hosszú távú kapcsolatát. Az eredmények rámutattak, hogy bizonyos területeken, így a mentális, illetve motoros fejlődésben, avagy a szülővel való interakcióban a koraszülött gyermekek különböznek időre született társaiktól. A vizsgálatok eredményei alapján elmondható, hogy a szövődmények nélküli koraszülés önmagában nem jelent sem idegrendszeri károsodást (Boda 1981), sem az intelligencia csökkenését, bár azok a tényezők, amelyek a koraszüléshez vezetnek, okozhatnak perinatális ártalmakat is. Az óvodáskori és iskolai hátrány hátterében is gyakran az eltérő viselkedésformák, illetve a nyelvi területen való lemaradás áll. Gallagher és Watkin (1998) kutatásai szerint a nyelvi nehézségeket gyakran nem veszik figyelembe a gyermek fejlődésének vizsgálatakor, hiszen a lehetséges idegrendszeri kihatások egészen addig nem mutatkoznak meg, amíg a nyelvi fejlődés nem éri el a kinyilatkoztatások szintjét.

Magyar vizsgálatok eredményei szerint a beiskolázás előtt, 5-6 éves gyermekek körében végzett beszédpercepciós vizsgálatok során a kis születési súlyúak gyengébb teljesítményt értek el, mint az éretten születettek, ám az előbbiek memóriája is jónak bizonyult (Beke et al. 2001). Az észlelés és a megértés eredményei összefüggést mutattak az intelligenciateszt eredményeivel, amely alapján elmondhatjuk, hogy a beszédfejlődés üteme jól jelzi előre számunkra a későbbi kognitív teljesítmény alakulását. Emellett érdemes kiemelni, hogy a különböző beszédészlelési és beszédmegértési tesztek összefüggést mutatnak az újszülöttkori komplikációk súlyosságával, a születési súllyal, a hipoxiával (a vér csökkent oxigéntelítettségével) és a lélegzetvétel idejével. Csiky, Rajk, Simon és Dorner kísérletük során (1981) 21 igen kis súlyú újszülöttet vizsgáltak az iskolaérettség szempontjából. A kommunikáció, illetve a beszéd területén azt tapasztalták, hogy az általuk felmért gyermekek esetében az artikuláció renyhébbnek, a beszédtempó gyorsabbnak bizonyult, emellett beszédük monotonabb, ritmustalanabb és dallamtalanabb volt, továbbá halkabb hangerőt használtak, mint azonos korú társaik. A szerzők szerint e mögött a megzavart biológiai érés, valamit a születéskor fellépő károsodás áll. Mindezen folyamatok gyakran kedvezőtlenül érintik a bal temporális agyféltekét, amely többek között a nyelvi funkciókat is raktározza (Csépe 2005).

A jelen kutatás célja ötéves koraszülött gyermekek beszédprodukciójának felmérése a beszédsebesség és a reakcióidő szempontjából, továbbá néhány beszédpercepciós folyamat és időzítésük elemzése. Arra a kérdésre igyekeztünk választ találni, hogy a 37. hét előtt és 2500 g súly alatt világra jött gyer-

mekek kísérleti adatai eltérnek-e azonos korú és időre született társaikétól vagy nem. Három hipotézist állítottunk fel. 1. A koraszülött gyermekek reakcióideje eltér a különböző típusú feladatok során, és minden esetben hosszabb az éretten született társaikéhoz képest. 2. A koraszülött gyermekek gyengébb teljesítményt érnek el a beszédpercepciós feladatok megoldása során. 3. A koraszülött gyermekek beszédtempója meghaladja az éretten született gyermekek beszédtempóját a három feladatban és a spontán beszéd során is.

Anyag, módszer, kísérleti személyek

Kutatásunkban 15 tipikus fejlődésű, ép hallású és ép artikulációjú koraszülött gyermek (8 lány és 7 fiú) vett részt, mindannyian ötévesek. Adataikat 15 normál időre és normál súllyal született, ép hallású és ép artikulációjú, tipikus fejlődésű kontroll-gyermekcsoport eredményeivel vetettük össze. Minden gyermektől spontán beszédanyagot rögzítettünk digitális hangfelvevő segítségével a megszokott otthoni, illetve óvodai környezetben. Elsőként egy képeskönyvet néztünk, majd kedvenc játékukról, időtöltésükről kérdeztük őket. Ezután elvégeztük velük a GMP sztenderdizált diagnosztikai vizsgálat (Gósy 1995/2006) GMP4, GMP5 és a GMP12 részfeladatát, amelyek alkalmasak mind a beszédtempó, mind a reakcióidő mérésére. A GMP4 feladat esetében 10 szűk frekvenciatartományú (2200–2700 Hz) mondat pontos megismétlése a cél, így vizsgálható a beszédészlelés fonetikai szintje. A GMP5 feladat hanganyagában 10 mondat esetében a természetes beszéd sebességét műszeresen 25%-kal felgyorsították. Az adatközlőknek az volt a feladatuk, hogy ismételjék meg az elhangzottakat a lehető legnagyobb pontossággal. A gyorsított mondatok visszamondásának feladata nehezített az óvodások számára, szókincse meghaladja az ötéves gyermekektől elvárható ismeretanyagot, ám a feladat megoldása során nem cél a szöveg megértése, csak megismétlése. A harmadik feladat, a GMP12-es vizsgálatában a gyermekek egy mesét hallottak, majd 10 megértést ellenőrző kérdésre kellett válaszolniuk. Ebben a feladatban a válaszadáskor önálló beszédprodukciót vártunk el, amelyhez szükséges volt a hallott szöveg megértése a szemantikai és szintaktikai szerkezetek értelmezésére, asszociációs működések végrehajtására; illetve a rövid és hosszú távú memória aktiválása is. Összesen 180 percnyi hanganyagot dolgoztunk fel, illetve annotáltunk a Praat szoftverrel (Boersma–Weenink 2010). A statisztikai elemzésekhez egytényezős varianciaanalízist és Pearson-korrelációt használtunk (az SPSS 19.0 szoftverben).

Mértük a gyermekek reakcióidejét a GMP részfeladatok megoldásában, összehasonlítottunk a koraszülött és a kontrollcsoportok eredményeit. A Praat szoftver segítségével a felvételekben megmértük a szüneteket és a jelidőt, meghatároztuk a beszédhangok számát, majd kiszámoltuk a gyermekek beszédtempó-értékeit a rögzített spontán beszédanyag alapján, végül összehasonlítottuk a három feladat adatait.

Eredmények

A koraszülött és az éretten született gyermekek reakcióértékeit hasonlítottuk össze a három GMP-feladatban a mondatismétlések és az ellenőrző kérdésekre adott válaszok tekintetében. A szűk frekvenciás és gyorsított mondatok reakcióértékei igen közel álltak egymáshoz, a mese ellenőrző kérdéseire adott válaszoknál közel másfélszeres reakcióidő-értéket adatoltunk a koraszülött gyermekeknél. Az eltérő feladatok közötti különbség a kontrollcsoport esetében is kimutatható, átlagosan 2,5-szer annyi időre volt szükség a szövegértést ellenőrző kérdések válaszaihoz, mint a szűk frekvenciás és a gyorsított mondatok megismétléséhez (1. táblázat).

1. táblázat: A két csoport reakcióideje (s): átlag és átlagos eltérés

	Kontrollcsoport				Koraszülött csoport			
	Átlag	Átl. elt.	Min.	Max.	Átlag	Átl. elt.	Min.	Max.
GMP4	0,65	0,34	0,00	3,11	1,02	0,97	0,05	8,33
GMP5	0,56	0,29	0,00	1,93	1,08	0,91	0,00	7,38
GMP12	1,48	1,49	0,00	9,37	1,56	1,29	0,00	9,96

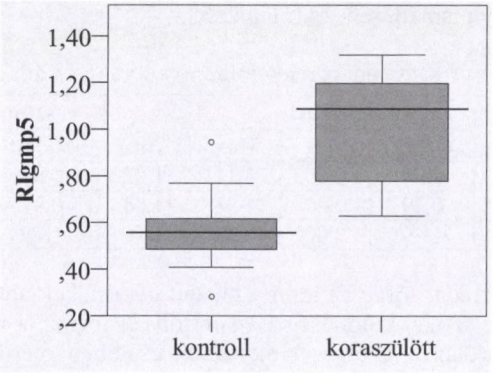
Mindhárom feladat során az időre született gyermekek adtak gyorsabb válaszokat, ám statisztikai különbséget a kontroll- és a rizikócsoport teljesítménye között kizárólag a gyorsított mondatok esetében mértünk ($F(1, 29) = 20{,}033$; $p = 0{,}001$), lásd 1. ábra.

A gyorsított mondatok megismétlése a beszélőhöz való alkalmazkodás jelenségéből adódóan arra késztetheti az adatközlőket, hogy maguk is gyorsabban ismételjék meg az elhangzottakat. Az időre és éretten született gyermekek esetében statisztikailag is igazolható közepes erősségű összefüggés (Pearson-korreláció: $r = 0{,}558$) volt kimutatható a gyorsított mondatok helyes ismétlése és a szövegértésre adott válaszok tempója között. Azok a gyermekek, akik kevesebb idő alatt képesek megválaszolni a megértést ellenőrző kérdéseket, nagyobb arányban képesek helyesen visszamondani a felgyorsított mondatokat.

A szűk frekvenciás, a gyorsított mondatok ismétlésében és a mesehallgatást követő válaszadásban mért reakcióidőket összehasonlítottuk a feladatok megoldásának eredményességével mindkét vizsgált csoport esetében. Mind a koraszülött, mind pedig az időre született gyermekek a szűk frekvenciás mondatokat mondták vissza a legnagyobb arányban helyesen, illetve ebben a feladatban statisztikailag igazolható különbség mutatkozott a koraszülött és időre született gyermekek teljesítménye között ($F(1, 29) = 4{,}699$; $p = 0{,}039$). A koraszülöttek helyes ismétlésének átlaga 70% (átl. elt.: 11,8%), a kontrollcsoporté pedig 84% (átl. elt.: 22%).

A koraszülött gyermekeknél a spontánul gyorsabban beszélők gyorsabban ismételték helyesen a szűkfrekvenciás mondatokat (az összefüggés erős: $r = 0{,}945$). Közepes erősségű összefüggést találtunk a koraszülött gyermekek esetében a szűk frekvenciás mondatok visszamondásának eredményessége és

a mese kérdéseire adott helyes válaszok aránya ($r = 0{,}590$), illetve erős összefüggést a szűk frekvenciás mondatok sikeres visszamondása és a gyorsított mondatok reakcióideje között ($r = 0{,}795$). A koraszülött gyermekek esetében a különböző részfolyamatok közepes és erős összefüggéseket mutattak az adott életkorban (hat esetben), ez nem csak a feladat sikeres teljesítésében, hanem az időzítésben is megmutatkozott. Az éretten született gyermekek körében az eredményesség csak a szövegértést ellenőrző kérdésekre adott válaszokban mutatott közepes erősségű összefüggést a reakcióidővel ($r = 0{,}634$).



1. ábra
Gyorsított mondatok ismétlésének reakcióidő-értékei

A koraszülött gyermekeknél a spontánul gyorsabban beszélők gyorsabban ismételték helyesen a szűkfrekvenciás mondatokat (az összefüggés erős: $r = 0{,}945$). Közepes erősségű összefüggést találtunk a koraszülött gyermekek esetében a szűk frekvenciás mondatok visszamondásának eredményessége és a mese kérdéseire adott helyes válaszok aránya ($r = 0{,}590$), illetve erős összefüggést a szűk frekvenciás mondatok sikeres visszamondása és a gyorsított mondatok reakcióideje között ($r = 0{,}795$). A koraszülött gyermekek esetében a különböző részfolyamatok közepes és erős összefüggéseket mutattak az adott életkorban (hat esetben), ez nem csak a feladat sikeres teljesítésében, hanem az időzítésben is megmutatkozott. Az éretten született gyermekek körében az eredményesség csak a szövegértést ellenőrző kérdésekre adott válaszokban mutatott közepes erősségű összefüggést a reakcióidővel ($r = 0{,}634$).

Elemeztük a válaszadók beszédtempóját a különböző feladatok végrehajtása során. A szűk frekvenciás mondatok esetében 12,5 hang/másodperc, a gyorsított mondatok visszamondása során átlagosan 15,0 hang/másodperc volt a beszédtempó. A spontán beszéd tempójához a mesékkel kapcsolatos kérdésekre adott válaszok beszédsebessége állt a legközelebb (2. táblázat),

ami a koraszülött gyermekek esetében statisztikailag is igazolható összefüggést mutatott ($r = 0,531$).

A legnagyobb beszédtempó-értékek nem a gyorsított mondatok visszamondásakor, hanem a szűkfrekvenciás mondatok esetében jelentek meg. A reakcióidő eredményeihez hasonlóan az eltérés a feladat típusával magyarázható, ugyanis a nehezített szavak és kifejezések visszamondása nagyobb koncentrációt, így hosszabb időt és nagyobb odafigyelést követelt meg a gyermektől, a tervezési szakasz több időt vett igénybe, mint maga a produkció. Mind a szűk frekvenciás, mind a gyorsított mondatok esetében az időre született gyermekek beszéde volt gyorsabb koraszülött társaikéhoz viszonyítva.

2. táblázat: A két vizsgált csoport beszédtempója (hang/másodperc)

	Szűk frekvenciás mondatok	Gyorsított mondatok	Mesére adott válaszok	Spontán beszéd
Koraszülött csoport	10,042	9,748	8,931	8,927
Kontrollcsoport	10,506	10,219	8,902	8,911

Következtetések

Kutatásunk a koraszülött és a velük azonos korú, időre és éretten született gyermekek reakcióidőit, beszédészlelését, illetve beszédtempóját hasonlította össze különböző típusú feladatok megoldása során. Noha a nemzetközi és hazai kutatásokban a kutatók eltérő eredményeket kaptak a koraszülöttség és a különböző mentális képességek, így a beszédkészség fejlődésének egymásra hatását illetően, a kutatás megkezdése előtt különbséget feltételeztünk az általunk vizsgált koraszülött gyermekek és a kontrollcsoport teljesítménye között. Számos vizsgálat eredménye szerint ugyanis a különbség nem a nyelvi teljesítményben, hanem az eltérő beszéd- és kommunikációs stratégiákban, a különböző viselkedési jegyekben keresendő.

Eredményeink szerint különbség mutatkozott a koraszülött, illetve az éretten született gyermekek beszédtempó-értékei között, emellett a koraszülött gyermekek esetében eltérést tapasztaltunk a három különböző feladat és a spontán beszéd beszédtempóját tekintve. Míg a kontrollcsoport esetében csak igen kevés összefüggést találtunk a különböző részképességek között, addig a koraszülött gyermekeknél számos különböző tényező, a reakcióidő, illetve a beszédsebesség adatai összefüggést mutattak. Míg az éretten született csoport esetében a kognitív fejlődéssel párhuzamosan a különböző részképességek önállósultak, addig a koraszülött gyermekeknél még ötéves korban is igen szoros összefüggést tapasztaltunk ezek között. Abban az esetben azonban, ha valamely képesség elmaradást vagy zavart mutat, a részképességek egymásra utaltsága akadályozza más területek fejlődését is, ilyen módon akár egy nagyobb terület elmaradását vagy globális elmaradást eredményezve. A szűk frekvenciás és a gyorsított mondatok esetében a koraszülött csoport a kontrollcsoporthoz viszonyítva alacsonyabb beszédtempó-értékeket ért el, míg a spon-

tán beszédben és az ahhoz legközelebb álló mese kérdéseire adott válaszok esetében nem mutatkozott jelentős eltérés a csoportok között. Azt gondoltuk, hogy a beszéd temporális sajátosságait érintő különbség nem elsősorban a koraszülöttség tényével magyarázható, nagyobb szerepe van a különböző feladattípusoknak, az eltérő megoldási stratégiáknak, illetve a nagy egyéni különbségeknek, amelyek mindkét vizsgált csoportra jellemzők voltak.

Különbséget feltételeztünk a két csoport között a beszédtempó tekintetében is. A hipotézis nem igazolódott, ugyanis közel azonos beszédsebességértékeket adatoltunk a három eltérő feladat és a spontán beszéd esetében is.

A reakcióidők tekintetében elmaradást vártunk a koraszülöttek teljesítményében az időre született gyermekekhez viszonyítva. Ez a hipotézis részben nyert igazolást. Míg a szűk frekvenciás mondatoknál a gyorsabb teljesítmény sikeresebb megoldást is jelentett, a gyorsított mondatok esetében a hosszabb reakcióidő eredményesebb stratégiának mutatkozott.

Noha a jelen kutatás csak részben mutatott különbségeket a koraszülött és az éretten született gyermekcsoport között, a beszédpercepció területén kimutatott elmaradás, valamint a részfolyamatok szoros összefüggése a nem időre született gyermekeknél a beiskolázás előtti vizsgálatok szükségességére hívja fel a figyelmet.

Irodalom

Aszmann Anna 2006. A nő-, anya-, csecsemő- és gyermek-egészségvédelem. In Ádány Róza (szerk.): *Megelőző orvostan és népegészségtan*. Medicina Kiadó, Budapest. 483–484.

Bailey, David – Wolery, Mark 1989. Assessing infants and preschoolers with handicaps. Merrill Publishing Company, Columbus.

Bata Sarolta – Gráczi Tekla Etelka 2009. Hatással van-e a beszédpartner életkora a beszélő beszédének szupraszegmentális jellegzetességeire. In Keszler Borbála – Tátrai Szilárd (szerk.): *Diskurzus a grammatikában, grammatika a diskurzusban*. Tinta Kiadó, Budapest. 74–83.

Beke Anna – Németh Tünde – Gósy Mária 2001. Koraszülöttek beszédészlelése és beszédmegértése óvodás korban. *Pediáter* 10. 71–74.

Boda Domokos 1981. *Gyermekgyógyászat*. Medicina Kiadó, Budapest.

Boersma, Paul – Weenink, David 2005. *Praat: doing phonetics by computer*. http://www.praat.org/ (Letöltés: 2010. szeptember 29.)

Bóna Judit 2007. *A felgyorsult beszéd produkciós és percepciós sajátosságai*. PhD-disszertáció. ELTE, Budapest. http://doktori.btk.elte.hu/lingv/bona/Phd_dolgozat_BonaJudit.pdf (Letöltés: 2012. augusztus 30.)

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

Csépe Valéria 2005. *Kognitív fejlődés-neuropszichológia*. Gondolat Kiadó, Budapest.

Csiky Erzsébet – Rajk András – Simon Mária – Dorner Márta 1981. Igen kissúlyú koraszülöttek utóvizsgálata az iskolaérettség szempontjából. *Magyar Pszichológiai Szemle* 38/4. 359–373.

DeHirsch, Katrina – Jansky, Jeannette – Langford, William 1965. The prediction of reading, spelling and writing disabilities in children, a preliminary study. Columbia University, New York.

Estefánné Varga Magdolna 2003. Koraszülött gyermekek longitudinális, pszichodiagnosztikai vizsgálatának és iskolai beválásának eredményei. *Alkalmazott Pszichológia* 5. 5–14.

Gallagher, Tanya – Watkin, Kenneth 1998. Prematurity and language developmental risk: Too young or too small? *Topics in Language Disorders* 18. 34–50.

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

Gósy Mária 1995/2006. *GMP-diagnosztika. A beszédészlelés és a beszédmegértés folyamatának vizsgálata, fejlesztési javaslatok*. Nikol, Budapest.

Greenberg, Mark – Crnic, Keith 1988. Longitudinal predictors of developmental status and social interaction in premature and full-term infants at age two. *Child Development* 59. 554–570.

Markó Alexandra 2005. A temporális szerkezet jellegzetességei eltérő kommunikációs helyzetekben. *Beszédkutatás 2005*. 63–77.

McCormick, Linda – Schiefelbusch, Richard 1990. *Early language intervention: An introduction*. Merrill Publishing Company, Columbus.

Szanati Dóra 2011. A logopédus-egészségpszichológus szerepe a koraszülött gyermekek utánvizsgálatában. *Gyógypedagógiai Szemle* 39/3–4. http://www.prae.hu/prae/gyosze.php?menu_id=102&jid=37&jaid=541 (Letöltés: 2012. 08. 02.)

Szanati Dóra – Nagy Beáta 2006. A koraszülöttség mint a preverbális képességek fejlődését befolyásoló tényező. *Gyógypedagógiai Szemle* 34/1. 23–37.

Wohlmuth Gertrud 1986. Koraszülöttek. In Véghelyi Péter – Kerpel-Fronius Ödön (szerk.): *Az újszülött*. Akadémiai Kiadó, Budapest. 455–517.

SPRACHREZEPTION IM BILINGUALEN KINDERGARTEN

Anja Bereznai

Einleitung

Viele Jahre lang war die sprachliche Förderung von Kindern im Kindergarten gleichbedeutend mit der Förderung der Muttersprache. Immer häufiger stehen Kindergärten jedoch auch vor der Aufgabe, Kinder in einer weiteren Sprache zu fördern bzw. ihnen deren Anfangsgründe als Fremd- bzw. Zweitsprache zu vermitteln; dabei kann die zweite Sprache die Umgebungssprache, eine Minderheitensprache oder eine (andere) Weltsprache sein. Dies hat auch dazu geführt, dass sich im In- und Ausland zunehmend mehr Forschungsarbeiten mit dem Thema früher Zweitspracherwerb befassen. Diese untersuchen beispielsweise Ähnlichkeiten und Unterschiede zwischen Erst- und Zweitspracherwerb und die Abhängigkeit des L2-Erwerbs von Faktoren wie Alter und Geschlecht des Kindes sowie von der Qualität und Quantität des Inputs (z.B. Kersten 2002; Müller 2004; Wenzel 2006; Imre–Gráczi 2009). Allerdings ist festzustellen, dass sich ein Großteil der Publikationen in diesem Bereich hauptsächlich oder ausschließlich mit der Sprachproduktion befasst, während die Zahl von Veröffentlichungen zum Thema Rezeption wesentlich geringer ist.

Werden Kinder bereits zu Beginn des Kindergartenalters mit einer zweiten Sprache konfrontiert, so entwickelt sich die Sprachrezeption in der L2 in vielerlei Hinsicht ähnlich wie in der Erstsprache (Rothweiler–Rugberg 2011; Rothweiler 2007). So lassen sich zahlreiche Strategien identifizieren, die sich auch bei Kindern im Erstspracherwerb beobachten lassen, wie z.B. das Erschließen von Bedeutungen aus der Situation, Schlüsselwortstrategien oder phonetisch-phonologische Assoziationen (Müller 2004). Gleichzeitig führen die bereits erfolgte kognitive Entwicklung, das sprachliche Wissen in der Muttersprache sowie das schon erworbene Weltwissen insbesondere beim Sprachverstehen häufig schon nach kurzer Kontaktdauer mit der L2 zu deutlichen Fortschritten (Schulz 2007). Als weitere Besonderheit des L2-Erwerbs ist anzusehen, dass das Verstehen sprachlicher Mittel nicht in jedem Fall der Produktion vorausgeht: So werden besonders im Kindergartenalltag oft verwendete Sätze und Satzbausteine sowie Lieder und Reime vielfach bereits von den Kindern (re)produziert, obwohl sie deren Bedeutung nicht oder nur annäherungsweise verstehen (Kersten 2002; Wenzel 2006). Es ist anzunehmen, dass diese von Pienemann (2002: 4) als „unanalysierte Einheiten" be-

zeichneten sprachlichen Mittel den Kindern dabei helfen, syntaktische Strukturen und morphologische Merkmale zu erschließen und sich diese anzueignen (Apeltauer 2007). Auch im Bereich der Sprachwahrnehmung zeigt sich, dass Kinder die zweite Sprache schon in diesem frühen Alter auf der Grundlage ihrer Erfahrungen mit der Erstsprache erwerben: Die im Verlauf des L1-Erwerbs entstandene muttersprachliche Perzeptionsbasis erschwert die Sprachwahrnehmung in der L2, insbesondere die Differenzierung von L2-spezifischen Lauten (Imre – Gráczi 2009; Steinbrink – Spitzer 2006); vermutlich liegen derartige Probleme bei der Sprachwahrnehmung auch Untersuchungsergebnissen zugrunde, denen zufolge schon bei drei- bis vierjährigen Kindern in Immersionskindergärten phonologische Interferenzen in der L2 festzustellen sind (Kersten 2002).

Ziel der vorliegenden Untersuchung war zum einen, festzustellen, inwieweit es bei einem L2-Erwerb im institutionellen Rahmen Unterschiede und Übereinstimmungen zwischen der Sprachrezeption in der L1 und der in der L2 gibt. Hierbei wurde davon ausgegangen, dass bei der Sprachwahrnehmung die in L1 und L2 erbrachten Leistungen stärker korrelieren als beim Sprachverstehen. Zudem sollte ausgelotet werden, welche Möglichkeiten und Grenzen sich bei der Testung der Sprachrezeption in der L2 mithilfe eines für Muttersprachler ausgearbeiteten Tests ergeben.

Untersuchungsdesign

Für die Untersuchung wurden zehn Kinder einer Gruppe eines zweisprachigen Kindergartens in Ungarn ausgewählt. Die getesteten vier Jungen und sechs Mädchen waren zum Untersuchungszeitpunkt 5;2 bis 6;4 Jahre alt und wurden seit 26 Monaten von zwei zweisprachigen Kindergartenpädagoginnen betreut. Beide Sprachen wurden jeweils in bestimmten Situationen im Kindergartenalltag verwendet. Alle Kinder sprechen Ungarisch als Muttersprache. Zwei der Mädchen werden von den Eltern bilingual erzogen (ungarisch-englisch bzw. ungarisch-französisch) (nähere Informationen zum Untersuchungsdesign s. Bereznai 2012).

Die Kinder wurden jeweils im Rahmen einer Sitzung nacheinander mit der ungarischen und der deutschen Version des GMP getestet. Der GMP ist ein von Mária Gósy ausgearbeiteter standardisierter Test zur Sprachrezeption für 3-12jährige Kinder ungarischer Muttersprache (Gósy 2006), der von Angéla Imre als GMP-d für Kinder deutscher Muttersprache adaptiert wurde (Imre 2006). Für die vorliegende Untersuchung wurden fünf Subtests zur Sprachwahrnehmung (GMP 3, 5, 6, 10, 16, 17) sowie der Subtest zum Satzverstehen ausgewählt. Die Aufgaben im ungarischen Teil wurden auf Ungarisch und die im deutschen Teil auf Deutsch gestellt. Mehrere Subtests wurden aufgrund von Überschneidungen zwischen den Versionen oder auch wegen schwindender Motivation und Aufmerksamkeit der Kinder um einige Wörter bzw. Sätze reduziert. Korrelationen zwischen einzelnen quantifizierbaren Ergeb-

nissen der Untersuchung wurden mithilfe der Tabellenkalkulation NeoOffice Calc bestimmt (Korrelationskoeffizient nach Pearson).

Ergebnisse

Im Subtest GMP 6 werden den Kindern natürliche Sätze präsentiert, die sie reproduzieren sollen. Hiermit soll die Sprachwahrnehmung insgesamt überprüft werden. Eine Nutzung der Verstehensebene ist in diesem Subtest möglich, aber nicht unbedingt notwendig. Im GMP 6 erreichten acht Kinder altersgemäße Ergebnisse, die anderen beiden lagen um zehn Prozentpunkte darunter. Bei der Auswertung des GMP-d 6 ergab sich das Problem, dass bei herkömmlicher Auswertung (10% pro vollständig fehlerlos wiederholtem Satz) nicht differenziert werden konnte, ob z.B. nur eine Silbe nicht korrekt reproduziert wurde oder das Kind die Antwort vollständig verweigerte. In der vorliegenden Untersuchung wurde daher ein alternativer Auswertungsmodus gewählt, bei dem in allen Subtests durchgängig die Silbe als elementare Perzeptionseinheit zugrunde gelegt wurde (Bredel 2009; vgl. auch Imre–Gráczi 2009). Abb. 1 und Abb. 2 zeigen die Ergebnisse der ungarisch- und der deutschsprachigen Version des GMP6 bei Auswertung auf der Satz- und auf der Silbenebene.

Erwartungsgemäß schnitten die Kinder auch bei Auswertung auf der Silbenebene im GMP-d 6 mit durchschnittlich 46,5% wesentlich schlechter ab als in der ungarischen Version (99,6%), jedoch ergibt sich für den GMP-d 6 ein wesentlich differenzierteres Bild mit einer größeren Streuung. Die nur schwache Korrelation von $r = 0{,}39$ ($p > 0{,}05$) ist möglicherweise darauf zurückzuführen, dass der GMP6 eigentlich für jüngere Kinder gedacht ist, die ungarische Version daher für die hier untersuchten Kinder zu einfach war und ein dementsprechend undifferenziertes Bild über ihre Leistungsfähigkeit geliefert hat.

Die qualitative Auswertung der Antworten in der deutschen Version zeigt die Verwendung verschiedener Strategien, wie sie zum größten Teil auch im Erstspracherwerb vorkommen. Teilweise lassen diese auf Wahrnehmungsprobleme schließen, beispielsweise im Falle der verzerrten oder unartikulierten Wiedergabe von Wörtern oder Sätzen, bei der größtenteils Betonung und Silbenzahl beibehalten oder einzelne Silben durch Pausen ersetzt werden. Nicht eindeutig der Wahrnehmungs- oder der Produktionsebene zuzuordnen sind die häufig vorkommenden Elisionen (Auslassungen) und Substitutionen (Ersetzungen) eines oder mehrerer Laute in sonst korrekt wiederholten Wörtern (diese Strategien sind im Übrigen auch in Testergebnissen jüngerer Kinder in der L1 zu finden, vgl. Gósy 2007).

Andere Strategien lassen annehmen, dass bei der Reproduktion der Sätze auch die Verstehensebene beteiligt war. Beispielsweise wurden mehrere Wörter zu einem klanglich ähnlichen und mutmaßlich bekannteren deutschen Wort zusammengezogen (z.B. *Mittagessen* für „zu Mittag gegessen"), unbe-

kannte Wörter durch bereits bekannte substituiert (z.B. *sagen* für „jagen") oder der unbestimmte durch einen bestimmten Artikel ersetzt.

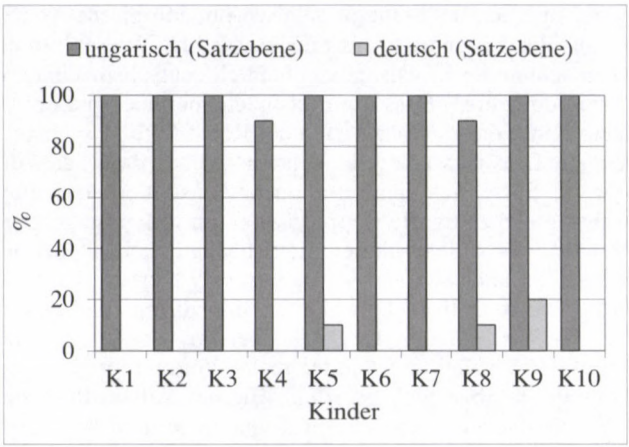


Abb. 1
Ergebnisse des GMP6 bei Auswertung auf Satzebene

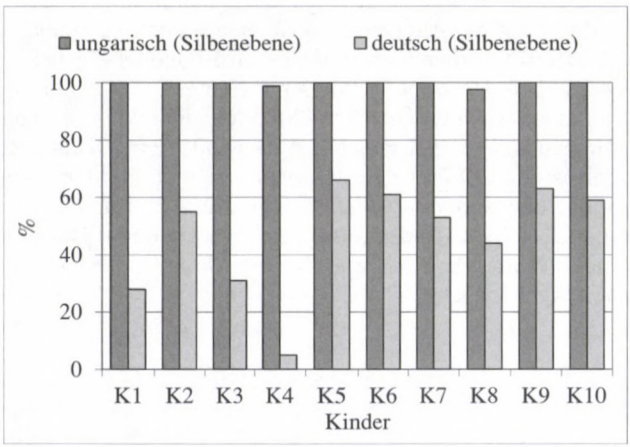


Abb. 2
Ergebnisse des GMP6 bei Auswertung auf Silbenebene

In vielen Fällen wurden nur einzelne Wörter eines Satzes wiederholt, wobei es sich fast ausschließlich um Substantive (teilweise auch Verben) handelte, beispielsweise *Kühe Wiese* für „Die Kühe stehen auf der Wiese". Bei

den in sonst fehlerfrei wiederholten Sätzen elidierten Wörtern überwiegen Funktionswörter im Verhältnis 8:2 gegenüber Inhaltswörtern. Dass die Kinder klar zwischen den beiden Sprachen zu unterscheiden wissen, zeigte sich daran, dass nur in einem Fall eine phonetisch-phonologische Assoziation mit einem Wort aus der Erstsprache auftrat (*vízesés* für „Wiese"); in allen anderen Fällen versuchten die Kinder, ausschließlich deutschsprachige Sätze bzw. Wörter zu (re)produzieren. Dass die Erstsprache dennoch präsent war, zeigte sich bei einem Kind, das die Sätze nach der Reproduktion spontan ins Ungarische übersetzte. Dies ist wiederum auch ein Zeichen dafür, dass das Kind in der Lage war, die Sätze nicht nur wahrzunehmen, sondern auch zu verstehen, wobei ihm aber gleichzeitig die Reproduktion nur teilweise gelang. Ein Beispiel hierfür ist der Satz „Jetzt blühen die Schneeglöckchen", der als *Kinőttek a virágok* übersetzt, aber als *Jetzt* (+ nicht transkribierbare Lautfolge) reproduziert wurde. Die rezeptiven Fertigkeiten sind daher zumindest in diesem Fall nicht aus der Reproduktion der Sätze ableitbar.

Im Subtest GMP3 sollen die Kinder reale Wörter reproduzieren, die mit einem Störgeräusch abgespielt werden. Hierbei sollen Prozesse auf der akustisch-phonologischen Ebene überprüft und außerdem Aussagen über den Zugriff auf das mentale Lexikon ermöglicht werden. In der ungarischen Version konnten 90% der Kinder alle Wörter korrekt reproduzieren, es ergab sich ein Durchschnittswert von 99% auf der Wort- und 99,4% auf der Silbenebene. Die Ergebnisse lagen um durchschnittlich 19 Prozentpunkte über der für die jeweilige Altersgruppe zu erwartenden Leistung. Im GMP-d streuen die Werte wesentlich stärker als im GMP und es ergaben sich 55,6% bei Auswertung auf der Silbenebene und 40,6% bei Auswertung auf der Wortebene. Letzterer Wert ist um 20 Prozentpunkte geringer als die durchschnittlichen Leistungen gleichaltriger Kinder mit Deutsch als Muttersprache, damit ist die Differenz zu den deutschsprachigen Kinder wesentlich geringer als bei den anderen Subtests (vgl. Imre 2006). Abb. 3 zeigt die Ergebnisse der einzelnen Kinder in beiden Testversionen. Das Kind, das im GMP nicht alle Wörter korrekt reproduzieren konnte, hat auch im GMP-d den geringsten Wert erreicht.

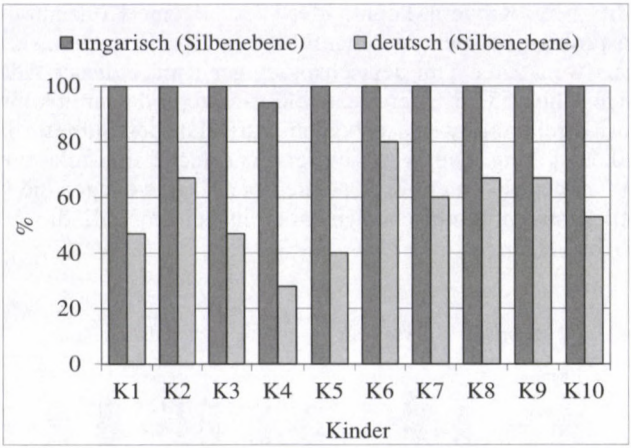


Abb. 3
Ergebnisse des GMP3

Auffällig ist, dass im GMP-d die Reproduktion mehrsilbiger Wörter mit 50% wesentlich häufiger gelingt als die einsilbiger Wörter (27,5%), was sich im Übrigen auch bei der Testung deutschsprachiger Kinder zeigte (Imre 2006). Da die Zahl einsilbiger Wörter im GMP-d doppelt so hoch ist wie im GMP, könnte dies mit zu dem schlechteren Ergebnis der Kinder in der deutschsprachigen Version beigetragen haben. Es zeigten sich Strategien wie Substitutionen und Elisionen von Konsonanten, wobei Vokale, Betonung und Silbenzahl nahezu immer beibehalten wurden. In einigen Fällen wurde das vorgegebene durch ein klanglich ähnliches, mutmaßlich bekannteres Wort ersetzt (z.B. *oben* für "Ofen"), so dass wohl auch hier die Verstehensebene eine Rolle gespielt hat.

Im Subtest GMP5 sollen die Kinder in einem erhöhten Tempo vorgespielte Sätze reproduzieren, die bzgl. Struktur und Bedeutungsgehalt ihr Sprachniveau übersteigen. Im GMP5 erreichten die Kinder durchschnittlich 58,5% bei Auswertung auf Satzebene. Die Leistungen lagen zwischen 23 und 90%, nur 30% der Kinder erreichten mindestens den für ihr Alter erwarteten Wert. Bei Auswertung auf Silbenebene ergibt sich ein Durchschnittswert von 81,3%. Ein ähnliches Ergebnis zeigte sich bei einer Untersuchung von Gósy–Horváth (2006); hier erreichten nur 12% der Kindergartenkinder ein altersgemäßes Ergebnis.

Die Leistung der Kinder im GMP-d 5 lag bei Auswertung auf der Silbenebene zwischen 5-48%, im Durchschnitt ergaben sich 29,1%. Zwei Kinder konnten 1-2 Sätze bis auf eine Silbe korrekt wiedergeben, die vollständig korrekte Wiedergabe eines Satzes gelang jedoch keinem der

Kinder. Mit zwei Kindern konnte der Test wegen Konzentrations- und Motivationsproblemen nicht durchgeführt werden. Erwartungsgemäß liegt das Ergebnis weit unter dem deutschsprachiger Kinder dieser Altersgruppe, die durchschnittlich 55,75% der Sätze fehlerfrei reproduzieren konnten (Imre 2006). Die Korrelation zwischen beiden Varianten des Subtests ist mit r = 0,28 ($p > 0,05$) gering. Die Werte in der Erstsprache sind relativ einheitlich und liegen – außer bei einem Kind – alle über 80%, wogegen sie sich in der L2 wesentlich stärker unterscheiden, aber in keinem Fall die 50%-Marke erreichen (Abb. 4).

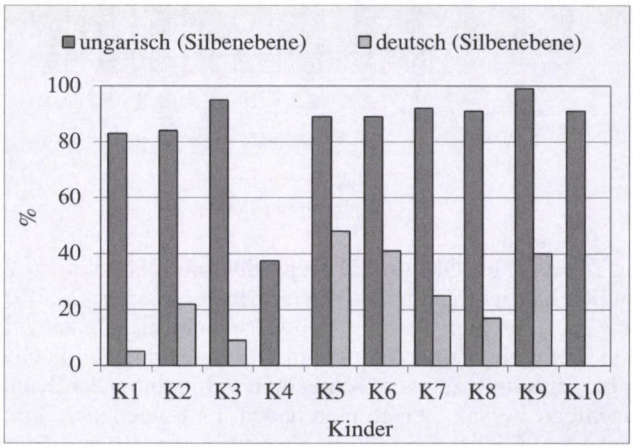


Abb. 4
Ergebnisse des GMP5

Bei der qualitativen Auswertung des GMP 5 zeigt sich, dass in etlichen Fällen Wörter durch andere, syntaktisch passende Wörter substitutiert wurden, die zumeist klanglich ähnlich waren und in der Silbenzahl übereinstimmten (z.B. *belesültek* für „felesküdtek"). Hingegen findet sich – abgesehen von auf Dyslalien beruhenden Lautsubstitutionen – nur ein Satz, in dem ein Laut ersetzt wurde.

Prozesse auf einer höheren Ebene zeigen sich erwartungsgemäß im GMP-d in weit geringerem Maße, doch finden sich auch hier Reproduktionen, in denen Wörter durch syntaktisch passende, z.T ähnlich klingende Wörter ersetzt wurden (z.B. *Trauben* für „Taube"). Wesentlich häufiger wurden Silben, Wörter oder ganze Sätze nur unartikuliert wiedergegeben, wobei Betonung und Sprachrhythmus zumeist beibehalten wurden. Zudem finden sich zahlreiche Elisionen und Substitutionen von Lauten in sonst korrekt reproduzierten Wörtern; vollständig korrekt wiederholt wurden nur 22,2% der Wörter.

Ziel des GMP10 ist die Testung der seriellen Wahrnehmung, also zu testen, ob ein Kind die Reihenfolge von Lauten in einem Logatom erkennen und reproduzieren kann. Dies gelang den Kindern in der ungarischen Version zu durchschnittlich 79,6% bei Auswertung auf der Wortebene und zu 88,9% bei Auswertung auf der Silbenebene. Die Leistung von vier Kindern entsprach der Altersnorm oder lag um zehn Prozentpunkte darüber, die anderen lagen zehn bis vierzig Prozentpunkte unter der Norm. Alle Kinder haben mindestens fünf Logatome korrekt reproduziert, einem Kind gelang dies bei allen Logatomen.

Im GMP-d 10 konnten die Kinder durchschnittlich 38% der Logatome korrekt wiederholen, die Auswertung auf der Silbenebene ergab durchschnittlich 62,5%. Damit unterscheiden sich die Ergebnisse in L1 und L2 mit gut 26 Prozentpunkten zwar immer noch deutlich, aber doch weniger als die der anderen bisher diskutierten Items. Die Streuung ist mit 18–93% korrekt wiederholter Silben wesentlich höher als beim GMP, die Kinder konnten zwischen einem und acht Logatome richtig wiedergeben. Bei den einzelnen Kindern unterscheiden sich die Ergebnisse in L1 und L2 nur in einem Fall um mehr als 40 Prozentpunkte, bei drei Kindern sind sie in beiden Sprachen fast gleich. Dies lässt sich damit erklären, dass eine Verarbeitung der Logatome auf höherer Ebene nicht möglich war. Die Differenzen zwischen den Ergebnissen in den unterschiedlichen Sprachen sind daher primär auf die mangelnde Vertrautheit der Kinder mit phonologischen, prosodischen und phonotaktischen Merkmalen der L2 zurückzuführen (vgl. Imre–Gráczi 2009) – allerdings sind auch Ermüdungserscheinungen nicht immer auszuschließen. Zwischen den Leistungen in beiden Sprachen gibt es mit $r = 0{,}73$ ($p < 0{,}01$) eine hohe Korrelation. Abb. 5 zeigt die Ergebnisse der einzelnen Kinder bei Auswertung auf der Silbenebene.

Beim GMP17 werden dem Kind verschiedene Logatompaare präsentiert, wobei die Logatome entweder identisch sind oder sich in einem Laut bzw. in der Länge eines Lautes unterscheiden. Hiermit soll getestet werden, inwieweit das Kind Laute diskriminieren kann. Die Entscheidung, ob sie gleiche oder verschiedene Logatome gehört hatten, gelang den Kindern im GMP17 durchschnittlich zu 70,2%, im GMP-d 17 zu 66,9%. Allerdings ist die Differenzierung von Lauten nach ihrer Länge in diesem Alter noch nicht erwartbar (Gósy 2006). Bleiben die entsprechenden Logatompaare unberücksichtigt, so erhält man ein wesentlich besseres Ergebnis von durchschnittlich 84,9% in der ungarischen und 71,3% in der deutschen Version. Die Unterschiede zwischen den Leistungen der einzelnen Kinder sind auch hier in der L2 bedeutend größer als in der L1 (Abb. 6). Bei sieben Kindern sind die Ergebnisse in der L2 um fünf bis fünfundzwanzig Prozentpunkte schlechter als die in der L1, zwei Kinder erreichten in der L2 ein besseres Ergebnis als in der L1. Die Leistungen in den beiden Sprachvarianten korrelieren mit $r = 0{,}69$ ($p < 0{,}05$) relativ hoch.

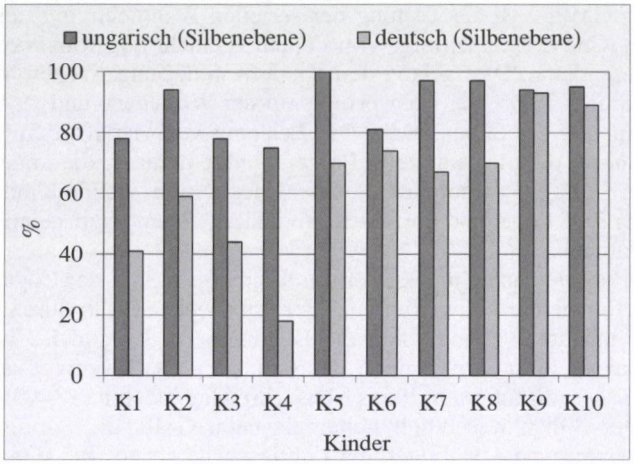


Abb. 5
Ergebnisse des GMP10

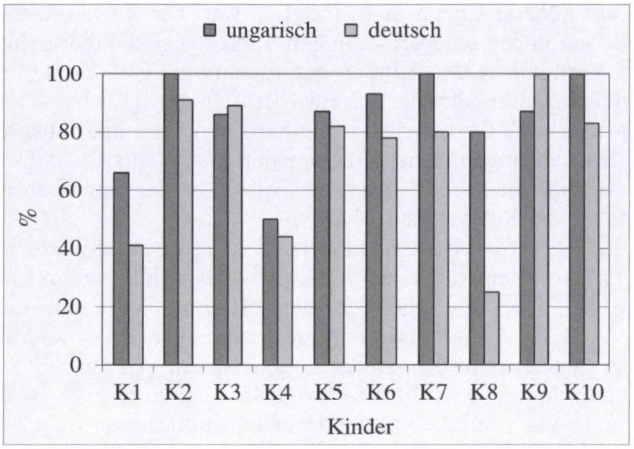


Abb. 6
Ergebnisse des GMP17 (ohne Differenzierung nach Lautlänge)

GMP16: Satzverstehen

Ziel des GMP16 ist die Testung des Satzverstehens: Mit Hilfe einer Satz-Bild-Zuordnung soll überprüft werden, inwieweit die Kinder semantische und syntaktische Merkmale verarbeiten können. Hierfür werden ihnen zehn Sätze präsentiert sowie zu jedem Satz ein Bildpaar, wobei sich die

Bildaussage innerhalb jeden Paares in einem Punkt unterscheidet. In der ungarischen Version gelang dies den Kindern zu durchschnittlich 89,1%, wobei alle Kinder mindestens 75% und zwei Kinder allen Sätzen das richtige Bild zuordnen konnten. Bei 80% der Kinder lag die Leistung über dem altersgemäßen Wert, bei 20% knapp darunter.

Im GMP-d 16 erreichten die Kinder durchschnittlich 63,4%, wobei die Werte zwischen 12% und 87% streuen. Die Ergebnisse in L1 und L2 unterscheiden sich bei einigen Kindern kaum, bei anderen differieren sie erheblich (Abb. 7). Die Leistungen in beiden Sprachvarianten korrelieren nicht miteinander ($r = -0{,}01$).

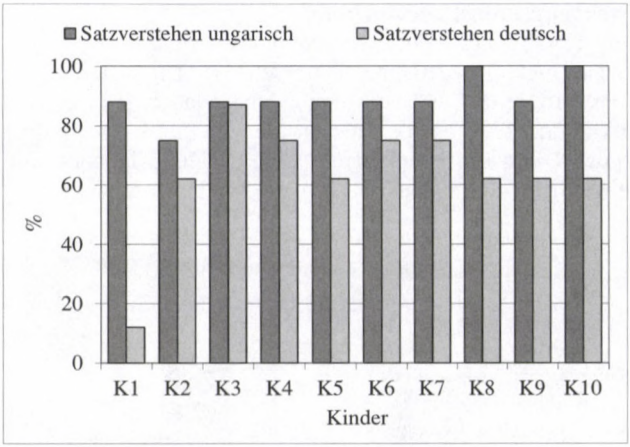


Abb. 7
Ergebnisse des GMP16

Wie die qualitative Auswertung zeigt, bereiteten die einzelnen Sätze den Kindern unterschiedlich große Probleme. So wurden im GMP fünf Sätze von allen Kindern dem richtigen Bild zugeordnet sowie je ein Satz von acht bzw. neun Kindern. Dem fünften Satz („Mielött a maci ivott, evett egy kicsit") wurde von 60% der Kinder das falsche Bild zugeordnet. In der L2 wurde zwar kein Satz von allen Kindern dem richtigen Bild zugeordnet, jedoch bereiteten vier Sätze nur je ein bis zwei Kindern Schwierigkeiten. Drei dieser Sätze enthielten eine Negation, der vierte eine Mengenangabe. Bezüglich der syntaktischen Konstruktion gibt es in diesen Sätzen keine Unterschiede zwischen der deutschen und der ungarischen Sprache, zudem reichte es jeweils aus, ein Schlüsselwort („nicht" bzw. „viele") zu verstehen, um das richtige Bild auszuwählen; während der Testung in der L2 war deutlich zu beobachten, dass die Kinder sich beim Hören dieses Schlüsselwortes für

eines der Bilder entschieden. Ob die Kinder z.B. verstanden, worauf sich der Negationspartikel im jeweiligen Satz bezog, ist aus den Antworten der Kinder nicht ableitbar. Bei falschen Antworten ließ sich nicht immer beurteilen, inwieweit die Probleme auf dem geringen rezeptiven Wortschatz der Kinder in der L2 oder auf mangelnder Kenntnis der Grammatik beruhten.

Zwei der Sätze wurden von nur je vier Kindern dem richtigen Bild zugeordnet, wobei der eine eine präpositionale Fügung enthielt, der andere die Lokalpräposition „hinter"; in diesen Sätzen entspricht die Konstruktion in der deutschen nicht der in der ungarischen Sprache. Dies trifft auch auf Satz 9 zu, der zudem die Lokalpräposition „unter" enthält, welche auch im L1-Erwerb später erworben wird als andere (Rickheit 2002), und von keinem der Kinder korrekt zugeordnet werden konnte.

Bei den Ergebnissen der Sprachwahrnehmungsitems ergeben sich erhebliche geschlechtsspezifische Unterschiede zugunsten der Mädchen. Dabei ist die Differenz zwischen Jungen und Mädchen in der L2 mit 22,2 Prozentpunkten deutlich größer als in der L1 (9,4 Prozentpunkte). Die Ergebnisse in L1 und L2 unterscheiden sich bei den Mädchen mit 33,2 Prozentpunkte weniger stark als bei den Jungen (46 Prozentpunkte).

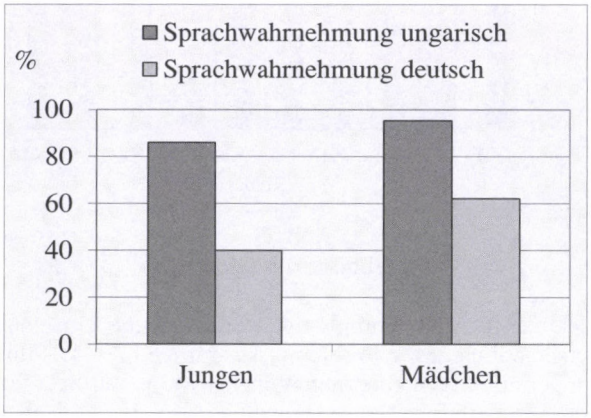


Abb. 8
Ergebnisse der Sprachwahrnehmungsitems bei Jungen und bei Mädchen

Beim Subtest zum Sprachverstehen zeigt sich in beiden Sprachen mit jeweils gut sieben Prozentpunkten ein geringerer Abstand zwischen Mädchen und Jungen. Inwieweit diese Ergebnisse auf eine höhere Kooperations- und Leistungsbereitschaft der Mädchen während der Testung oder auf bessere Sprachkompetenzen zurückzuführen ist, lässt sich hier nicht abschließend klären.

Zusammenfassung und Schlussfolgerungen

Die interindividuellen Unterschiede für die einzelnen Subtests sind – trotz gleichen Erwerbszeitraums und im Prinzip gleichen Erwerbsbedingungen für alle Kinder – in der L2 größer als in der L1. Dies kann bedeuten, dass das erreichte Niveau beim Erwerb der phonotaktischen und phonologischen Regeln der deutschen Sprache unterschiedlich ist, möglicherweise haben sich aber auch die unterschiedlichen Sprachverstehenskompetenzen auf die Ergebnisse in den Subtests zur Sprachwahrnehmung in der L2 stärker ausgewirkt. Als Gründe für die Differenzen zwischen den Leistungen einzelner Kinder kommen z.B. individuelle Faktoren wie Geschlecht, Motivation, Intelligenz und Kontakt zur deutschen Sprache außerhalb des Kindergartens sowie Ermüdung und Aufmerksamkeitsprobleme während der Testung infrage.

Auch die intraindividuellen Unterschiede zwischen den einzelnen Subtests zur Sprachwahrnehmung sind bei Testung in der Erstsprache wesentlich geringer als bei Testung in der deutschen Sprache. Betrachtet man die durchschnittlichen Ergebnisse in einzelnen Subtests, so ergeben sich die größten Unterschiede zwischen L1 und L2 bei den Subtests, in denen ganze Sätze reproduziert werden sollten (GMP6: 53,2 Prozentpunkte und GMP5: 52,2 Prozentpunkte). Die geringsten Unterschiede finden sich bei Subtests, für die statt realer Wörter Kunstwörter verwendet wurden (GMP10: 26,4 Prozentpunkte und GMP17: 13,6 Prozentpunkte). Diese Ergebnisse sind vermutlich auf Prozesse auf der Verstehensebene bei der Reproduktion realer Wörter zurückzuführen. Sie korrespondieren im Übrigen mit Ergebnissen aus einer ähnlichen Untersuchung (Imre–Gráczi 2009). Zwischen den Gesamtergebnissen der Kinder in der L1 und denen in der L2 ergibt sich bei den Sprachwahrnehmungsitems mit $r = 0{,}83$ ($p < 0{,}01$) eine wesentlich höhere Korrelation als bei dem Item zum Sprachverstehen ($r = -0{,}01$), sodass die Ausgangshypothese verifiziert werden konnte. Bei Testung in der Erstsprache ist kein Zusammenhang zwischen der Sprachwahrnehmung und dem Sprachverstehen zu erkennen ($r = 0{,}07$), hier sind bei 60% der Kinder die Ergebnisse der Sprachwahrnehmungsitems besser. Hingegen ergibt sich für die L2 eine geringe, wenn auch nicht signifikante Korrelation zwischen Sprachwahrnehmung und Sprachverstehen von $r = 0{,}22$ ($p > 0{,}05$), und nur 30% der Kinder erzielten bessere Leistungen in den Sprachwahrnehmungsitems. Letzteres Ergebnis wie auch die in den deutschsprachigen Subtests verwendeten Strategien korrespondieren mit den Ergebnissen der Testung jüngerer Kinder in der Erstsprache (Gósy 2007).

Bereits während der Testung fiel auf, dass die Kinder im deutschsprachigen Teil bei Aufgaben, in denen deutsche Wörter (auch Kunstwörter) oder Sätze reproduziert werden sollten (GMP 3, 5, 6, 10), größere Probleme hatten als bei Subtests, bei denen die Aufgabe in Form einer Ja/Nein-Entscheidung oder einer motorischen Antwort lösbar war (GMP 16, 17). Dies

spiegelt sich auch in den Testergebnissen wieder: 90% der Kinder erzielten bei Aufgaben ohne reproduktiven Anteil bessere Ergebnisse (Abb. 9). Zum Teil könnten diese Differenzen dem Umstand geschuldet sein, dass die nicht-reproduktiven Aufgaben aufgrund des Materials motivierender waren, jedoch trat dieser Effekt bei Testung in der Erstsprache nicht auf: Hier waren die Leistungen aller Kinder bei Aufgaben mit reproduktivem Anteil besser.

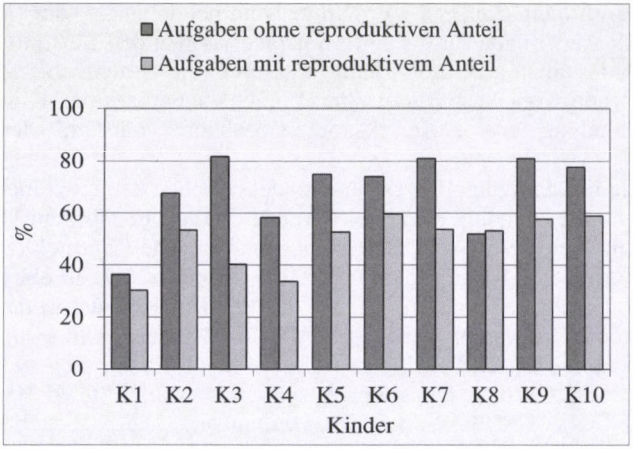


Abb. 9
Ergebnisse des GMP-d bei Aufgaben mit und ohne reproduktiven Anteil

Auch die Ergebnisse der qualitativen Auswertung zeigen, dass offenbar auch bei der Bewältigung der Subtests zur Sprachwahrnehmung Prozesse auf einer höheren Ebene eine Rolle spielten. Die tatsächlichen Zusammenhänge zwischen der Sprachrezeption in der L1 und der in der L2 konnten daher im Rahmen der vorliegenden Untersuchung nicht abschließend bestimmt werden.

Es bleibt festzustellen, dass die Leistungen in den L2-Sprachwahrnehmungsitems deutlich vom jeweiligen Aufgabentyp abhängen und bei der Reproduktion von realen Wörtern und Sätzen in der L2 auch Verstehensprozesse ablaufen. Als Folgerung für die Testung der Sprachrezeption von Kindergartenkindern in der L2 mit dem GMP ergibt sich, dass die Aufgaben den Besonderheiten des frühen L2-Erwerbs angepasst werden müssen. Insbesondere sind für Sprachwahrnehmungsitems Aufgabenstellungen zu präferieren, die unabhängig vom rezeptiven Wortschatz bewältigt werden können, also Aufgaben mit Logatomen. Zudem sollten bevorzugt Aufgaben ohne einen reproduktiven Anteil angeboten werden. Denkbar sind hier Aufgabenstellungen, bei denen das Kind motorisch Ja/Nein-Entscheidungen

mitteilt oder auf gleich- bzw. verschiedenfarbige Würfel zeigt (z.B. bei Items, die den Vergleich von Wörtern erfordern). Eine derartige Aufgabenstellung ohne reproduktiven Anteil beinhaltet auch der GMP16 zur Überprüfung des Satzverstehens, bei dem jeweils einem Satz ein Bild zugeordnet werden soll; allerdings muss hier berücksichtigt werden, dass Kinder im L2-Erwerb sehr geübt in der Anwendung der Schlüsselwortstrategie sind. Daher kann es sinnvoll sein, das Kind unter mehr als zwei Bildern auswählen zu lassen. Um feststellen zu können, inwieweit das Kind die Struktur eines Satzes durchdringt, empfiehlt es sich außerdem, vorab in einem eigenen Subtest (oder über eine Befragung der Kindergartenpädagogin) zu überprüfen, ob den Kindern der in den präsentierten Sätzen verwendete Wortschatz bekannt ist. Darüber hinaus ist bei der Durchführung zu berücksichtigen, dass eine Testung in der L2 eine höhere Konzentration erfordert und die Kinder hierfür entsprechend motiviert sein müssen.

Literatur

Apeltauer, Ernst 2007. Das Kieler Modell: Sprachliche Frühförderung von Kindern mit Migrationshintergrund. In Ahrenholz, Bernt (Hg.): *Deutsch als Zweitsprache – Voraussetzungen und Konzepte für die Förderung von Kindern und Jugendlichen mit Migrationshintergrund*. Fillibach, Freiburg, 91–115.

Bereznai, Anja 2012. *Sprachrezeption bei Kindern in zweisprachigen Kindergärten*. Unveröffentlichte Masterarbeit. ELTE, Budapest.

Bredel, Ursula 2009. Sprachstandsmessung – eine verlassene Landschaft. In Ehlich, Konrad et al.: Anforderungen an Verfahren der regelmäßigen Sprachstandsfestellung als Grundlage für die frühe und individuelle Förderung von Kindern mit und ohne Migrationshintergrund. Bundesministerium für Bildung und Forschung, Berlin, 77–119. http://www.bmbf.de/pub/bildungsreform_band_elf.pdf (Stand: 25. 10.2012).

Gósy Mária 2005. Pszicholingvisztika. Osiris Kiadó, Budapest.

Gósy Mária 1996/2006. *GMP-Diagnosztika. A beszédészlelés és a beszédmegértés folyamatának vizsgálat, fejlesztési javaslatok*. Nikol Kht., Budapest.

Gósy Mária 2007. A beszédészlelés és a beszédmegértés összefüggései kisgyermekkorban. *Magyar Nyelvőr* 131. 129–143.

Gósy Mária – Horváth Viktória 2006. Beszédfeldolgozási folyamatok összefüggései gyermekkorban. *Magyar Nyelvőr* 130. 470–481.

Höhle, Barbara 2010a. Sprachwahrnehmung. In Höhle, Barbara (Hg.): *Psycholinguistik*. Akademie Verlag, Berlin. 39–50.

Höhle, Barbara 2010b. Erstspracherwerb: Wie kommt das Kind zur Sprache? In Höhle, Barbara (Hg.): *Psycholinguistik*. Akademie Verlag, Berlin. 125–137.

Imre Angéla 2006. *Beszédészlelési és beszédmegértési vizsgálatok német anyanyelvű gyermekek körében*. Dissertation. ELTE, Budapest.

Imre Angéla – Gráczi Tekla Etelka 2009. Egynyelvű, magyar anyanyelvű iskolások beszédpercepciós folyamata idegen nyelven. In Lengyel Zsolt – Navracsics Judit (szerk.): *Tanulmányok a mentális lexikonról. Nyelvelsajátítás – beszédprodukció – beszédpercepció*. Tinta Könyvkiadó, Budapest. 326–335.

Kersten, Kristin 2002. *Äquivalenzklassifizierungen im Zweitspracherwerb in bilingualen Kindertagesstätten.* Magisterarbeit. http://www.kristin-Kersten.de/media/Kersten$202002$20Aequivalenzklassifizierungen.pdf (Stand: 11.03.12).

Müller Mónika 2004. Kommunikációs és nyelvtanulási stratégiák óvodás korban. In *Habakukk alapítványi óvoda: Módszertani füzetek 1.* Budapest. 21–31.

Pienemann, Manfred 2002. Unanalysierte Einheiten und Sprachverarbeitung im Zweitspracherwerb. *Zeitschrift für angewandte Linguistik* 2002/37. 3–26.

Rickheit, Gert et al. 2002. *Psycholinguistik.* Stauffenburg Verlag, Tübingen.

Rothweiler, Monika 2007. Bilingualer Spracherwerb und Zweitspracherwerb. In Steinbach, Markus Albert, Ruth – Girnth, Heiko – Hohenberger, Annette – Kümmerling-Meibauer, Bettina, Meibauer, Jörg – Rothweiler, Monika – Schwarz-Friesel, Monika: *Schnittstellen der germanistischen Linguistik.* Metzler, Stuttgart–Weimar. 103–130.

Rothweiler, Monika – Rugberg, Tobias 2011. *Der Erwerb des Deutschen bei Kindern mit nichtdeutscher Erstsprache. Sprachliche und außersprachliche Einflussfaktoren.* Deutsches Jugendinstitut e.V. http://www.weiterbildungsinitiative.de (Stand: 15.02.12).

Schulz, Petra 2007. Sprache verstehen – Herausforderungen für Kinder mit Deutsch als Zweitsprache. Plenarvortrag, Fachtagung Sprache, Mainz, auf Einladung des Ministeriums für Bildung,Wissenschaft, Jugend und Kultur Rheinland-Pfalz. http://kita.bildung-rp.de/fileadmin/downloads/Prof._Schulz_Mainz_Sprachfachtagung_2007_Vortrag.pdf (Stand: 17.10.12)

Steinbrink, Claudia – Spitzer, Manfred 2006. Sprache lernen – Lernprozesse in der kindlichen Sprachentwicklung. *Frühes Deutsch* 07. 34–37.

Wenzel, Veronika 2006. Zweisprachigkeit im Kindergarten Pusteblume in Gronau. Evaluation. http://www.pusteblume-gronau.de/index.php?option=com_jdownloads&Itemid=131&view=finish&cid=2&catid=1&m=0&lang=de (Stand: 20.12.11)

A BESZÉDFELDOLGOZÁS JELLEMZŐI MAGYAR DOMINÁNS KÉTNYELVŰ GYERMEKEKNÉL

Bartha Krisztina

Bevezetés

A kétnyelvűség kérdése sokakat érint Romániában a román és a magyar nemzetiségű csoportok együttélése kapcsán. A kisebbségben élő magyar nemzetiségűek rá vannak utalva arra, hogy a többségi nyelvet is minél jobban elsajátítsák. A magyar családból származó gyermekek, bár dominánsan magyar nyelvűek, a román nyelvvel két-három éves koruktól ismerkednek. Minél jellemzőbb a többségi román közösség egy-egy településen, annál korábban ismerkednek meg a magyar gyermekek a román nyelvvel a játszótéren, a szomszédokat hallva vagy más közösségi színtereken. Óvodába lépésüktől fogva kötelező módon tanulják is az államnyelvet. A kétnyelvű környezet maga után vonja azt, hogy a gyermek már élete korai szakaszában, kisgyermekkorától kezdve, az anyanyelvén kívül legalább még egy nyelvvel ismerkedjen, s ez befolyásolja az anyanyelvi mechanizmusait is (Imre 2005).

A kétnyelvűség fogalma a szó hétköznapi és tudományos értelmében is ismert, meghatározása mégsem egységes. Léteznek tágabb és szűkebb meghatározások, de előfordul az is, hogy az egyértelműség okán skálák, jellemzők mentén vizsgáljuk inkább a kétnyelvűséget (Bartha 1999, Edwards 2006). A legtöbb definícióban azt tekintik kiindulópontnak, hogy mennyire kell birtokolnia valakinek mindkét nyelvet ahhoz, hogy kétnyelvűnek tekintsük (pl. Bloomfield 1933, Haugen 1953, Diebold 1961, Grosjean 1992). Tág értelmezés fűződik Diebold (1961) nevéhez, aki szerint az a személy kétnyelvű, aki képes az írott nyelv passzív megértésére. Ezzel ellentétben Bloomfield (1933) meghatározásában az számít kétnyelvűnek, aki mindkét nyelvet anyanyelvi szinten beszéli és folyamatosan használja. Mások e két véglet között fogalmazzák definícióikat. A jelenség meghatározásában felmerülő nehézségek oka legfőképpen az, hogy sokszínűsége és egyéni változatai megnehezítik a pontos körvonalazását.

A kétnyelvűség kifejezés egyénekre és közösségekre egyaránt vonatkoztatható fogalom. A magyar szerzők is különböző nézőpontból vizsgálják: a kisebbségi környezetben élők kétnyelvűségével Lanstyák István és Szabómihály Gizella (1997), Göncz (2004), Karmacsi (2007) kutatásai foglalkoznak, a kétnyelvűek mentális lexikonával Navracsics (2010), a beszédészlelés és a beszédértés kérdéskörével többek között Vančoné Kremmer (2002), Imre

(2005), a kétnyelvűség egyéni és közösségi kérdéseivel egyaránt Bartha (1999) foglalkozik. Bartha Csilla Grosjean nyomán úgy fogalmaz, hogy az a személy tekinthető kétnyelvűnek, aki „a mindennapi érintkezései során két vagy több nyelvet kommunikatív, szociokulturális szükségleteinek megfelelően (szóban és/vagy írásban, illetőleg jelelt formában) rendszeresen használ" (Bartha 1999: 40).

Az egyéni kétnyelvűvé válás számos módon következhet be. Vannak olyan helyzetek, amelyek egy nagyobb közösséghez való tartozás következményei, mások viszont kizárólag a családi háttérből adódnak. Ezen lehetőségek közül Hoffmann rendszerezte az alapeseteket (Hoffmann 1991): bevándorlás, migráció, a kétnyelvű családban felnövő gyermek, szoros érintkezés más nyelvi csoportokkal, az iskola szerepe.

A romániai magyar kisebbség esetében ezek közül három ok jelenlétéről beszélhetünk. A szoros érintkezés más nyelvi csoportokkal a kétnyelvűvé válás azon típusa, amikor két vagy több nemzet él együtt, ezért az egyes nemzetek érdeke, hogy a mellettük élő nemzet nyelvét elsajátítsák. Romániában a magyar kisebbség esetében az érvényesülésnek elengedhetetlen feltétele a román nyelv ismerete.

A másik mód az iskolai nyelvtanulás. Az iskola szerepe a kétnyelvűvé válásban Romániában már akkor elkezdődik, amikor a kisgyermek óvodába kezd járni. Ettől a pillanattól kezdve kötelező a magyar gyermekek számára a román nyelvű oktatás. A kisgyermekkorban elkezdett második nyelv elsajátítása mentális erőfeszítést (munkamemória, figyelem, gondolkodási műveletek) igényel a gyermektől, ugyanakkor alakítja az anyanyelvi működéseket, a beszédészlelést és a beszédmegértést is (Imre 2005). A kérdés az, hogy az erőfeszítés, amit a gyermek a második nyelv elsajátításának érdekében kifejt, milyen hatással van az egyes nyelveken történő beszédészlelésére és beszédmegértésére.

A vegyes házasság is vezethet kétnyelvűséghez. Mivel vegyes házasságban született gyermekeket a jelen kutatás nem vizsgál, ezért nem térünk ki a jelenség elméleti vonatkozásaira sem.

A beszédészlelés és a beszédmegértés két egymással szoros összefüggésben lévő folyamat. A beszédészlelés a beszédhangok felismerése és differenciálása, a hangkapcsolatok és hangsorok pontos felismerése és azonosítása. A beszédmegértés szavak, szókapcsolatok, mondatok és ezek összességének a helyes felismerése és értelmezése, vagyis a bennük foglalt tartalmak feldolgozása (Gósy 2000).

Ahhoz, hogy a beszédmegértéshez fogható komplex folyamat végbemenjen, nem elég előfeltétel a jó hallás megléte. Egymásra épülő részfolyamatokon keresztül jut el a gyermek a megértésig. A hallást a beszédészlelés szintje követi, majd a megértés és az asszociációs szint (Gósy 2000).

A beszédmegértés a jelrendszer értelmezését jelenti. A jó beszédmegértésben szükség van a pontos szómegértésre. Ebben játszik jelentős szerepet a

mentális lexikon aktiválása. A mentális lexikon három részre tagolódik: aktív, passzív és aktivált szókincsre. A beszédmegértés szempontjából fontos az elemek száma és a hozzáférés minősége, pontossága, gyorsasága (Gósy 2005). A mondatmegértés is szükséges feltétele a jó beszédmegértésnek. Ezen a szinten szükséges a szavak szószerkezetekké kapcsolása, a szerkezetekhez mondattani funkciók rendelése, a szerkezetek jelentésének értelmezése. A magyar nyelv gazdag a toldalékok szempontjából, így a mondat feldolgozása során jelentős szerep jut az alaktani elemzésnek (Gósy 2005). A mondatértés akkor is bizonytalanná válhat, ha a gyermek ismeri ugyan a szótőt, de nem ismeri vagy bizonytalan a toldalékok szerepének azonosításában. Ez megnehezítheti a tanulási folyamatot, az elhangzó mondatok pontos megértését. A szövegmegértés szintjén lényeges folyamat a részletek és azok összefüggésének megértése, a világról való ismeret mozgósítása, a közlési helyzet felismerése, az új információ behelyezése a már meglévő tágabb ismeretanyagba (Büky – Egyed – Pléh 1984, Pléh 1998, Gósy 2000). A beszélt szöveg megértése komplex folyamat, melynek megfelelő működésében fontos szerepe van az emlékezetnek is, emellett az egyénnek egyszerre kell dekódolnia a részletek jelentéseit és az ezek közötti összefüggéseket. Itt tehát jelentéstani, szerkezeti összefüggéseket kell teremteni ahhoz, hogy megértésről beszéljünk. Mindezek a folyamatok és az egyéni tudás is, az előzetes információk egészébe helyezkednek bele, s így nyernek igazi értelmet (Gósy 2000). Ha e folyamatok bármely szintjén zavar támad, problémák jelenhetnek meg a beszédészlelésben és/vagy a beszédmegértésben, és ez a tanulási folyamatok működtetésére is kihat.

A beszédészlelés és a beszédmegértés folyamatainak alakulását egy- és kétnyelvű gyermekeknél egyaránt vizsgálták már (Gósy 2000, Macher 2002, Vančoné Kremmer 2002, Imre 2005; 2007, Gósy–Horváth 2006, Markó 2007, Menyhárt 2007, Bakk-Miklósi 2009). Az eredmények azt mutatják, hogy mind az egynyelvű magyar gyermekek, mind a kétnyelvűek bizonyos beszédfolyamaira az elmaradás jellemző. Mind az egynyelvűek, mind a kétnyelvűek jobban teljesítenek a mondatértésben, de ezek az eredmények elmaradnak az elvárttól. A két csoport teljesítménye között nincs kimutatható különbség, de más-más szerkezetek okoznak problémát az egy- illetve a kétnyelvűeknek. A szövegértésben az elmaradások jelentősebbek az egyes csoportokon belül, a kétnyelvűek szövegértését pedig az iskola tannyelve befolyásolja jelentősen.

Markó Alexandra (2007) tipikusan fejlődő 6–9 éves egynyelvű gyermekeket vizsgálva kimutatta, hogy mind a mondatértési, mind a szövegértési képesség elmarad a korosztályukban elvárt eredményektől. A két részfolyamat közül a mondatértés bizonyult sikeresebbnek. Ezek az eredmények megegyeznek más egynyelvű gyermekeket felmérő vizsgálatok eredményeivel (Gósy–Horváth 2006, Imre 2007). A vizsgálatok azt is kimutatták, hogy a mondat- és a szövegértés eredményei csak kis mértékben függnek össze, ami

nemcsak azt bizonyítja, hogy a szövegértésben a gyermekek kevéssé támaszkodnak a mondatértési képességeikre, hanem arra is utalhat, hogy a megfelelő szintű szövegértésben számos más részfolyamatnak is jól kell működnie.

A magyar–szlovák felmérések szerint a kétnyelvű gyermekek mondatértése jó. Nincs szignifikáns elmaradás az egynyelvű gyermekekhez képest. Ugyanakkor vannak olyan nyelvi szerkezetek, melyeknek megértése problémásabb a kétnyelvűek számára, de vannak olyanok is, amelyek egyaránt nehézséget jelentenek az egy- és kétnyelvűeknek, ilyen például a vonatkozó kötőszóval kifejezett időviszony. A szövegértésre nagyobb elmaradás jellemző, de itt sincsenek látványos eltérések a magyar egynyelvű és a magyar-szlovák kétnyelvű gyermekek között. Az azonban megállapítható, hogy a magyar tannyelvű iskolába járó gyermekek szövegértése szignifikánsan jobb, mint a szlovák tannyelvű osztályba járóké (Vančoné Kremmer 2002). Romániában a magyar domináns kétnyelvű gyermekek beszédfolyamatait Bakk-Miklósi Kinga vizsgálta először, s azóta nem születtek újabb kutatások ebben a témában. A magyar-román kétnyelvű gyermekek felméréséből szintén az derült ki, hogy a mondat- és szövegértés szintjein nincsenek szignifikáns eltérések a magyar egynyelvű gyermekekhez viszonyítva. Itt az elvárthoz képest jelentősebb elmaradással találkozunk mindkét részképesség esetében (Bakk-Miklósi 2009).

A szóbeli mondat- és szövegértési képességek vizsgálatát Romániában több okból találjuk hatványozottan fontosnak. Egyrészt, mert kevés vizsgálatot végeztek eddig ebben a témakörben (Bakk-Miklósi 2009), másrészt, mert szükséges tudnunk azt, hogy milyen fejlett beszédmegértési folyamatokkal rendelkeznek az itt élő gyermekek, mennyire biztos az anyanyelvi tudásuk, mennyire tudnak sikeresen részt venni a tanulási folyamatokban.

Felmerül a kérdés, hogy vajon az egynyelvűek számára kidolgozott tesztek alkalmazhatóak-e kétnyelvűek esetében is. Azok a magyar domináns kétnyelvű gyermekek, akik Romániában magyar tannyelvű iskolába járnak az iskolában magyarul (kivéve a román nyelv és irodalom tantárgyat) tanulnak. Nincs speciálisan kétnyelvűeknek kifejlesztett tananyag, vagyis amennyiben a gyermek magyar tannyelvű osztályba jár, magyar egynyelvűnek tekintik, amennyiben a szülők román tannyelvű osztályba íratják, román egynyelvűként kezelik. Éppen ezért a magyar domináns, magyar iskolába járó gyermekek teljesítményét össze lehet vetni az egynyelvűek teljesítményével, hiszen az iskolában sem tesznek különbséget közöttük. Az életkornak megfelelő beszédmegértési folyamatok tehát elengedhetetlenek a magyar nyelven történő tanulás és az írott anyanyelv elsajátítása során.

Ugyanakkor az is kérdésként merülhet fel, hogy a román nyelv hatása mennyire jelentős a beszédfolyamatok működésében. Az eddigi felméréshez képest több új probléma is előtérbe kerülhet a kétnyelvű gyermekek beszédfeldolgozási folyamatainak vizsgálata során. Befolyásolja-e a gyermekek magyar nyelvű beszédértését, ha a település, ahol élnek túlnyomó többségében román? Van-e különbség a beszédfolyamatok tekintetében a magyar-

lakta és a románlakta településeken élő gyermekek között? Vannak-e eltérések a nagyvárosi és a kisvárosi gyermekek beszédértési folyamatai között?

A kutatás célja a kétnyelvű környezetben élő, de magyar domináns II. osztályosok szóbeli mondat- és szövegértési képességeinek megismerése. Célunk az is, hogy az eredmények nyomán rámutassunk arra, hogy a beszédfolyamatok vizsgálata fontos a sikeres tanulásban.

A kutatás hipotézisei a következők:

1. A második osztályos kétnyelvű gyermekek mondatértési és szövegértési képessége elmarad a korosztályukban elvártaktól.
2. A kisvárosi gyermekek eredményei rosszabbak a nagyváradi gyermekek eredményeinél.
3. A nagyrészt románok által lakott kisvárosban a magyar nyelvű beszédmegértésben gyengébben teljesítenek a diákok.

Adatközlők, anyag, módszer

A kutatásban 48 Bihar megyei tanulót mértünk fel, három településen. A csoport átlagéletkora 7,8 év, mindannyian második osztályos tanulók. A lakhely szerinti eloszlást tekintve 16 nagyváradi, 16 székelyhídi és 16 élesdi gyermek vett részt a felmérésben. A megyeközpont (Nagyvárad) mellett két olyan kisváros, Székelyhíd és Élesd vett részt a vizsgálatban, melyek egyaránt körülbelül 40 kilométerre helyezkednek el Nagyváradtól. Székelyhíd a román–magyar határ mellett helyezkedik el, Élesd Bihar megye központjában található.

A 2011-es népszámlálási előrejelző adatok szerint Nagyvárad lakosságának 24,51%-a magyar anyanyelvű. A kutatás másik helyszíne Székelyhíd, ahol a magyarság 77,49%-ban van jelen, Élesden a magyarság aránya már jóval alacsonyabb, mindössze 16,04%. Nagyváradon és Székelyhídon a magyarság aránya tehát 20% fölötti, vagyis átlépi a kisebbségi nyelvi jogok gyakorlásához szükséges alsó határt, így ezeken a településeken több lehetőség nyílik az anyanyelv gyakorlására. Ez azt is jelenti, hogy az említett településeken élő tanulók számára az anyanyelv használata jellemzőbb, mint Élesden, ahol a kisebbségi nyelvhasználat korlátozottabb.

19 fiú és 27 lány vett részt a felmérésben. A gyerekek mindannyian éphallók, nincs beszédhibájuk, és magyar tannyelvű iskolába járnak.

A tanulók mindannyian magyar családból származnak. Azt, hogy dominánsan magyar kétnyelvűek-e, három kérdés segítségével mértük fel: *Milyen nyelven beszélsz az édesanyáddal? Milyen nyelven beszélsz az édesapáddal? Milyen nyelven beszélsz a legszívesebben?* A tanulók mindhárom kérdés esetében minden településen egyöntetűen (100%) a magyar nyelvet választották. A felmért gyermekeket mégis kétnyelvűeknek tekinthetjük, hiszen intézményes formában 3 éves koruktól tanulnak románul, és a mindennapokban, de természetesen eltérő gyakorisággal (vásárlás, orvosi rendelő, játszótér stb.) használják azt.

A felmérést a GMP-diagnosztika két tesztjével végeztük el (Gósy 1995). A GMP-diagnosztika egy standardizált eljárás, két tesztje, a GMP 16 és a GMP 12 a beszédértés részfolyamatait vizsgálja. A GMP 16 a mondatértés tesztelésére szolgál. Ez a teszt az adott életkorban a jelentéstani sajátosságok és a szintaktikai/grammatikai struktúrák megfelelő működéséről informál. A gyermeknek a majdnem egyforma rajzpárok közötti egyetlen eltérés felismerésére és az elhangzott mondatnak megfelelő rajz rámutatására kell képesnek lennie. Például az egyik képen az látható, hogy a nyuszi kockás nadrágban van, míg a másikon csíkosban. A gyermeknek arra a képre kell rámutatnia, amelyik megfelel a *Nem a nyuszi vette föl a kockás nadrágot* mondatnak. 7 éves korra már hibátlan teljesítményt várunk, azaz mind a tíz alkalommal a mondatnak megfelelő rajzra kell rámutatnia a gyermeknek. Így csak a 100%-os teljesítményt tekintjük jónak, az összes többi eredményt gyengének nevezzük.

A GMP 12 célja a szövegértés feldolgozásának, a szemantikai, szintaktikai, ok-okozati viszonyok felismerésének, az asszociációs szint működésének vizsgálata. Ezt egy rövid mese meghallgatásával és az ehhez kapcsolódó tíz ellenőrző kérdéssel vizsgáljuk (például: *Miért szaladt ki a nyuszika a házikójából?* vagy *Hogyan változott meg az időjárás az erdőben?*). 8 éves kortól az összes kérdés helyes megválaszolását várjuk el.

A romániai Bihar megyében tudomásunk szerint nem készültek beszédmegértési vizsgálatok magyar nyelvű gyermekeknél, sem a GMP-diagnosztika tesztjeit nem használták felmérésekben.

Az adatok statisztikai elemzése az SPSS 17.0 szoftverrel készült, egytényezős variancianalízist (ANOVA), független mintás *t*-próbát, korrelációszámítást és kereszttábla elemzést végeztünk.

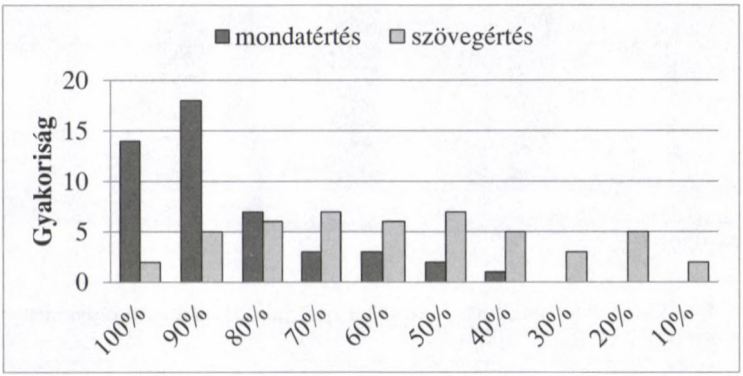
Eredmények

Az eredmények az elvárásoknak megfelelően alátámasztják, hogy a mondatértésben a tanulók az elvárt szint alatt teljesítettek. Az összes tanuló átlageredménye a mondatértési teszten 85,6%. A korosztályukban elvárt 100%-os teljesítményt a gyermekek közül mindössze 14 tanulónak sikerült teljesítenie. 90%-os eredményt 18 tanuló produkált. 80%-ot 7 diák, 70 és 60%-ot 3-3 diák ért el. Két gyermek mindössze 50%-os teljesítményt nyújtott, míg egy diák 40%-os teljesítményt ért el a mondatértési teszten.

A szövegértési tesztben a gyermekek gyengébben teljesítettek, mint a mondatértésben, az átlageredmény itt 57%. Mivel ennek a tesztnek a teljesítésében 8 éves korra a gyermekektől 100%-os eredményt várunk el, a felmérésben a 90%-os és a 100%-os eredményt fogadtuk el jónak, az összes többit gyenge eredménynek tekintettük. A tanulók közül mindössze ketten érték el a 100%-ot. 90%-ot is csak 5 gyermek teljesített. 80%-os eredményt 6-an, 70%-ot 7-en, 60% ismét 6-an értek el. 50%-os teljesítménye 7 gyermeknek volt,

40-ot 5-en teljesítettek. Volt 3 olyan gyermek is, aki csak a kérdések 30%-ára tudott helyesen válaszolni. 20%-ot 5-en és 10%-ot 2-en értek el.

Ahogyan azt a számszerű eredmények bizonyítják, a tanulók mondatértési képessége fejlettebb ebben a korcsoportban, mint a szövegértésük (l. 1. ábra). A mondatértési eredmények szignifikánsan magasabbak, mint a szövegértésben nyújtott teljesítmény [$t(47) = 8{,}31$, $p < 0{,}01$]. A Pearson-féle korrelációszámítás alapján szignifikáns, közepes erősségű korreláció mutatkozik a két tesztben elért eredmények között ($r = 0{,}355$; $p < 0{,}05$).

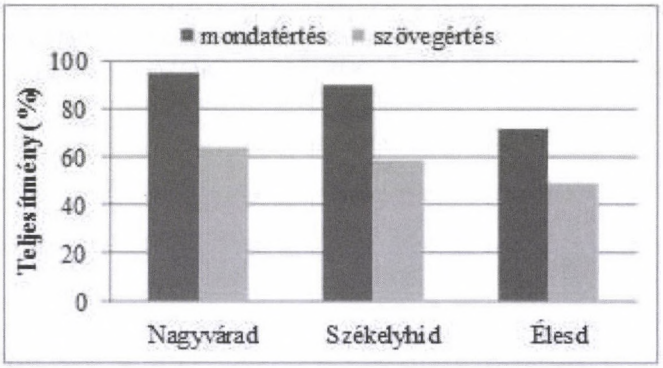


1. ábra
A tanulók mondat- és szövegértési teljesítményének eloszlása

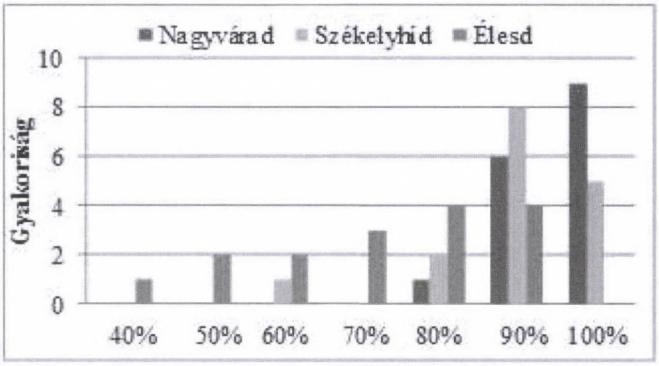
A második és a harmadik hipotézisben azt feltételeztük, hogy a lakhely szerint kimutathatóak lesznek bizonyos eltérések a diákok eredményei között, melyek összefüggésbe hozhatók azzal is, hogy mennyire van jelen a mindennapjaikban a román nyelv. A három, lakhely alapján kialakított csoport közül a nagyváradiak eredménye a legkiemelkedőbb (l. 2. ábra). A nagyváradi tanulók átlagosan 95%-ot értek el a mondatértési teszten, szövegértésük 63,7%-os. A székelyhídi gyermekek mondatértési teszten elért átlagteljesítménye 90%, míg a szövegértési teljesítményük 58,7%. Az élesdi gyermekek mindkét tesztben gyengébben teljesítettek az előző két csoporthoz képest, itt a mondatértés átlagos értéke 71,8%, a szövegértésé viszont csupán 48,7%-os. Az ANOVA teszt szignifikáns különbséget mutatott településtípus szerint a mondatértési tesztben elért eredmények között, $F(2, 47) = 18{,}27$, $p < 0{,}01$, de nem mutatható ki szignifikáns összefüggés a szövegértési teszt esetében.

A keresztáblás elemzésből kiderült, hogy a nagyváradi csoport a mondatértési feladatban 80–100% között teljesített. 9 gyermek érte el az elvárt 100%-os teljesítményt, 6-an csupán egy hibát vétettek, így eredményük 90%-os, és 1 tanuló 80%-ot teljesített. A székelyhídi gyermekek eredményei 60%-nál kezdődnek: 100%-ot csak 5 gyermek teljesített, 90%-ot 8-an, 80%-ot 2-en és

60%-ot egy gyermek ért el. Az élesdi tanulók közül 100%-os eredményt senki nem ért el, a 90%-os teljesítményt is csak a csoportból 4-en értek el. Szintén 4-en teljesítettek 80%-ot, 3-an 70%-ot. 60 és 50%-ot 2-2 tanulónak sikerült elérnie. A leggyengébb, 40%-os eredmény 1 tanulóra jellemző (3. ábra).



2. ábra
A mondat- és szövegértési eredmények lakhely szerinti eloszlása

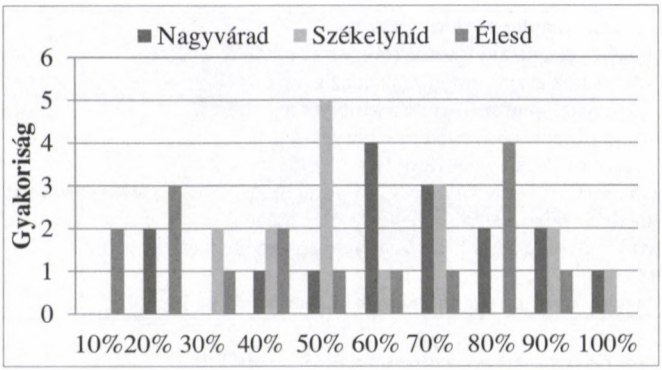


3. ábra
A mondatértés eredményei lakhely szerint

A szövegértési teszt eredményei nagyobb szórást mutatnak, hiszen itt a leggyengébb teljesítmény 10%-os volt, a legjobb pedig 100%-os. A nagyváradi és a székelyhídi tanulók nagyon alacsony arányban (1-1 gyermek) válaszoltak 100%-ban helyesen a mesével kapcsolatos ellenőrző kérdésekre, Élesden nem volt 100%-os teljesítmény. A jónak mondható 90%-os eredményt Nagyváradon és Székelyhídon 2-2 gyermek, míg Élesden 1 gyermek

érte el. 80%-ot értek el Nagyváradon 2-en és Élesden 4-en. 70%-ot teljesítettek 3-an Nagyváradon és ugyanennyien Székelyhídon, Élesden 1 gyermek. 60%-ot legtöbben, 6-an Nagyváradon értek el, Székelyhídon és Élesden 1-1 gyermek. 50%-ot teljesített 1 nagyváradi tanuló, 5-en Székelyhídról és szintén 1 élesdi gyermek. 40%-ot Nagyváradon 1 gyermek ért el, 2-2-en teljesítettek ugyanígy Székelyhídon és Élesden. 30%-ban helyesen válaszolt 2 székelyhídi tanuló és 1 élesdi. Voltak 20%-os eredmények is: 2 gyermek Nagyváradról és 3 gyermek Élesdről érte el ezt az eredményt, míg 10%-os helyes válaszadás csak Élesden volt 2 gyermek esetében.

A kereszttáblás eredmények mutatják, hogy az élesdi tanulóknál nagyobb a gyengén teljesítő tanulók száma és a jól (90–100%-ban) teljesítők aránya alacsonyabb.



4. ábra
A szövegértés eredményei lakhely szerint

Mindkét tesztben jellegzetesek voltak a gyermekek által elkövetett hibák. A mondatértési tesztben, amint azt már említettük, a kép és a mondat közötti megfelelés felismerését várjuk el. Ezzel ellenőrizni lehet, hogy a gyermek képes-e bizonyos szófaji, szintaktikai és morfológiai szerkezetek azonosítására és megértésére.

A legtöbben az ötödik mondat azonosításában tévedtek: *Mielőtt a maci ivott, evett egy kicsit.* Ebben az időviszony szintaktikai kifejezésének megértését ellenőrizzük. A teljes csoport 31,25%-a adott helytelen választ (1. táblázat). Szintén sokan (27,08%) tévesztettek a negyedik mondatban, ahol a részeshatározós szerkezetek felismerése volt a feladat. Ehhez hasonló eredmény született a hetedik mondat esetében, ahol egy tagadó szerkezetet kellett felismerni. A tanulók 25%-a hibázott. Közepesen nehéznek bizonyult az első számú mondat. Ebben a módosítószó jelentésének azonosítását vizsgáljuk. A gyermekek 10,41%-a adott helytelen választ. A teszt hatodik mondata fo-

lyamatos melléknévi igeneves szerkezetet tartalmaz, ezt a tanulók 6,25%-a tévesztette, a nyolcadik mondatot, melyben célhatározói alárendeléssel találkozunk csupán 4,16%-uk hibázta el. Biztonságosan ismerték fel a tanulók a második, a harmadik és a tizedik mondat jelentését. A második mondat határozatlan névmást tartalmazott, a harmadik a névutóhasználatot vizsgálja a *mögül* szerkezet segítségével, míg a teszt tizedik mondata mindkét tagmondatban múlt idejű feltételes mondatot tartalmaz. Egy-egy hibás válasz született mindhárom esetben, vagyis a csoport 2,08%-a teljesített tévesen. Legjobban a kilencedik mondat, a feltételes módú alárendelő szerkezet, sikerült, itt ugyanis senki nem hibázott.

1. táblázat: Az egyes mondatok téves azonosítása a teljes csoportban

	Mondatok	Hibák száma
1.	Az egérke majdnem eléri a sajtot.	5
2.	A maci és a nyuszi fára másztak, és az egyikük leesett.	1
3.	A macska az asztal mögül húzza az egeret.	1
4.	A kislánynak oda kell adnia a könyvet a kisfiúnak.	13
5.	Mielőtt a maci ivott, evett egy kicsit.	15
6.	Az asztalról leeső gyertyáról beszélnek.	3
7.	Nem a nyuszi vette föl a kockás nadrágot.	12
8.	A medve szalad, nehogy megcsípjék a méhek.	2
9.	Mivel nagyon esett a hó, a kislány mégsem ment el szánkózni.	0
10.	A kislány megette volna a tortát, ha elérte volna a tálat.	1

A szövegértési tesztben a szemantikai, szintaktikai, ok-okozati viszonyok felismerését, az asszociációs szint működését mérjük, de ahhoz, hogy helyes válaszok szülessenek, szükség van a munkamemória jó működésére, az összpontosításra, az előző ismeretek mozgósítására is.

Négy kérdés a csoport több mint 50%-ának okozott nehézséget. A legproblematikusabb kérdés a negyedik volt, 85,41%-uk hibázott. A tanulók vagy nem a mesében szereplő állatokat sorolták fel, vagy kihagytak, illetve más állatokkal egészítették ki a felsorolást. A gyermekek 62,50%-ának nehézséget okozott felidézni a harmadik kérdéshez kapcsolódó állatot. Többségük a mesében szereplő állatok közül neveztek meg egyet, de nem a megfelelőt. Az ok-okozati összefüggéseket ellenőrző ötödik kérdésben 54,16%-uk tévesztett. Az elhangzó válaszok között sok volt, amelyet a gyermekek inkább saját világismeretükből aktiváltak, kevésbé a mese cselekményéből (pl. *Mert kinn szaladgálhatnak. Mert hosszú a szőrük. Tudnak szerezni eledelt.*) A nyolcadik kérdésre 52,08%-uk adott hibás választ, annak ellenére, hogy a helyes válasz explicit módon jelen van a mesében. Közel egyforma hibaarány jelenik meg a tizedik (35,41%) és a második (33,33%) kérdés esetében. A tizedik kérdésre a helyes választ a tanulóknak ki kellett következtetniük, hiszen az nem volt jelen

konkrétan a mesében. Ennél a kérdésnél többen nem válaszoltak, aki mégis adott választ, szintén inkább saját élettapasztalatára, ismereteire hagyatkozott (pl. *Gyorsan tud szaladni. Kicsi, és befér az odújába. Tud ugrálni. Mert medvének nem jó lenni.*). Meglepő azonban a második kérdésre adott válaszok nagy arányú tévesztése, hiszen ebben az esetben a szövegben megtalálható a helyes válasz. A kilencedik és a hetedik kérdésben a tévesztések aránya 29,16 és 27,08%. Könnyebbnek bizonyult az első kérdés, ahol a csoport 20,83%-a tévesztett, és a hatodik, melyre 14,58% helytelen válasz érkezett.

2. táblázat: A szövegértést ellenőrző kérdések téves megválaszolása a teljes csoportban

	Ellenőrző kérdések	Hibák száma
1.	Miért szaladt ki a nyuszika a házikójából?	10
2.	Hová ment?	16
3.	Kivel találkozott először?	30
4.	Milyen állatokkal találkozott még?	41
5.	Miért irigyelte az állatokat?	26
6.	Mi szeretett volna lenni?	7
7.	Hogyan változott meg az idő az erdőben?	13
8.	Mit csináltak az erdei állatok?	25
9.	Mit csinált a nyuszika?	14
10.	Miért mondta a nyuszika végül, hogy jó nyuszinak lenni?	17

Következtetések

A kutatásban célul tűztük ki, hogy olyan magyar gyermekeket vizsgáljunk meg a GMP-diagnosztika tesztjeivel, akik kétnyelvű környezetben élnek, dominánsan magyar anyanyelvűek, és életükben, környezetükben kevésbé vagy jelentősebb mértékben jelen van az államnyelv, a román.

Az előző kutatások eredményei azt mutatják, hogy a gyermekek nagy százalékánál felmerülnek beszédfeldolgozási nehézségek (Gósy 2000; Macher 2002; Vančoné Kremmer 2002; Imre 2005, 2007; Gósy–Horváth 2006; Markó 2007; Bakk-Miklósi 2009). Ez egyaránt jellemző az egynyelvű magyar, illetve a magyar–szlovák, magyar–román kétnyelvű gyermekekre. A felmérés első hipotézisében ezekre a korábbi kutatásokra támaszkodva feltételeztük, hogy a felmért tanulóknak mind a mondatértési, mind a szövegértési képessége elmarad a korosztályukban elvártaktól. Az eredmények alapján ez a feltevés beigazolódott, a teljes felmért csoportban jelentős elmaradás figyelhető meg. A mondatértési tesztben az elvárt 100%-os teljesítményt a csoport 29,2%-a teljesítette. A mondatmegértés nélkülözhetetlen részfolyamat a beszédmegértés során, ezért az elmaradásokkal rendelkező diákok esetében nagy a veszélye annak, hogy a tanulási folyamatban és az információszerzés során pontatlanul értik meg az elhangzó beszédet. A szövegértésben jónak mondható 90–100%-ot összesen 14,6%-uk tudta teljesíteni. A tipikus fejlődé-

sű egynyelvű gyermekek körében végzett vizsgálatokkal (Markó 2007) összehasonlítva az itt kapott eredményeket az is megállapítható, hogy az elmaradás a standardhoz képest még jelentősebb, hiszen míg a fent említett felmérésben a mondatértésben elért átlageredmény 7 éveseknél 90,60%, a mostani felmérésben ez már csak 85,60%. A kétnyelvű gyermekek körében elvégzett vizsgálatokhoz képest is alacsonyabb a mondatértés szintje az általunk felmért csoportban, hiszen korábban mind a magyar–szlovák, mind a magyar–román kétnyelvű gyermekek 90% fölött teljesítettek ebben a tesztben (Vančoné Kremmer 2002, Bakk-Miklósi 2009). Hasonlóképpen csökkent a szövegértésben elért átlagérték is az egynyelvű 75% feletti, a magyar–szlovák 80% feletti és a magyar–román 70% feletti eredményekhez képest 57%-ra (Vančoné Kremmer 2002, Bakk-Miklósi 2009).

Ahogyan azt előző kutatások is alátámasztják, iskoláskorra a mondatmegértés sokkal jobban fejlődik, mint a szövegértés, de ezt az eredményességet a tanulók nem mozgósítják a szövegértésben (Gósy 2000; Vančoné Kremmer 2002; Imre 2007; Gósy–Horváth 2006; Markó 2007). Ez tapasztalható a felmért csoportban is, hiszen a mondatértési eredmények szignifikánsan magasabbak, mint a szövegértésben nyújtott teljesítmény [$t(47) = 8{,}31, p < 0{,}01$].

A mondatértési teszt hibaelemzése szintén összehasonlítható az előző kutatások eredményeivel. A legnehezebb szerkezeteknek az előidejűséget kifejező tagmondat megértése, az azonos alakú toldalékokkal ellátott részeshatározós mondat, valamint a ritka szórendiséget tartalmazó és az alanyt tagadó szerkezet bizonyult. A sorrend megegyezik más vizsgálatok eredményeivel (Vančoné Kremmer 2002; Markó 2007). A három mondatban bekövetkezett nagyszámú tévesztés arra enged következtetni, hogy a vizsgált életkorban a gyermekeknek még nehézséget okoz a bonyolult logikai szerkezetek nyelvi kifejezése.

A szövegértési eredmények alapján egyértelművé vált, hogy a gyermekek a hallott szövegek jó részét nem értik pontosan, nem ismerik fel az okokozati összefüggéseket, az időrendet, nem emlékeznek a fontos szereplőkre és az eseményekkel kapcsolatos részletekre. A hibaelemzés során kiderült, hogy a legnagyobb bizonytalanságok a kikövetkeztetendő válaszoknál és az összefüggések közötti kapcsolat megteremtésében van. Ez a jelenség, bár más mesével vizsgálva, az előző kutatásokban is felismerhető (Vančoné Kremmer 2002; Markó 2007). Emellett problémát okoz a tanulóknak az időrendiség felismerése (pl. a *Kivel találkozott először?* kérdésben), a részletező információk felsorolása (pl. *Milyen állatokkal találkozott még*? Helyes válasz: *rókával, vaddisznóval*), a lényegkiemelés (pl. *Mit csináltak az erdei állatok?* Helyes válasz: *igyekeztek elbújni*).

A meghallgatott mese a gyermekek életkori sajátosságaihoz leginkább alkalmazkodó műfaj. A tesztben elhangzó mindössze 3 perces, jól strukturált szöveg sokkal könnyebben hozzáférhető, mint azok a magyarázatok, a hangos beszédben elhangzó szövegek, amelyekkel a diákok az iskolában talál-

koznak. Ennek ellenére sem tudtak az életkoruknak megfelelő eredményeket elérni, ez azt jelzi számunkra, hogy a tanulási folyamat során is számos problémájuk akadhat a tananyag elsajátításában.

A második és harmadik feltevésünk a lakhely szerinti eloszlással függött össze. Azt feltételeztük, hogy a kisvárosi gyermekek eredményei gyengébbek a nagyváradi gyermekek eredményeinél. Ez a hipotézis részben beigazolódott, hiszen a nagyváradi gyermekek valóban jobban teljesítettek az Élesden felmért tanulókhoz képest, de nem mutatható ki különbség Nagyvárad és Székelyhíd szempontjából. A mondatértési eredményekben szignifikáns eltérés figyelhető meg a nagyváradi és élesdi, valamint a székelyhídi és élesdi gyermekek teljesítménye között. Ez azt jelentheti, hogy a harmadik hipotézis – melyben azt feltételeztük, hogy a nagyrészt románok által lakott kisvárosban a magyar nyelvű beszédmegértésben gyengébben teljesítenek a diákok – is részben igazolódott. Az élesdi tanulók eredményei elmaradnak mind a nagyváradiak, mind a székelyhídiak eredményeitől. A nagyváradi és a székelyhídi gyermekek jobb teljesítménye összefügghet azzal, hogy az ő környezetükben elsősorban magyarul beszélnek, míg Élesden jellemzőbb, hogy a gyermekek az iskolán és a családon kívül sok színtéren (szomszédok, a lakhely környéki pajtások, bevásárlás stb.) a román nyelvvel szembesülnek. Ez a nyelv az ő életükben jelentősebb hatást gyakorol. Ugyanakkor, mivel az iskolában magyarul tanulnak, a gyenge beszédértési folyamatok megakadályozzák őket a sikeres tanulásban. Bár kimutatható az élesdi eredmények alapján a magyar nyelvű megértésbeli elmaradás a két jellemzőbben magyarlakta településhez képest, további vizsgálódás lenne szükséges az elmaradások konkrét okainak feltérképezésére.

A jelen kutatás adatai is megerősítik az előző vizsgálatok (Gósy 2000; Macher 2002; Vančoné Kremmer 2002; Imre 2005, 2007; Gósy–Horváth 2006; Markó 2007; Bakk-Miklósi 2009) során már kialakult álláspontot, mely szerint a gyermekek nagyobbik részének beszédfeldolgozási képessége elmarad attól a szinttől, amely elvárható lenne, s amelynek alapján az iskolai követelmények strukturálódnak. A kétnyelvű környezet tovább színezi a beszédfeldolgozás folyamatát, s az eredmények azt látszanak igazolni, hogy másként alakulnak ilyen helyzetben a beszédértés folyamatai, ezért szükséges nagyobb hangsúlyt fektetni a magyar–román kétnyelvű gyermekek diagnosztizálására, s legfőképpen annak a nyelvnek a megerősítésére, fejlesztésére, melyet a gyermekek a tanulási folyamatban is használnak.

Irodalom

Bakk-Miklósi Kinga 2009. *Kétnyelvűvé válásunk útjain.* Ábel, Kolozsvár.

Bartha Csilla 1999. *A kétnyelvűség alapkérdései.* Nemzeti Tankönyvkiadó, Budapest.

Bloomfield, Leonard 1933. *Language.* Holt, New York.

Büky Béla – Egyed András – Pléh Csaba 1984. *Nyelvi képességek, fogalomkincs, megértés. A pszicholingvisztika gyakorlati lehetőségei.* Tankönyvkiadó, Budapest.

Diebold, Richard 1961. Incipient bilingualism. *Language* 37. 97–112.
Edwards, John 2006. Foundations of bilingualism. In Bathia, Tej K. – Ritchie, William C. (eds): *The handbook of bilingualism*. Blackwell Publishing, Malden–Oxford–Victoria. 7–31.
Gósy Mária 1995/2006. *GMP-diagnosztika. A beszédészlelés és a beszédmegértés folyamatának vizsgálata, fejlesztési javaslatok*. Nikol Kkt., Budapest.
Gósy Mária 2000. *A hallástól a tanulásig*. Nikol Kkt, Budapest.
Gósy Mária 2005. *Pszicholingvisztika*. Osiris, Budapest.
Gósy Mária – Horváth Viktória 2006. A beszédészlelés és a beszédmegértés összefüggései kisgyermekkorban. *Magyar Nyelvőr* 130. 470–481.
Göncz Lajos 2004. *A vajdasági magyarság kétnyelvűsége. Nyelvpszichológiai vonatkozások*. MTT, Szabadka.
Grosjean, Françios 1982. *Life with two languages. An introduction to bilingualism*. Harvard University Press, Cambridge.
Haugen, Einar I. 1953. *The Norwegian language in America*. University of Pennsylvania Press, Philadelphia.
Hoffmann, Charlotte 1991. *An introduction to bilingualism*. Longman, London.
Imre Angéla 2005. Kétnyelvű gyermekek beszédpercepciós teljesítménye. *Beszédkutatás 2005*. 123–133.
Imre Angéla 2007. A beszédmegértés és az olvasás összefüggése. In Gósy Mária (szerk.): *Beszédészlelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol Kkt., Budapest. 184–201.
Karmacsi Zoltán 2007. *Kétnyelvűség és nyelvelsajátítás*. Poliprint, Ungvár.
Lanstyák István – Szabómihály Gizella 1997. *Magyar nyelvhasználat – iskola – kétnyelvűség*. Kalligram, Pozsony.
Lengyel Zsolt 1981. *A gyermeknyelv*. Gondolat, Budapest.
Macher Mónika 2002. Cigány gyermekek beszédészlelésének és beszédmegértésének vizsgálata. *Beszédkutatás 2002*. 118–130.
Markó Alexandra 2007. A mondat- és szövegértés jellemzői és összefüggése 6–9 éves korban. In Gósy Mária (szerk.): *Beszédészlelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol Kkt., Budapest. 285–301.
Menyhárt Krisztina 2007. A magyarországi bolgár közösség nyelvi helyzete. *Kisebbségkutatás* 9/1. http://www.hhrf.org/kisebbsegkutatas/kk_2007_01/cikk.php?id=1775.
Navracsics Judit 2010. *Egyéni kétnyelvűség*. Szegedi Egyetemi Kiadó, Szeged.
Pléh Csaba 1998: *A mondat megértés a magyar nyelvben. Pszicholingvisztikai kísérletek és modellek*. Osiris, Budapest. 13–39.
Vančoné Kremmer Ildikó 2002. A beszédészlelés és a beszédmegértés vizsgálata magyar–szlovák kétnyelvű gyermekeknél. In Lanstyák István – Simon Szabolcs (szerk.): *Tanulmányok a kétnyelvűségről*. Kalligram Könyvkiadó, Pozsony. 73–96.

AKUSZTIKAI HANGOSZTÁLYOK FELISMERÉSÉN ALAPULÓ, NEMLINEÁRIS IDŐVETEMÍTÉS MEGVALÓSÍTÁSA A MONDATHANGLEJTÉS ÉS A SZÓHANGSÚLYOZÁS OKTATÁSÁHOZ

Kiss Gábor – Vicsi Klára

Bevezetés

Kutatási célunk egy olyan általános, több nyelvre működő, nemlineáris időbeli vetemítési eljárás kifejlesztése volt, ami azonos tartalmú beszédminták összehasonlítására alkalmas, főleg a beszédminták prozódiai jellemzőinek összevetésére oktatás-alkalmazásokhoz. Cél volt az is, hogy az eredményt mind vizuálisan, mind számértékkel, illetve egyéb szöveges módon megjelenítsük a tanulónak.

A prozódiának a beszédfolyam automatikus tagolásában, a szintaktikai és szemantikai információ helyes közlésében igen nagy szerepe van (Olaszy 2010). A magyar beszédben a prozódiával fejezzük ki a mondatok modalitását is. A prozódia sok esetben a közlés értelmezését is elősegíti.

Ennek a munkának így egyik lehetséges gyakorlati haszna, hogy az általunk kifejlesztett eljárás az audiovizuális nyelvoktatásban vagy a hallássérültek kiejtéstanításában a prozódia fejlesztéséhez jól felhasználható.

A BME TMIT Beszédakusztikai Laboratóriumában korábban kifejlesztettek egy HMM alapú, automatikus intonációosztályozó rendszert, amelyet hallássérültek beszédterápiájában használtak fel (Sztahó et al. 2009; Szaszák et. al. 2009). Ez a rendszer automatikusan különbséget tudott tenni a helyes és a helytelen mondatintonáció között, de a vizuális visszajelzésnél nehezen volt értelmezhető az etalon és az aktuális minta összehasonlítása. Ennek oka az volt, hogy ez az eljárás csupán mondatok időtartamát egyenlítette ki egyformára, és ez egyfajta lineáris skálázásnak (lineáris vetemítésnek) tekinthető.

Ahhoz, hogy ugyanazon mondat két egyén által kiejtett formáját össze lehessen hasonlítani prozódiai szempontból, azokat hangszinten azonos időbeli hosszra kell skálázni, azaz egymáshoz vetemíteni. A prozódia összehasonlításánál a konkrét probléma az, hogy a beszéd tartalmaz nemlineáris megnyúlásokat és rövidüléseket, és ezek nem feltétlenül számítanak hibás ejtésnek. Amennyiben tehát lineáris időskálázást használnánk a prozódiai jellemzők összehasonlításához, akkor az egyes hangok nagy valószínűséggel nem pontosan ugyanazon két időpont közé esnének. Így a jellemzők összehasonlítása-

nál ott is eltérés mutatkozna, ahol a tényleges mintákban a bemondás jóságában nem lenne lényeges különbség, hiba. Csupán abból fakadna az eltérés, hogy az időbeli eltolódás miatt a beszéd nem megfelelő szakaszából mérnénk az adatokat.

Az ideális időbeli összehasonlítás az, amikor a két mintát az egyes beszédhangok mentén illesztjük egymáshoz. Ezt alkalmazzák gépi beszédfelismerésnél, de ez a megoldás nyelvfüggő (Baker et. al. 2009; Huang–Deng 2010).

Több nyelvre használható prozódiai összehasonlításhoz olyan vetemítő eljárás kell, ahol szigorúan a beszédhangok mentén történik az illesztés, és csak azokon belül lineáris a skálázás. Ezzel nemcsak azt tudná közölni a program a tanulóval, hogy az általa bemondott mondat prozódiai jellemzői nagyban eltérnek a referenciamondattól, hanem képes lehetne a hibát lokalizálni is, és pontosan megfogalmazni, hogy mit ejt rosszul a tanuló. Például, hogy az adott szónak nem az első, hanem az utolsó szótagját ejtette hangsúlyosan. A tanulmányban bemutatunk egy általunk kifejlesztett, olyan több nyelvre működő automatikus szegmentáláson alapuló nemlineáris vetemítő eljárást, amely lehetővé teszi a különböző prozódiai jellemzők meglehetősen pontos összehasonlítását mondat egységű kiejtést tanítási folyamatokban.

Anyag és módszer

A kutatásunk során három, különböző nyelvű beszédadatbázist használtuk, az MRBA magyar (Vicsi et. al. 2004), a Kiel német (Benno 2005) és a TIMIT angol (Garofolo et al. 1993) nyelvű adatbázisokat. Mind a három adatbázist beszédhang egységben szegmentáltunk és címkéztünk. Ezek az adatbázisok olvasott, folyamatos, fonetikailag kiegyensúlyozott szöveget tartalmaznak. A feladat szempontjából fontos jellemzőik leolvashatóak az 1. táblázatról.

1. táblázat: Az eljárás kifejlesztése során felhasznált adatbázisok

Az adatbázis neve	A beszélők száma	Bemondási környezet	Az adatbázis hossza percben
MRBA	100	iroda, otthon	330 perc
Kiel	50	iroda, otthon	180 perc
TIMIT	40	iroda, otthon	310 perc

A tárolt hangminták digitalizálása 16 kHz-es mintavételi frekvenciával, 16 bites lineáris kvantálással készült. Az eljárásunkban három alapvető módszert alkalmazunk: gépi szegmentálás, mintaillesztés és vetemítés a referencia és a bemondott mondat összehasonlítására; valamint az előbbiek tesztelése. Az eljárásunk újdonságát az jelenti, hogy a gépi szegmentálás folyamatában általános akusztikai hangosztályokat használunk, ami lehetővé teszi a szegmentálást több nyelv esetén is általános akusztikai környezetben. Ebből következik, hogy a nemlineáris idővetemítést is általánosan több nyelvre alkalmasan oldottuk meg.

A gépi szegmentálásban használt újszerű hangosztályozás alapötletét a Vicsi Klára és Vig Attila által publikált tanulmány (1998) adta. Ennek a lényege, hogy képesek voltak neurális háló segítségével, több nyelven 80-90%-os szegmentálási pontosságot elérni az általuk definiált akusztikai hangosztályokban süketszobai körülmények között rögzített hangfelvételek esetén.

Eljárásunk során mi is ezeket az akusztikai hangosztályokat használtuk a gépi szegmentáláshoz. Az akusztikai hangosztályok és a hozzájuk tartozó beszédhangok, beszédhangrészek a 2. táblázatból olvashatóak le.

2. táblázat: A definiált akusztikai hangosztályok felsorolása és annak jelölése, hogy mely osztályba mely beszédhangok, vagy összetett hangok esetén mely beszédhangrészek kerültek magyar nyelv esetén [+*: a zöngés zárszakasz (zönge), –*: a zöngétlen zárszakasz [csend])

Az akusztikai hangosztály neve	Az akusztikai osztály jele	Osztályba sorolt beszédhangok (SAMPA-jele)
Magánhangzók 1	mv	O, A:, E, o, o:, u, u:, 2, 2:
Magánhangzók 2	hv	e:, i, i:, y, y:
Zöngétlen spiránsok	s-	f, s, S, h,
Zöngés spiránsok	s+	v, z, Z
Nazálisok és likvidák	na	r, l, j, m, n, J
Zöngés zárszakasz	b+	+*
Zöngés felpattanás	vo	b, b:, d, d:, g, g:, dz, dz:, dZ, dZ:, d', d':
Zöngétlen zárszakasz	b-	–*
Zöngétlen felpattanás	uv	p, p:, t, t:, k, k:, ts, ts:, tS, tS:
Csend	si	

Az osztályozás lényege, hogy a nyelvtől független artikulációs konfigurációkat határoz meg, amelyek általánosak. Amennyiben két ugyanazon osztálybeli egység áll egymás mellett, akkor a szegmentáló nem különíti el a két elemet. Ez azonban nem olyan súlyos probléma, ugyanis prozódiai szinten elégséges a szótagszintű pontosság is. Szótaghibát csak az okoz, ha két egy osztálybeli magánhangzó kerül egymás mellé (*boa*), ez azonban ritkán fordul elő, és egyszerű szabályrendszerrel nagyrészt a hiba kiküszöbölhető (elemhosszfigyelés).

A gépi szegmentáláshoz Hidden Markov Model (HMM) elvű felismerő motort használtunk (Mihajlik 2010), amellyel a 10 fonetikai osztály akusztikus modelljét tanítással hozzuk létre. Az összehasonlításhoz nemlineáris vetemítéssel egyforma időskálára hozzuk a mondatok hangelemeit. Adott a referenciamondat, ahol ismert az időszerkezet, és a felhasználó ugyanezt a mondatot mondja be, ahol viszont nem ismerjük a beszélő ejtési sajátosságait (hadar, lassan beszél stb.). Eljárásunk feltételezi, hogy címkézési szinten ismerjük a referenciamondat időszerkezetét, illetve hogy a felhasználó kooperál.

A megfelelő vetemítéshez szükséges, hogy ismerjük a bemondott mondat beszédhang- vagy szótagszintű időszerkezetét is. Ehhez a bemondást automatikusan szegmentálni kell. A szegmentálás után összehasonlítjuk a bemondás és a referencia magánhangzóit, és a megfelelőket összekapcsoljuk. Ezután a referencia- és a bemondott mondatot vetemítjük egymáshoz, amit a magánhangzók időkoordinátái alapján végezzük el. Ezután teszteljük a vetemítés jóságát a mondatintonáció és a mondatintenzitás dinamikája alapján. Az eredmény összehasonlíthatóságához a tesztelést elvégezzük lineáris skálázás esetén is. Ehhez kirajzoljuk a mondatintonáció és mondatintenzitás dinamikáját, illetve kiszámoltatjuk ezeknek a referenciához képesti négyzetes távolságát. (Vagyis vettük mintapontonként a négyzetes eltérést, ezeknek az átlagát, majd gyököt vontunk az eredményből.)

Az eljárás megvalósítása

A szegmentáláshoz szükséges HMM modellek tanításához az MRBA adatbázis 30%-át használtuk fel. A program bemenetként a tanításra felhasznált beszédmintákat, illetve a hozzájuk tartozó annotációs fájlokat kapja meg. Első lépésben előfeldolgozást végez, majd a hangfájlokat feldarabolja az annotációs fájl ismeretében. Az előfeldolgozás során a periférikus hallásnak megfelelő felbontásban (kritikus sávokban) színképelemzést végez. Kiszámítja a hangminta kritikus sávokban lévő energiaértékeit, illetve azok deriváltjait. 20 Hztől 7,7 kHz-ig huszonegy sávra bontja a hangmintát. Ezek az értékek alkotják a jellemzővektorokat a modelleket létrehozó szoftver bemeneteként. A program hangosztályonként háromállapotú Markov-lánccal dolgozik. Az előfeldolgozással nyert jellemzők alapján és az annotációs fájl segítségével létrehozza az egyes akusztikai hangosztályok modelljeit. Így minden fonetikai hangosztályhoz létrehoz a tanító eljárás egy külön akusztikai modellt. Ennek a módszernek az egyik előnye például, hogy bármikor bármelyik modell könnyen cserélhető, illetve bővíthető vagy csökkenthető a modellek száma.

Ezután a felismerés, illetve a szegmentálás a következőképpen történik. Adott a bemondás hipotetikus leirata (feltesszük, hogy kooperatív a bemondó, és törekszik a referenciamondat bemondására, legfeljebb kisebb hibákat ejt). A leiratot egy fonetikai átíró átírja a megfelelő akusztikai hangosztályok szerint. Ezután a HMM elvű gépi felismerő megkapja a bemondást, az akusztikai hangosztályokká átírt leiratot és az akusztikai hangosztályok akusztikai modelljeit. Ezeknek a segítségével elvégzi a felismerést, szegmentálást a bemondáson, és azt a PRAAT program (Boersma–Weenink 2001) annotációs formátumában visszaadja.

Ezután valósítottuk meg a nemlineáris idővetemítést. Ideális esetben a vetemítési feladat egyszerű. Tegyük fel, hogy a két mondat beszédhangjainak sorozata megegyezik, ebben az esetben a vetemítés csupán annyi lenne, hogy a bemondott mondat beszédhangjainak hosszát lineárisan úgy módosítanánk egyesével, hogy azok a referenciával azonos hosszúak legyenek. Sajnos ezzel a feltételezéssel nem élhetünk, mivel a beszéd rendkívül változatos (beszéd-

hangbeszúrások, kihagyások, elharapások, másképp ejtések). Még akkor sem egyezik meg feltétlenül a két mondat, ha a bemondó törekszik a pontosságra.

A szegmentáló a legjobban a magánhangzókra működött 90%-os körüli pontossággal (magyar nyelv esetén). Másrészt a beszéd prozódiai elemzéséhez elég a magánhangzónkénti pontosság, vagyis elégséges, ha magánhangzónként végzünk időbeli vetemítést nemlineárisan, és két magánhangzó között pedig lineárisan. Így kézenfekvőnek látszott úgy dönteni, hogy a két mondatban a magánhangzók helyzetét illesztjük dinamikusan, és a köztük lévő részt pedig lineárisan.

Kulcspontoknak neveztük el azokat a pontokat a referencia- és a vetemítendő mondatban, amelyek illesztése dinamikus lesz. Ilyen kulcspontok a magánhangzók, illetve a beszéd kezdete és vége. Mivel a kulcspontok sorozata nem feltétlenül egyezik meg, szükséges egy mintaillesztési eljárás. Amennyiben a két mintasorozat ugyanaz, akkor az egymás alatti kulcspontokat illeszti. Probléma akkor adódik, ha a két mintasorozat eltér egymástól. Feltehetjük, hogy az eltérés nem lesz jelentős (a bemondó kooperatív), emiatt háromféle eltéréssel számolhatunk:

– beszúrás: a bemondó beszúrt egy, esetleg több hangot, amely eredetileg nincs jelen;

– kihagyás: a bemondó kihagyott egy, esetleg több hangot (elharapta), amely a referenciamondatban megtalálható;

– másképp ejtés: a bemondó más hangot ejt ki, mint a referenciában eredetileg volt.

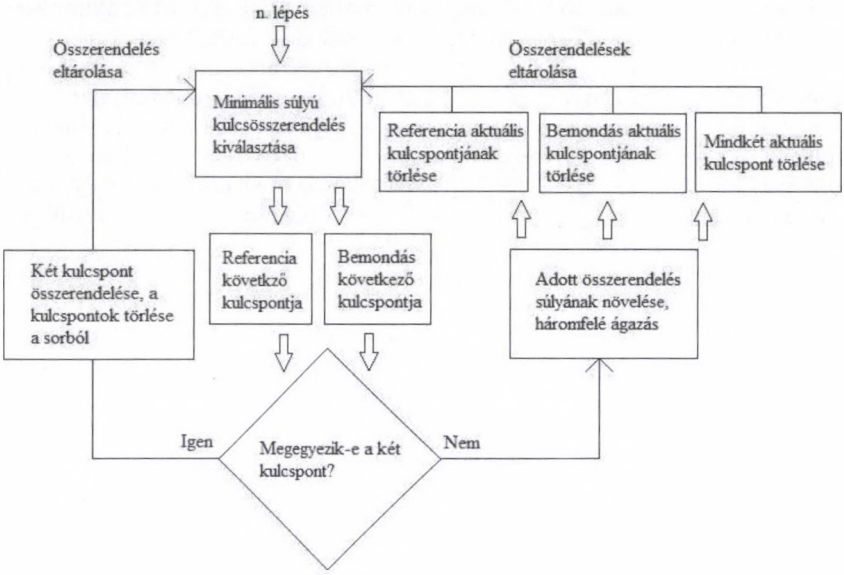
Ezek alapján az illesztést a következő gépi eljárás szerint végeztük. Felvettük kulcspontoknak a magánhangzókat, illetve a mondat elejét és végét. Az algoritmus végigment a referencia- és a vetemítendő mondat kulcspontjain egyesével (mondat eleji kulcsponttal kezdve), ha a két elem megegyezett, mindkettőn lépett egyet előre, és összekapcsolta őket, ha nem egyeztek meg, az adott út súlyát növelte eggyel, és háromfelé ágazott az algoritmus:

– a referenciamondatban lép egyet előre az algoritmus, míg a vetemítendőben nem, és ehhez a referenciaponthoz semmit sem kapcsol (törli a kulcspontot);

– a vetemítendő mondatban lép egyet előre az algoritmus, míg a referenciában nem, és ehhez a vetemítendő ponthoz semmit sem kapcsol (törli a kulcspontot);

– mindkét sorban lép egyet az algoritmus, és ezeket a pontokat senkivel se kapcsolta össze (mindkét kulcspont törlésére került).

Ezután iteratívan folytatódott az algoritmus futása. Ha megint eltért egymástól a két mintasor, akkor megint háromfelé ágazott, míg az adott útvonal súlyát növelte eggyel. Mindig az aktuálisan minimális súlyú úton végezte a számítást. Az algoritmus akkor fejeződött be, ha mindkét sor esetén elérte a vége kulcspontokat (ezek nem törölhetőek). A mintaillesztő algoritmus egy lépése látható az 1. ábrán.



1. ábra
A mintaillesztő algoritmus egy lépése

Ha a két mintasorozat megegyezik, akkor valóban az ideális működést biztosítja. Egyébként a vetemítés működésének helyességét a végeredményekből adódóan láthatjuk, ahol kifejezetten szélsőséges példán is kipróbáltuk az eljárást. Ezután a következő a vetemítési algoritmust, programot implementáltuk az időben nemlineáris vetemítéshez:

– beolvassa a referencia- és a vetemítendő mondat szegmentálását;

– bejelöli a referencia- és a vetemítendő mondatban a beszéd elejét és a végét, ami nem feltétlenül egyezik meg a hangminta elejével, illetve végével;

– levágja a két mondat felesleges részét (vagyis ami a beszéd kezdete előtt, illetve után van), és ezután a maradékot 0–1 időintervallumra skálázza;

– a magánhangzók közepei kulcspontok lesznek, az „mv"-beliek A típusú, a „hv"-beliek B típusú kulcspontok;

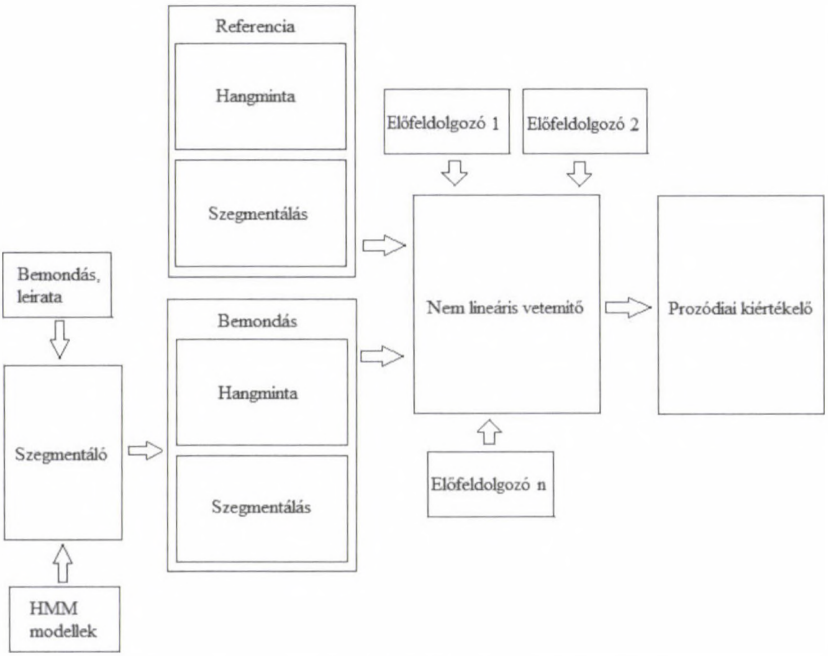
– ezután elvégzi a korábban tárgyalt mintaillesztési eljárást, ahol véglegesíti, illetve összerendeli a kulcspontokat;

– a vetemítendő mondatban a kulcspontokat a referenciamondatban lévő időszerkezetre helyezi át, és a köztük lévő részt lineárisan skálázza.

Ezután pedig egy megfelelő távolságszámító algoritmussal a két mondat prozódiai távolsága meghatározható, így eldönthető, hogy helyes-e a bemondott mondat prozódiai szempontból vagy sem. Egy ennél még „okosabb" ki-

értékelő program akár azt is közölheti a felhasználóval, hogy milyen tévedéseket ejtett, és azokon hogyan képes változtatni.

Az időben nemlineáris vetemítő eljárást az eddigiek alapján implementáltuk, illetve megalkottunk egy rendszert, ami képes két mondat prozódiai jellemzőinek összehasonlítására. A teljes eljárás folyamatábrája 2. ábrán látható.



2. ábra
A prozódiai kiértékelő folyamatábrája

Az eljárás bemenete a referenciamondat hangmintája, illetve a hozzátartozó annotáció. Ez látható az ábrán a referenciarésznél (középen fent). Az eljárás további bemenete a bemondott mondat hangmintája, illetve a hozzátartozó leirat. Ezután az eljárás a leirat alapján elkészíti a bemondott mondat hangosztályszintű szegmentálását. A szegmentálás eredményét és a bemondott hangmintát a referenciamondat hangmintájával és szegmentálásával együtt megkapja a nemlineáris vetemítő modul, amely a korábban tárgyaltak szerint elvégzi az időben nemlineáris vetemítést az egyes előfeldolgozók által mért jellemzők mentén. Ezután a prozódiai kiértékelő elvégzi a görbék elemzését, és a felhasználó számára valamilyen visszajelzést nyújt.

Eredmények

Az eredményeket a feldolgozási sorrend szerint mutatjuk be. Először a gépi szegmentálás, majd a vetemítési eljárás eredményeit közöljük. A gépi szegmentáló kiértékelése a következő automatikus eljárással készült:

– először végignézte az eljárás a szegmentálást a referenciához képest. Ha bárhol is két szegmenshatár csupán 25 ms-on belül tért el, azt nem tekintettük hibának. Ilyenkor a felismert mondatban a szegmenshatárt áthúzta a referencia-szegmenshatár időpontjába. Erre azért volt szükség, mert a kézi szegmentálás is legalább ekkora hibahatárral rendelkezett.

– A javítás után 10 ms-os lépésközzel összehasonlította a referenciamintát és a felismert eredményt. Ha az adott mintavételi pontban megegyezett a két annotáció, akkor az abban az időpontban történő felismerést sikeresnek vette, ha eltért, akkor sikertelennek, és eltérés esetén megjegyezte a tévesztést is, így képes volt tévesztési mátrixot építeni.

A fent definiált kiértékeléssel az eljárást három nyelvre teszteltük: magyar, német és angol. Bár a tanítás során csak magyar beszédmintákat használtuk fel az akusztikai modellek létrehozásához, így várható, hogy az angol és német nyelv esetén valamivel alacsonyabb lesz a gépi szegmentálás pontossága. Magyar nyelv esetén az MRBA adatbázis azon beszédmintáin teszteltük, amelyek nem szerepeltek a tanítás során. Vagyis a teljes adatbázis 70%-án, ez körülbelül 230 perc. A német nyelven való teszteléshez a Kiel Korpuszt használtuk, ez körülbelül 180 perc. Az angol nyelven való teszteléshez a TIMIT adatbázist használtuk, ez körülbelül 310 perc. A tesztelési eredményeket a 3., 4. illetve az 5. táblázat tévesztési mátrixaiból lehet leolvasni.

A magyar nyelv esetén az eredmény valamivel jobb, mint a már korábban megvalósított neurális hálókkal való felismeréssel kapott eredmények (85%) (Vicsi–Vig 1998), annak ellenére, hogy ott süketszobai, zaj nélküli felvételeket szegmentáltak. A mi esetünkben pedig természetes irodai körülményeknek megfelelő zajos környezetben rögzített hangfelvételeken történik a szegmentálás. A többi nyelv esetén viszont alacsonyabb a szegmentálási pontosság, bár az összehasonlítást nem lehet tökéletesen elvégezni, mivel Vicsi Klára és Vig Attila tanulmányában a többi nyelvből is tanítottak a felismerés esetén. Feltehetőleg ha nyelvenként külön modelleket használnánk, akkor a magyarhoz hasonló szegmentálási pontosságot kapnánk. A táblázatból jól látható, hogy a magánhangzók (mv és hv) felismerési eredménye a legjobb (80–93%). Így ez megfelelő lehet a későbbi időben nemlineáris vetemítéshez. Az összes európai nyelv esetén az akusztikai hangosztályok használata elegendő az adott nyelv összes hangjának a besorolásához. A prozódiai összehasonlításhoz a tíz osztály használata elégséges, ugyanis a célunk nem felismerni a bemondott hangsorozatot, csupán megtalálni a beszédhangok, kiváltképp a magánhangzók határait. A vetemítő eljárásunk tesztelésénél a következő szempontokat vettük figyelembe:

– az eljárásban az újdonság a „nemlineáris idővetemítés";

– az eljárásunk helyességének a tesztelése a cél, ezért törekedtünk, hogy a bemondás dallammenete és intenzitásgörbéje hasonló legyen, miközben tartalmazzon nemlineáris megnyúlásokat, rövidüléseket.

3. táblázat: Magyar nyelven történt tesztelés tévesztési mátrixa

	mv	hv	s+	s-	na	vo és b+	uv és b-	si
mv	**93%**	2%	0%	0%	2%	1%	1%	1%
hv	4%	**90%**	0%	0%	4%	1%	1%	0%
s+	9%	5%	**74%**	2%	5%	3%	2%	0%
s-	5%	1%	1%	**86%**	2%	1%	4%	0%
na	17%	9%	0%	0%	**68%**	3%	2%	1%
vo és b+	5%	4%	1%	0%	6%	**82%**	1%	1%
uv és b-	4%	1%	0%	1%	3%	1%	**88%**	2%
si	4%	1%	0%	4%	2%	3%	2%	**84%**

Összesített felismerés: 86% (jó minták/összes minta).

4. táblázat: Német nyelven történt tesztelés tévesztési mátrixa

	mv	hv	s+	s-	Na	b+	b-	vo	uv	si
mv	**83%**	2%	2%	3%	4%	2%	0%	1%	2%	1%
hv	6%	**80%**	1%	1%	6%	3%	1%	1%	1%	0%
s+	8%	13%	**74%**	0%	2%	2%	1%	0%	0%	0%
s-	6%	4%	0%	**81%**	4%	1%	2%	1%	1%	0%
na	12%	7%	1%	3%	**69%**	5%	1%	1%	1%%	0%
b+	9%	7%	1%	2%	2%	**73%**	0%	2%	0%	4%
b-	2%	2%	1%	6%	6%	3%	**76%**	0%	4%	0%
vo	4%	3%	3%	0%	2%	10%	4%	**74%**	0%	0%
uv	7%	6%	0%	6%	6%	0%	20%	1%	**51%**	3%
si	5%	0%	1%	0%	3%	4%	2%	0%	0%	**86%**

Összesített felismerés: 78% (jó minták/összes minta).

5. táblázat: Angol nyelven történt tesztelés tévesztési mátrixa

	mv	hv	s+	s-	na	b+	b-	vo	uv	si
mv	**93%**	3%	0%	1%	1%	0%	1%	0%	0%	0%
hv	6%	**86%**	1%	1%	3%	1%	1%	0%	0%	0%
s+	11%	4%	**76%**	1%	2%	1%	1%	1%	1%	2%
s-	5%	3%	1%	**87%**	1%	0%	1%	0%	1%	0%
na	26%	7%	1%	2%	**59%**	2%	1%	1%	1%	0%
b+	5%	2%	1%	3%	5%	**74%**	2%	7%	1%	1%
b-	5%	2%	1%	5%	2%	1%	**79%**	1%	4%	1%
vo	7%	4%	2%	3%	7%	2%	2%	**70%**	2%	1%
uv	8%	5%	2%	1%	5%	1%	4%	1%	**71%**	1%
si	2%	2%	1%	3%	2%	1%	6%	1%	1%	**80%**

Összesített felismerés: 80% (jó minták/összes minta).

Tehát kettős célunk volt, egyrészt megmutatni, hogy az eljárás nem rosszabb, mint a lineáris skálázás, másrészről korrektebb az általunk vizsgált prozódiai jellemzők távolságának meghatározása esetén. Háromféle bemondásra teszteltük le az eljárást, majd végeztük el a prozódiai kiértékelést. A bemondott mondat így hangzott: *Az északi szél nagy vitában volt a nappal, hogy kettejük közül melyikük az erősebb.* A bemondás módjai:

– normális tempóban: referenciának (6320 ms),

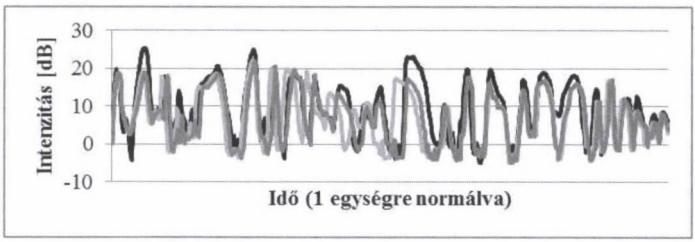
– normális tempóban (törekedve, hogy a bemondás időszerkezete megegyezzen a referenciával): bemondottnak (6213 ms),

– más tempóban (szüneteket beleütemezve) bemondva, ezáltal erős, nemlineáris megnyúlásokat garantálva a mondatban (8101 ms).

A két mondat prozódiai távolságának kiszámításához, szemléltetéshez az alaphang és az intenzitás dinamikáját használtuk fel. A távolságszámítás során négyzetes távolságszámítást választottuk, és ennek átlagát számoltuk. A mondatintonáció (dallammenet), illetve az intenzitás kiszámításához PRAAT programot használtuk. Az alaphang kiszámítása során az ablak mérete 50 ms, a lépésköz 10 ms volt. Az intenzitás kiszámítása során az ablak mérete 100 ms, a lépésköz 10 ms volt.

Az adott mondat 26 magánhangzót tartalmaz, így 28 pont dinamikus vetemítése történik meg, és a pontok között a lineáris skálázás lesz. A kiértékelő eljárással kirajzoltattuk az alaphangot, illetve az intenzitást lineáris skálázással és nemlineáris vetemítéssel is, összehasonlítva a referenciával mindkét bemondás esetén. Az alaphangot, illetve az intenzitást normáltuk az első valódi értékre (az első mért értéket kivontuk a többiből), illetve az alaphangot interpoláltuk azokon a szakaszokon, ahol nem volt zönge. (A zöngétlen szakasz két végpontját összekötöttük egy egyenes szakasszal.)

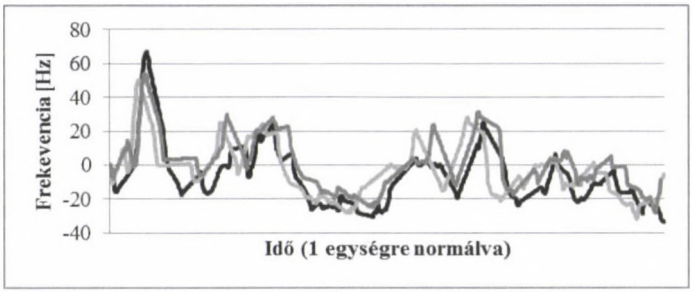
A 3., 4., 5. és 6. ábrákon a vízszintes tengelyen a referenciamondat szerint (6320 ms) 1 egységre lett skálázva, és a tengely 0–1 intervallumban lineáris.



3. ábra

A referenciamondat és a „jól" bemondott mondat alaphang-dinamikájának összehasonlítása, különböző vetemítések esetén (fekete: referencia, világosszürke: bemondott mondat lineáris vetemítéssel, sötétszürke: bemondott mondat dinamikus vetemítéssel)

Az átlagos négyzetes távolság lineáris vetemítés esetén 13,5 Hz lett, ezzel szemben dinamikus vetemítés esetén 10,1 Hz.

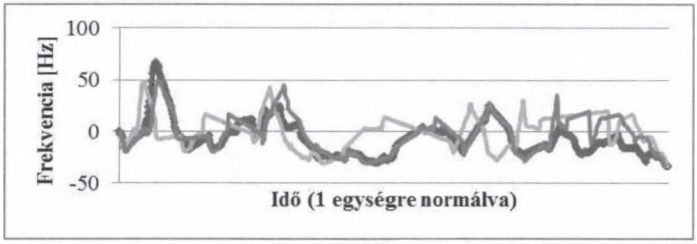


4. ábra
A referenciamondat és a „jól" bemondott mondat intenzitás dinamikájának összehasonlítása, különböző vetemítések esetén (fekete: referencia, világosszürke: bemondott mondat lineáris vetemítéssel, sötétszürke: bemondott mondat dinamikus vetemítéssel)

Az átlagos négyzetes távolság lineáris vetemítés esetén 6,9 dB, ezzel szemben dinamikus vetemítés esetén 5,5 dB lett. A 3. ábrán látható a két „ugyanúgy" bemondott mondat esetén a mondatintonáció (dallammenet), illetve a 4. ábrán az intenzitás dinamikája az egyes vetemítések esetén. Az ábrák alapján levonhatóak az első teszt tanulságai az időben nemlineáris vetemítésről. Jól látható, hogy az időben nemlineárisan vetemített görbe (sötétszürke) lényegében megegyezik a referenciagörbe dinamikájával (fekete). A lineárisan vetemített görbe (világosszürke) is nagyjából követi a referenciagörbe dinamikáját, bár nem teljesen, az eltérés láthatóan nem jelentős, attól még az időben nemlineárisan vetemített görbe láthatóan pontosabban követi a referenciát. Ez abból adódik, hogy ebben az esetben a referenciabemondás és a vetemítendő mondat időszerkezete lényegében megegyezik, és így láthatjuk, az időben nem lineáris vetemítés nem ad rosszabb eredményt, mint a lineáris.

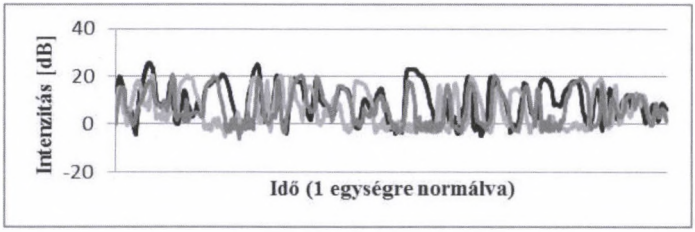
Az előzővel szemben, a második teszt esetén szembetűnő különbség mutatkozik a két vetemítési eljárás között (5., 6. ábra). Az ábrákon kifejezetten jól látszik az időben nem lineáris vetemítés előnye. A sötétszürke görbe (időben nemlineárisan vetemített) lényegében együtt mozog a referenciagörbével (fekete). Ezzel szemben a lineárisan vetemített görbe (világosszürke) a referenciától lényegesen eltérően mozog helyenként ellenkezően, mint a referenciagörbe. Ez a jelenség a nemlineáris megnyúlásokból és rövidülésekből adódik. Tehát ilyen esetben, ha lineáris vetemítéssel hasonlítanánk össze a két mondatot, akkor a prozódiai kiértékelő program több helyen is hibát jelezne a

lineáris skálázás esetén, míg az időben nemlineáris vetemítés esetén feltehetőleg azt közölné, hogy a bemondás prozódiai szempontból helyes.



5. ábra
A referenciamondat és az időben „rosszul" bemondott mondat alaphang-dinamikájának összehasonlítása, különböző vetemítések esetén (fekete: referencia, világosszürke: bemondott mondat lineáris vetemítéssel, sötétszürke: bemondott mondat dinamikus vetemítéssel)

Az átlagos négyzetes távolság lineáris vetemítés esetén 22,1 Hz, ezzel szemben dinamikus vetemítés esetén 11,5 Hz lett.



6. ábra
A referenciamondat és az időben „rosszul" bemondott mondat intenzitás dinamikájának összehasonlítása, különböző vetemítések esetén (fekete: referencia, világosszürke: bemondott mondat lineáris vetemítéssel, sötétszürke: bemondott mondat dinamikus vetemítéssel)

Az átlagos négyzetes távolság lineáris vetemítés esetén 11,4 dB, ezzel szemben dinamikus vetemítés esetén 7,4 dB lett.

Összegzés

Az eddig tárgyalt eredmények alapján kijelenthetjük, hogy az általunk tervezett eljárás működik, és az alkalmas lehet két mondat prozódiai távolságának kiértékelésére, illetve azok korrekt módon való vizuális ábrázolására több nyelv esetén is. Ez az eljárás segítheti a beszédoktató programok hatéko-

nyabb használatát, pontosabb, illetve összetettebb visszajelzést adhat a tanulóknak.

A 6. táblázatban látható, hogy az időben lineáris, illetve az időben nemlineáris vetemítés esetén, egy referenciamondattól mekkora a prozódiai távolság az egyes jellemzők alapján (alaphang, intenzitás), ha a bemondó törekszik arra, hogy „jól" mondja vissza a mondatokat úgy, hogy az első esetben időszerkezetileg is hasonlóan mondja be, míg a második esetben erre nem törekszik.

6. táblázat: Az alaphang, illetve az intenzitás átlagos négyzetes távolsága a referenciától lineáris és dinamikus vetemítés esetén

	Azonos tempóban bemondott	**Más tempóban bemondott**
Alaphang:		
Lineáris vetemítés esetén	13,5 Hz	**22,1 Hz**
Dinamikus vetemítés esetén	10,1 Hz	**11,5 Hz**
Intenzitás:		
Lineáris vetemítés esetén	7,2 dB	**11,4 dB**
Dinamikus vetemítés esetén	5,5 dB	**6,8 dB**

A 6. táblázatból leolvasható, hogy a nem megfelelő vetemítés nagyban megnöveli az átlagos négyzetes távolságot. A 6. táblázat, a 6. és a 7. ábra alapján jól látható, hogy az időben nemlineáris vetemítés jobb eredményt ad vissza, mint a lineáris. Az is látható, hogy az időben nemlineáris vetemítés szükséges a helyes összehasonlításhoz, elengedhetetlen a korrekt vizuális megjelenítéshez. Ezek alapján kijelentjük, hogy egy időben nemlineáris vetemítő eljárást használó tanító program interaktívabb, magasabb szintű tanulási élményt nyújthat a tanuló számára. Részletesebb, pontosabb visszajelzéseket adhat, ami a lineáris vetemítés esetén nem volna lehetséges. És mindez több nyelv esetén is megvalósítható ugyanazzal az eljárással.

Irodalom

Baker, Janet – Deng, Li – Glass, James – Khudanpur, Sanjeev – Lee, Chin Hui – Morgan, Nelson 2009a. Updated MINDS Report on Speech Recognition and Understanding, Part I. *IEEE Signal Processing Magazine* 26/3. 75–80. Part II. *IEEE Signal Processing Magazine* 26/4. 76–85.

Benno, Peters 2005. *The Kiel Corpus of spontaneous speech.* http://www.ipds.uni-kiel.de/kjk/pub_exx/aipuk35a/aipuk35a_1.pdf. (A letöltés ideje: 2013.január.12.)

Boersma, Paul – Weenink, David 2001. Praat, a system for doing phonetics by computer. *Glot International.* 341–345.

Garofolo, John S. – Lamel, Lori – Fisher, William – Fiscus, Jonathan – Pallet, David – Dahlgren, Nancy – Zue, Victor 1993. *TIMIT Acoustic-Phonetic Continuous Speech Corpus.* Linguistic Data Consortium, Philadelphia.

Huang, Xuedong – Deng, Li 2010. Overview of modern speech recognition. In Indurkhia, Nitin – Damerau. Fred (eds.): *Handbook of natural language processing.* CRC Press Boca Raton, London–New York.

Mihajlik Péter 2010. Rejtett Markov-modellek. In Németh Géza – Olaszy Gábor (szerk.): *A magyar beszéd. Beszédkutatás, beszédtechnológia, beszédinformációs rendszerek.* Budapest, Akadémiai Kiadó. 242–243.

Olaszy Gábor 2010. A beszéd szupraszegmentális szerkezete. In Németh Géza – Olaszy Gábor (szerk.): *A magyar beszéd. Beszédkutatás, beszédtechnológia, beszédinformációs rendszerek.* Akadémiai Kiadó, Budapest. 171–204.

Szaszák György – Nagy Katalin – Sztahó Dávid – Vicsi Klára 2009. Automatikus intonációs osztályozó felhasználása hallássérültek beszédterápiájában. In Tanács Attila – Szauter Dóra – Vincze Veronika (szerk.): *VI. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY2009).* Szeged. 381–385.

Sztahó Dávid – Nagy Katalin – Vicsi Klára 2009. Subjective tests and automatic sentence modality recognition with recordings of speech impaired children. In Esposito, Anna – Campbell, Nick – Vogel, Carl – Hussein, Amir – Nijholt, Anton (eds.): *Developement of multimodal interfaces: Active listening and synchrony: Second COST 2102 International School Dublin.* Dublin, Ireland. 397–405.

Vicsi Klára – Vig Attila 1998. LIAS: Language Independent Automatic Segmentation technique using SAMPA labeling of phonemes. In: *Proceedings on First International Conference on Language Resources & Education.* Granada. 1317–1323.

Vicsi Klára – Kocsor András – Teleki Csaba – Tóth László 2004. Beszédadatbázis irodai számítógép-felhasználói környezetben. In Alexin Zoltán – Csendes Dóra (szerk.): *Second Conference on Hungarian Computational Linguistics (MSZNY 2004).* Szeged. 348–359.

Ez a kutatás az Alap- és alkalmazott kutatások hallássérültek Internetes beszédfejlesztésére és az előrehaladás objektív mérésére TÁMOP-4.2.2.C-11/1/KONV projekt és a „FuturICT.hu" TÁMOP-4.2.2.C-11/1/KONV projekt keretein belül készült.

PRECÍZIÓS, PÁRHUZAMOS, MAGYAR BESZÉDADATBÁZIS FEJLESZTÉSE ÉS SZOLGÁLTATÁSAI

Olaszy Gábor

Bevezetés

A beszédkutatásban világszerte egyre nagyobb teret kapnak az előre elkészített, annotált és szegmentált beszédadatbázisok. A gyarapodás oka, hogy a kutatásokban és a fejlesztésekben egyre inkább statisztikai eljárásokkal vizsgálják a beszédet, azaz nagy adattömeg vizsgálatával határozzák meg a beszédre jellemző paramétereket, azok változási tendenciáit. Mivel egy ilyen adatbázis nagyszámú adatot tartalmaz, az annotálását és szegmentálását csak gépi eljárások támogatásával lehet elvégezni. Ebből következik, hogy az adathalmazban lesznek hibás adatok is. A kutató számára viszont az lenne a kívánatos, hogy gyakorlatilag hibamentes adathalmazra építhesse a vizsgálatait. Az ideális megoldás tehát az lenne, ha az annotálási és címkézési hibák számát a minimumra lehetne csökkenteni a beszédadatbázisokban.

A jelen tanulmányban olyan beszédadatbázist mutatunk be, amely az ideális esetet közelíti, mivel a gépi hibázásokat félautomatikus támogatással feltártuk és manuális javítással megszüntettük. Ilyen részletességgel és pontossággal feldolgozott beszédadatbázis korábban még nem készült magyar nyelvre. Számos beszédadatbázist készítettek már az elmúlt évtizedekben Magyarországon, mindegyiket más-más céllal (Gósy et al. 2012).

Kiemelendő a legutóbbi fejlesztés, amely az MTA Nyelvtudományi Intézetében folyik, egy BEszélt nyelvi Adatbázis (BEA) létrehozása. Itt a kitűzött cél, hogy 500 beszélőtől gyűjtsenek felolvasott és spontán beszédet is. Jelenleg a fejlesztés a felénél tart, de már most is látszik, hogy számos fonetikai kutatás alapját képezi (Gósy szerk. 2012) a hatalmas adattár. A BEA hanghullámot és szövegszintű átiratot tartalmaz.

Anyag, módszer, beszélő személyek

A most ismertetett beszédadatbázis nyelvi anyagának gerincét egy korábbi kutatáshoz gyűjtött szöveges mondatkorpusz képezi (Vicsi–Vígh 1998), amely a BABEL projekt keretében készült el (1992 mondat, 20873 szó, 85424 hang, 33783 magánhangzó), és a magyar nyelvet jól reprezentálja hangstatisztikailag. Ezt egészítettük ki célzottan, egyrészről különböző hosszúságú és típusú kérdő mondatokkal (*Ő? Én? Még? Baba? Babával?* stb.), valamint rövid kijelentő mondatokkal az egyhangos mondattól kezdve a há-

rom szótagos egyszavas mondatig. Ez a szöveges anyag 522 mondatot, 1985 szót, 5534 hangot, ezen belül 1844 magánhangzót tartalmaz. A kibővítéssel az volt a célunk, hogy időszerkezetileg is, és a beszéddallam vonatkozásában is minél széleskörűbben reprezentáljuk a magyar beszéd sajátosságait. Mindkét szöveges anyagot 10 beszélő, 5 nő (NO) és 5 férfi (FF) olvasta fel, életkoruk 30–65 év között oszlik el, mindannyian Budapestiek, a köznapi magyart beszélik. Foglalkozásuk: tanár, színész, adminisztrátor, kutató, mérnök, zenész, énekművész. A teljes beszédadatbázis tartalma mintegy 900 000 beszédhang, ebből közel 350 000 magánhangzó. A beszédadatbázis rövidített neve PPBA. A hangfelvételeket egységesen a BME Híradástechnikai Tanszék Rezgésakusztikai Laboratóriumának professzionális hangstúdiójában készítettük 44,1 kHz-es mintavételezéssel és 16 bites lineáris kvantálással. A hangrögzítés közvetlenül számítógépre történt. A hangfelvételeket annotáltuk és szegmentáltuk saját fejlesztésű kényszerített gépi beszédfelismerő programmal (Mihajlik et al. 2002). Ezután, szintén saját fejlesztésű hibadetektálóval (Olaszy–Bartalis 2008) ellenőriztük a kapott adatokat és a hibás hangkódokat, illetve hanghatárokat kézzel, vizuális és auditív ellenőrzés segítségével korrigáltuk. Az ellenőrzéshez és javításhoz a Praat program megjelenítési algoritmusait használtuk (Boersma–Weenink 2005). A kombinált annotálási és szegmentálási módszer végeredménye, hogy a beszédadatbázis pontos adatokat tartalmaz (ezért precíziós), tehát referenciaként használható. A hanghullám akusztikai tartalma és a hozzá rendelt hangszimbólumok és címkék pontosan megfelelnek egymásnak.

Az annotálás háromféle jelölési csoportot foglal magába: beszédhangszimbólumok, nem beszédhez tartozó részek jelölése „sil" jellel (szünet, krákogás, levegővétel stb.), ejtési eltérés jelölése (például glottalizáció) a hangszimbólum mellé tett csukó zárójellel. A hangjelöléseket mind SAMPA, mind TMIT jelöléssel (Németh–Olaszy 2010: 77) megadtuk. A beszédhullámmal párhuzamos szegmentálás (címkézés) feldolgozási szintjei a következők: zöngés hangperiódusok (pitch marks: PM) egyenkénti jelölése, hanghatárok és a „sil" jellel jelölt szakaszok határai, valamint szóhatárjelölések. Az annotálásra és címkézésre támaszkodva formánsadatok meghatározását is elvégeztük. Összességében mintegy 4,5 millió formánsadatról van szó. Itt is célunk volt a precíz adatmeghatározás. Ezért első lépésben a Praat programot használtuk, majd a kapott formánsértékeket saját algoritmussal ellenőriztük (Abari–Olaszy 2012), és a hibás értékeket manuálisan javítottuk. A mondat szerkezetéből adódó hangsúlyok jelölése is célunk volt. A beszédatbázis szövegkorpuszának kijelentő mondatait szó szinten bináris (hangsúlyos/nem hangsúlyos) címkékkel láttuk el. A feldolgozáshoz gépi támogatásként a Profivox TTS rendszer hangsúlymeghatározó modulját használtuk, mivel más ilyen szoftver nem áll jelenleg rendelkezésre. A Profivox mintegy 72%-os pontossággal végzi a hangsúlybesorolást (Tamm–Olaszy 2005). Manuálisan is ellenőriztük

az eredményt, és ha kellett, javítottuk. Ezzel létrehoztuk az első magyar hangsúlyjelölési szöveges adattárat.

A PPBA szerkezeti felépítése egységes, minden beszélő adatai ugyanolyan könyvtárszerkezettel rendelkeznek, tehát a párhuzamos adatkeresés és az esetleges összehasonlítás biztosított.

A precíziós feldolgozás és a tisztán gépi címkézés kérdésköre

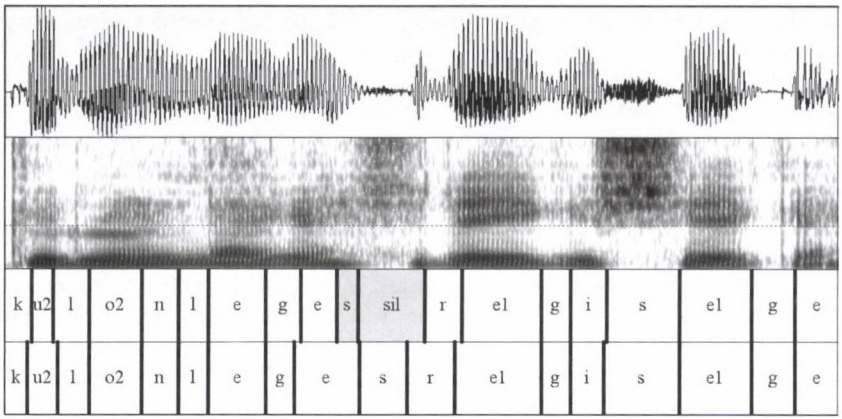
Vannak olyan kutatások, amelyek igénylik a precíz annotálást és címkézést. Ilyenek lehetnek a hangidőtartamokkal és a beszéd egyéb időszerkezeti elemeivel kapcsolatos statisztikai mérések. A beszéd spektrális felépítésével kapcsolatos statisztikai kutatások is ide tartoznak, például a formánsmenetek modellezése, a zárfelpattanások vizsgálata stb. Ha ilyen vizsgálatokat egy beszédadatbázisból származó adathalmazra akar alapozni a kutató, akkor joggal várja el, hogy minden hanghatár címkéje korrekt időpontra legyen bejelölve, valamint azt, hogy a jelölt hangnak megfelelő akusztikai tartalom legyen a bejelölt beszédszakaszon.

A tisztán gépi annotálás egyik hátránya lehet, hogy a mondatkezdő zöngétlen zárhangok virtuális kezdetét a gépi címkézés általában nem jelöli. Ugyanez a helyzet a beszédszüneteket jelölő „sil" címkék utáni ilyen hangokra is. Az ilyen esetekben a hangsorkezdő zöngétlen zárhang kezdési pontját jelentő hanghatárcímkét balra kell mozgatni, hogy a virtuális hangidőtartamot érzékeltessük, vagyis a nem látható néma fázist is a hanghoz jelöljük. Ha nem így van jelölve, akkor például hangidőtartam-méréseknél hibásan mérjük ezeknek a hangoknak az időtartamát (túl rövidre).

A VV, VVV kapcsolatok határát csak vizuális ellenőrzéssel és meghallgatással lehet jó közelítéssel bejelölni.

A fonetikai átíró program is hibázhat, nem azt a hangot jelöli az adott helyen, ami elhangzik a hangsorban (*arccsont, technika, meggyújt*). A gép által automatikusan jelölt „sil" szakaszok esetében a beszédhangokon túli részek (összefoglalva: szünetek) gépi jelölésének kritériumrendszere nem kiforrott. Gyakran kell a kézi ellenőrzés során „sil"-t betenni, kivenni, a „sil"-hez jelölt határ helyét módosítani. A kényszerített felismerő gyakran jelöl a hanghullámban automatikusan olyan „sil"-szakaszt, ami valójában nem szünetet reprezentál, hanem egy hang része (1. ábra). Ennek ellenkezője is előfordult, hogy be kellett iktatni egy sil jelölést.

Emberi mulasztásból adódó hibák is keletkeznek egy nagyméretű beszédadatbázis készítése során. A felolvasó személy mást mond, mint a szöveg. Ez még akkor is előfordul, ha a szöveg teljesen helyesen van írva. A beszélő automatikusan átfogalmazza kissé a leírt mondatot, és más szavakkal mondja azt (például beszúr egy névelőt, és elhagy egy ragot, de a mondat teljesen értelmes marad). Előfordult, hogy egy mondatban a *fogok* helyett *fognak*-ot mondott, egy másikban a *valamilyen* helyett *milyen*-t ejtett a bemondó. Ezt is korrigálni kell.



1. ábra
Példa egy felesleges szünetre (fent) és a kézi javítás utáni helyzetre (lent)

A szöveg is lehet hibás (például kimarad egy betű, elmarad egy rag stb.), de a felolvasó személy automatikusan korrigálja azt az ejtés során. Ilyenkor a kényszerített felismerő program a szöveg karaktereit veszi figyelembe, és akár több hangon keresztül is hamis hanghatárokat jelöl meg, mivel a szöveg által megadott beszédhangra jellemző akusztikai tartalom nem egyezik az ejtett tartalommal. A 2. ábrán látható egy ilyen példa, a karakterhibás szöveg: *felségét;* a felolvasott szó helyesen: *feleségét.* Ezt a hibát például a gépi hibakereső segítő algoritmusunk a feltűnően hosszú hangidőtartam alapján találta gyanúsnak, és a vizuális ellenőrzés során ki is derült, hogy kimaradt egy hangjelölés. Programhiba is előfordult a kényszerített beszédfelismerő működésében. Ilyenkor a helyes ortografikus szöveg fonetikai átírása nem korrekt, és más hangokat tesz a hangsorba a gép, mint amilyenek a szövegből adódnának. A bemondó viszont helyesen olvassa fel a szöveget, tehát a hangazonosítások a spektrális tartalom és a hangjelölés alapján nem lesznek korrektek. Ilyenkor a felismerő csak több hang után tud újra visszatérni a helyes döntésekhez (3. ábra). A példákból látható, hogy az annotálás és a címkézés utólagos ellenőrzése és javítása számos későbbi munkához szükségszerű, és emeli a beszédadatbázis tudományos értékét.

Az adatbázis szolgáltatásai

Az adatbázisban tárolt adatok a következők.

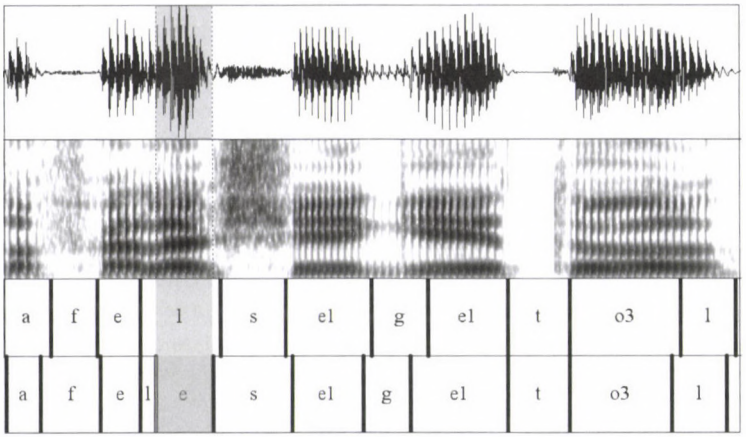
Szöveg: a mondat helyesírással megadott formája xx.txt fájlban.

Fonetikus átirat: a szöveges forma alapján készült el, és a felolvasott mondat (xx.wav fájl) pontos tartalmát rögzíti a négyféle jelölési csoporttal. Az automatikus feldolgozásból eredő annotálási és címkézési adatokat az

xx.TextGrid fájl tartalmazza, a kézi javítással módosított adatokat az xx_man.TextGrid fájl.

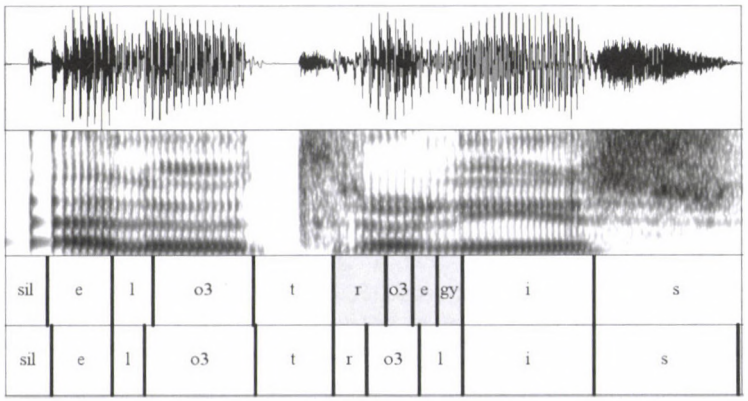
Hangnyomás-idő függvény: a mondat felolvasott formája az xx.wav fájlban.

Zöngeszinkron-jelek: Ennek megfelelően a zöngés szakaszok belsejében kiszámítható az alapfrekvencia, illetve az esetleges irregularitások is kereshetők. Az információt az xx.pit fájl tartalmazza.



2. ábra

A betűkimaradás rossz hanghatárjelölést eredményez (fent) A kézi javítás után áll helyre a helyes hang- és hanghatárjelölés (alsó hangsor)



3. ábra

A kényszerített beszédfelismerő hibázott (felső hangsor). A kézi javítás utáni helyes hangsor és hanghatárok alul láthatók

Hanghatárok: a mondat beszédhangjainak és „sil" jelzéseinek kezdőpontjai (a pontos időkoordinátákkal). Az automatikus feldolgozásból eredő adatokat az xx.TextGrid fájl tartalmazza, a kézi javítással módosított adatokat az xx_man.TextGrid fájl. A hanghatárok pontossága 10 ms-on belüli.

Szóhatárok: A mondat szavainak végét külön címkével jelöljük. Tehát szószintű vizsgálatok is végezhetők (szóhossz, szóhelyzet stb.). Ezt az információt az xx.ssw fájl tartalmazza. A jelölések kezelik azt az esetet is, amikor a szóhatár nem különíthető el hangszinten, hanem egyetlen hang képezi az előző szó végét és a következő szó elejét is.

Formánsfrekvenciák: a magánhangzók F_1, F_2, F_3 formánsadatai minden magánhangzóra meghatározásra kerültek a hangon belüli 5 mérési pontban (10%, 25%, 50%, 75%, 90%). Az 5 pontos reprezentáció lehetővé teszi a jellemző formánsmenetek meghatározását is, például a szűkebb és tágabb hangkörnyezet függvényében. A formánsadatok külön fájlban szerepelnek minden bemondóra vonatkoztatva.

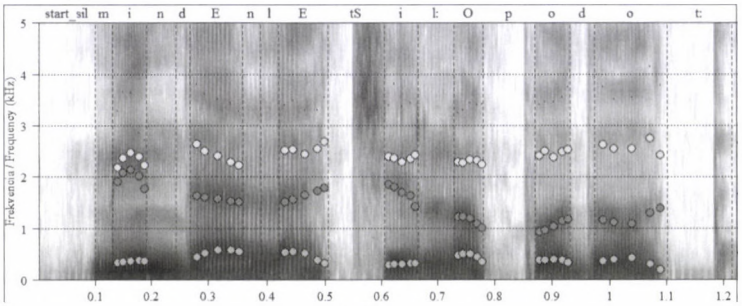
Hangsúlyjelölés: az adatázis szöveganyagát képező mondatkorpusz összes kijelentő mondatában, a mondat minden szavára elméleti hangsúlyozási címkét helyeztünk el bináris formában ([:W]=hangsúlyos/[:N]=nem hangsúlyos). Egy példamondat a feldolgozás bemutatására: [:W]*ehhez* [:N]*jön* [:N]*még* [:N]*a* [:W]*közjogi* [:N]*keret,* [:N]*és* [:N]*esetleg* [:N]*az* [:W]*új* [:W]*alkotmány.* A jelölések pontossága magasfokú, amit úgy kell érteni, hogy nincs benne jelölési hiba, vagyis ahol hangsúlyt jelöltünk, ott a hangsúlyos ejtés nem okoz megértési zavart és fordítva. Vannak olyan mondatok, amelyek többféle hangsúlykiosztással is ejthetők az értelmezés, illetve a közlési szándék szerint. Ezeknél a mondatoknál az egyik helyes formát tartalmazzák a jelölések.

A felsorolt nyolc szolgáltatás adatai szoros szinkronban vannak egymással. Ez biztosítja azt, hogy akár keresztvizsgálatok is végezhetők (formáns-szóhossz; formáns-beszélő; szóhossz-beszélő stb.).

Az adatbázis felhasználása

A PPBA sokrétű adattár, amely hatékonyan felhasználható a beszédkutatásban és a beszédtechnológiai fejlesztéseknél. Precíziós feldolgozottsága biztos alapot jelent a jövő kutatóinak. Nem kell a kutatónak adatgyűjtéssel, előzetes adatfeldolgozással foglalkozni, ezzel a kutatás hatékonysága növelhető.

Az adatbázis már most fontos szerepet tölt be a statisztikai parametrikus beszédszintézis-fejlesztéseknél (Tóth 2013), valmint a magyar beszéd frománsterének újszerű modellezésénél és vizsgálatánál (Abari 2013). Egy külön erre a célra készített szoftver segítségével minden mondat formánsmenetei vizuálisan is tanulmányozhatók (4. ábra).



4. ábra
A *Minden lecsillapodott* ....férfi ejtésű hangsor formánsmenetei a mérési pontok értékeivel megjelenítve

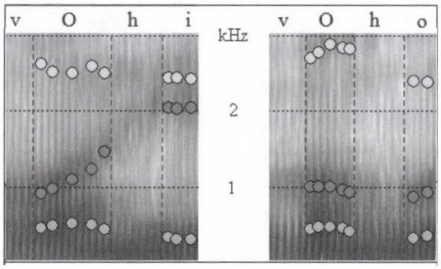
Bármely formánsmeghatározó algoritmus hatásosságának vizsgálatát is alapozni lehet a formánsadatbázisra. A Praatra vonatkozó ilyen méréseket elvégeztük (1. táblázat), és az adatok azt mutatják, hogy beszélőfüggő a hibázás foka. A Praat automatikus mérőrendszere jól teljesített, 94,35%–99,22%-ban korrekten határozta meg a formánsadatokat. Ehhez jelentősen hozzájárult a precíziós címkézés is.

1. táblázat: A formánsjavítások száma %-ban megadva a teljes formánsadatbázisra vonatkoztatva a 10 bemondó szerinti elosztásban

FF1	FF2	FF3	FF4	FF5	NO1	NO2	NO3	NO4	NO5
2,50	5,36	0,78	2,37	1,77	1,35	2,42	5,72	4,47	6,65

Ugyancsak új eredmény, hogy a magánhangzó F_2 formánsának mozgását nemcsak a közvetlenül hozzá kapcsolódó hangok befolyásolhatják, hanem a tőle balra és jobbra elhelyezkedő –2. és +2. hangok is. Ez azt mutatja, hogy az artikulációs mozgások az egyes hangok, a képzési módok és helyek esetében már a közvetlenül csatlakozó hangon túl mutatva befolyásolják a magánhangzó képzését. Például más az [ɔ] F_2-jének a mozgása a [vɔhi] hangsorban (*...szavahihető...*), mint a [vɔho]-ban (*...tolva hozzá...*). Ezt a V utáni +2. hang befolyása okozza (5. ábra). Az előbbiben felfelé mozog az F_2, az utóbbiban lefelé. Ez a jelenség más hangkapcsolatoknál is kimutatható. A formánsadatbázis felhasználásával az ilyen jelenségek szisztematikusan felderíthetők és osztályozhatók. Ez hozzájárulhat a pontosabb formánsmozgások meghatározásához. A teljes rendszerezés után eljuthatunk oda, hogy a szöveg alapján a hangsor formánsmenetei jósolhatók lesznek sematikus formában, ami fontos lehet a beszédtechnológiai algoritmusok finomításához. Egy további felhasználási példa a glottalizált hangzást érinti. Az adatbázisban elhelyezett glottalizált hangra vonatkozó jelölések felhasználásával segítséget

nyújtottunk egy új rendszerű glottalizált hangzást megszüntető javító algoritmus kidolgozásához.



5. ábra
Példa arra, hogy a *...vah...* CVC szekvencia ejtése során a magánhangzó F_2-jének menetét milyen erősen befolyásolja a magánhangzó utáni második [i], illetve [o] beszédhang

Vizsgálatok folytathatók a gépi kényszerített beszédfelismerési algoritmusok hatékonyságának meghatározására is. Összehasonlítottuk a gépi feldolgozás eredményeit és a kézi javítások helyét és számát (xx.TextGrid és xx_man.TextGrid fájlok), ezzel képet kaptunk arról, hogy mely pontokon gyengébb a gépi algoritmus teljesítőképessége, ez függ-e például a beszélő személy hangjától stb. (2. táblázat).

2. táblázat: A kézzel javított hanghatárok száma a 10 beszélőre a kézi eltolások időintervallumaival

	10-19 ms	**20-29 ms**	**30-39 ms**	**40-49 ms**	**50-59 ms**	**< 60ms**
FF1	17128	5315	1657	552	185	169
FF2	9155	1869	760	383	203	262
FF3	23249	6285	1406	467	167	240
FF4	17850	4845	1240	428	198	272
FF5	9042	2491	941	525	331	558
NO1	10851	2450	1005	439	221	318
NO2	11062	2816	1003	432	190	456
NO3	11156	2603	924	435	227	418
NO4	13866	2321	658	228	91	93
NO5	9362	2401	878	442	253	375
Összesen	132721	33396	10472	4331	2066	3161

A táblázatból több dolog kiolvasható. Legtöbbször az egy-két zöngeperiódusra jellemző időtartamsávban kellett a hanghatárt a valós helyére tolni. Látható az is, hogy a gépi beszédfelismerő algoritmus hibázásai függnek a

beszélő hangjától. Sok egyéb más mérés is végezhető lenne, azonban ennek a cikknek a terjedelme ezt nem teszi lehetővé.

Az adatbázis adataira támaszkodva egyedi számítási és keresési algoritmusokat bárki kidolgozhat és alkalmazhat. Ezzel szinte korlátlanul bővíthetjük a kutatási témák sorát. Csak néhány példát adunk: hangidőtartamok vizsgálata; fonológiailag hosszú hangok és kiejtésbeli megvalósulásaik felmérése; szünetek és osztályozásuk; szóhosszak kinyerése különböző szempontok alapján; bemondófüggő vizsgálatok; ritmikai mérések; hangstatisztika más paraméterek függvényében; formánsmenetek tulajdonságainak sokrétű vizsgálata; az elvi és gyakorlati hangsúlyozás közötti összefüggések elemzése; formánsmérő algoritmusok tesztelése; hangrealizációs összehasonlítások; oktatási adattárak kialakítása (hányféle akusztikai megvalósulásban jelennek meg ugyanazon hangok a folyamatos beszédben); betű-hang átalakító algoritmusok pontosságának vizsgálata.

Összefoglalás

A bemutatott beszédadatbázis sokoldalúan előkészített és feldolgozott adatok halmazait tartalmazza. Különlegessége a precíz adatfeldolgozás, ami jó alapot nyújt sokféle kutatáshoz. A PPBA 2012-ben a CESAR projekt részeként bekerült a META-SHARE nemzetközi nyelv- és beszédtechnológiai adatbázistárba, így része lett a nemzetközi kutatási beszédadatbázisoknak.

Irodalom

Abari Kálmán – Olaszy Gábor 2012. Új módszer formánsadatok meghatározására és formánsadatbázis létrehozására nagyméretű beszédadatbázisokhoz. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest. 251–268.

Abari Kálmán 2013. *A formánsmozgások statisztikai vizsgálata és modellezése a magyar magánhangzókban*. Doktori disszertáció. DE, Debrecen.

Boersma, Paul – Weenink, David 2005. *Doing phonetics by computer*. [Computer software]. www.praat.org

Gósy Mária – Gyarmathy Dorottya – Horváth Viktória – Gráczi Tekla Etelka – Beke András – Neuberger Tilda – Nikléczy Péter 2012. BEA: beszélt nyelvi adatbázis. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest. 9–24.

Gósy Mária (szerk.) 2012. *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest.

Mihajlik, Péter – Révész, Tamás – Tatai, Péter 2002. Phonetic transcription in automatic speech recognition. *Acta Linguistica Hungarica* 49/3–4. 407–425.

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

Olaszy Gábor – Bartalis Mátyás 2008. Jelfeldolgozási és fonetikai algoritmusok kombinációja a gépi hanghatárjelölés javítására. *Beszédkutatás 2008*. 208–220.

Tamm Anne – Olaszy Gábor 2005. Kísérlet automatizált szövegelemzési módszerek kialakítására a szóhangsúlyok meghatározásához. In Alexin Zoltán – Csendes Dóra

(szerk.): *III. Magyar Számítógépes Nyelvészeti Konferencia*. Szegedi Tudományegyetem Informatikai Tanszékcsoport, Szeged. 383–393.
Tóth Bálint 2013. *Rejtett Markov-modell alapú gépi beszédkeltés*. Doktori értekezés. BME, Budapest.

A szerző köszönetet mond Laczkó Klárának az adatrendezés korrekt elvégzéséért, Bartalis Mátyásnak a kényszerített gépi felismerésben nyújtott segítségéért, valamint Tóth Bálintnak a hanghatárok szegmentálási adatain végzett összehasonlító mérés elvégzéséért. Mindhárman a BME TMIT munkatársai. Köszönet illeti Abari Kálmánt is, a Debreceni Egyetem kutatóját, aki a formánsok meghatározásában és a javításukhoz szükséges interaktív szoftvertámogatás készítésében végzett komoly munkát. A kutatást támogatta a CESAR project (Grant No. 271022).

A FOGSZABÁLYOZÓ KÉSZÜLÉKEK HATÁSA A HANGKÉPZÉSRE

Götz Gergely

Bevezetés

A gégének és a toldalékcsőnek a beszédprodukcióban kettős szerepe van: hangforrás és hangmódosítás. A tüdőből kiáramló levegő áthaladva a hangszalagokon (vagy megrezegtetve azokat, vagy nem) a szájüregben valamilyen akadályba ütközik, ennek következtében valamiféle zörej keletkezik. E zöngés vagy zöngétlen zörejek jellemzik a mássalhangzókat (pl. Gósy 2004). Az állcsontok és a fogazat típusa, mérete, intaktságának foka, a fogsorok záródásának milyensége jelentős hatással van az artikulációra (Vallino et al. 1993). A kiejtés szempontjából súlyos és meghatározó változásokat okoz a fogak elvesztése, azaz a részleges vagy teljes foghiány. Kevésbé súlyos, ugyanakkor a beszélő szempontjából zavaró problémák adódhatnak a fogszabályozási eltérések miatt is (a fogmedernyúlvány, illetve a fogsorok deformitása). Ilyenkor is zavar keletkezhet az ejtés során, korlátozottan működtethető a tipikus artikulációs gesztus; az ajkak helyzete, mozgása megváltozik, mindez oda vezet, hogy a mássalhangzók egy részének képzése zavart szenvedhet (Fábián et al. 1997).

A szájüreg, mint rezonátorüreg, térfogatának, alakjának megváltozásával befolyásolja a gégében keletkező zönge hangszínét, illetve a szájüregi akadálytól függően módosíthatja a zörejhangok minőségét. A fogakat érintő súlyosabb deformitás esetén megváltozik a szájüreg alakja, köbtartalma, mert módosul a kemény és a lágy részek aránya, helyzete, valamint a szájüreg térfogatát legjobban befolyásoló nyelv és ajkak helyzete, formája és funkciója (Mason 2011). Emiatt nemcsak a zöngés mássalhangzók, hanem valamennyi beszédhang hangszíne megváltozhat. A beszéd megszokott, tudat alatti, reflektorikus mozgáselemei már nem elegendőek a konvencionális hangképzéséhez.

Kiejtési zavarokat okozó szájüregi elváltozások nem feltétlenül kell, hogy látványra is drasztikus, keményszövet-hiánnyal járjanak. Ugyanakkor komoly deformitások is sikeresen, funkcionálisan is rehabilitálhatók modern fogszabályozási eljárással. Hangsúlyozandó, hogy nemcsak a keményszövetek eltérései, hanem a lágyrészek anatómiai rendellenességei, illetve diszfunkciói is módosíthatják a hangképzést. Ilyen típusú eltérések például az alsó és felső ajakfék lenövése, amelynek fennállása esetén a bilabiális és a labiodentális hangok képzése szenvedhet zavart. Hasonló helyzet alakulhat ki a nyelvfék

túlzott rövidülése, lenövése miatt, ilyenkor inkább a pergőhangok ejtése nehezített. Mindkét lágyrészelváltozás megoldása könnyű, mivel kisebb fajta szájsebészeti beavatkozással megoldható. A felső ajakfék túlzott megvastagodása miatt kialakuló (több milliméteres) rés a két felső nagymetszőfog között (1. ábra) általában kedvezőtlenül befolyásolja a labiodentálisok, az alveolárisok és a posztalveolárisok artikulációját. A fogazati eltérést korrigálva az izomműködés és a lágyrészek pozicionálása is spontán rendeződik. Az esetek nagy részében, ami az emberi szem számára szép és természetes, az funkcionálisan is jó, ez igazolható a fogszabályozásban elért eredményekkel is.



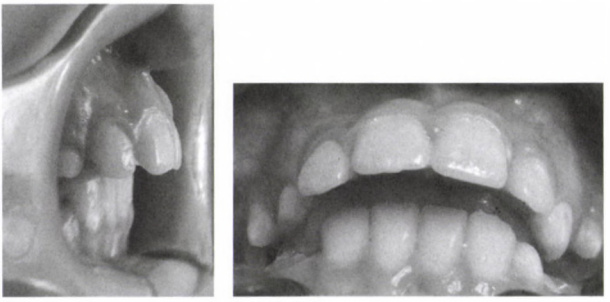
1. ábra
Réses fogazat fogszabályzó kezelés előtt (bal oldal) és után (jobb oldal)

Genetikai tényezők is generálhatnak réses fogazatot (széles fogmedernyúlvány és ugyanakkor kisméretű fogak öröklése, avagy akár többszörös maradó fog csírahiány); ezek megoldása is lehet fogszabályzó kezelés, azaz fogszabályzó készülékkel össze lehet zárni a réseket (Surján et al. 1982). A fogmedernyúlványoknak a fogakhoz, a másik állcsonthoz, illetve a nyelv kiterjedéséhez viszonyított mérete szintén befolyásolhatja az artikulációt. Ha túl nagy a méretkülönbség az alsó és a felső állcsont között (bármelyiknek a javára), akkor állandóan meglevő, jelentős rés lesz köztük (2. ábra), melynek megszüntetése a helyes hangképzés elérése céljából igen nehézkes (a leggyakoribb beszédhangképzési zavar ilyen esetekben a diszlália, diszglosszia, vö. Laine 1992). Az állcsontok egymáshoz képest nem megfelelő pozíciója is okozhat nehezen kompenzálható rést az alsó és a felső fogsor között, ami hasonló artikulációs zavart okozhat.

A túl szűk felső állcsontnak egyéb további, az artikulációs gesztusokat befolyásoló következményei lehetnek (3. ábra). Mivel a felső állcsont nemcsak a szájüreg csontos teteje, hanem az orrüreg alapja, a nem megfelelően tág, kemény szájpad és a szűk fogmedernyúlvány potencírozza az esetleges orrsövényferdülés hátrányos, a zöngeképzésre és a hangszínezetre gyakorolt hatását (orrhangzós beszéd). Ez a rendellenesség egyúttal nehezíti az orrlégzést is.

Jellegzetesek az ajak- és szájpadhasadékos emberek artikulációs zavarai. Náluk az orrhangzós beszéd – a többszörös korrekciós műtétek okozta erősen heges lágyszájpad elégtelen működése és a nem megfelelően növekvő, rendszerint alulfejlett felső állcsont – következtében jelentkezik. Hasadékos embereknél a többszörösen heges garatba torkollik bele a hallókürt (tuba auditi-

va) is, ami rendszerint szintén beszűkült, és gyakori középfülgyulladást és ezáltal halláskárosodást is okozhat. Éppen a percepciós visszacsatolás hiánya miatt fordulhatnak elő náluk beszédészlelési zavarok, jellemző lehet a szűk szókincs és a hajlam az agrammatizmusra (Surján et al. 1982). Mindezek másodlagosan kihathatnak az artikulációs gesztusok kivitelezésére.



2. ábra
Előrefelé jelentős rés van az alsó és felső fogsor között



3. ábra
Extrém felső fogívszűkület nazális színezetű ejtéshez vezethet

A fogmedernyúlvány és az állcsontok szerzett módon is deformálódhatnak, különféle balesetek vagy akár rossz szokások következtében. Az utóbbiak közé tartoznak például az ujjszopás, a cumizás, a nyelvlökéses nyelés, az ajakszorításos nyelés, a kizárólagos szájlégzés. Az említettek miatt leggyakrabban ún. nyitott harapás alakul ki, teljesen csukott száj esetén sem feltétlenül érintkeznek az alsó és felső fogcsoportok. A súlyosságtól függően ezeket a beszélőket erős szigmatizmus jellemezheti, mivel az alveoláris, posztalveoláris réshangok és affrikáták képzésekor a nyelvhegy, illetve a nyelv elülső része nem a megfelelő helyen, hanem valamivel előrébb tolódik. Megszokják ugyanis, hogy a fogsorok záródási hiányát nyelés közben a nyelvhegynek a résbe nyomásával korrigálják, a nyelvhegyet a végfázisban erős horizontális mozgás jellemzi a vertikális-horizontális helyett. Cumizás, ujjszopás esetén a kemény szájpad „magassága" is változhat, azaz ún. gótikus szájpad alakulhat ki, amely ugyan vertikálisan kiterjedtebb, de horizontálisan szűkebb az átla-

gosnál, csökkentve a nyelvteret, mely nem megfelelő nyelvmozgások kialakulásához vezet egyes beszédhangok képzésekor (Moyers 1973). Mindezen eltéréseket időben (gyermekkor, fiatal tinédzserkor) felfedezve, logopédus segítségével és fogszabályzó kezeléssel együttesen tökéletesen helyre lehet hozni, beleértve a rágás- és artikulációs funkciókat is (Bloomer 1963).

Azok a jellemző hangképzőszervi állások, amelyek az egyes hangok létrehozásához nélkülözhetetlenek, valamint azok a mozgások, amelyek az egyik beszélőszervi állásból a másikba való átmenethez szükségesek, az élet folyamán megszokottá válnak, beidegződnek (l. neuromuszkuláris egyensúly). Bármely alkotóelem változása, változtatása hangképzési módosulásokhoz, változásokhoz vezet. Bizonyos értelemben beszédzavarhoz vezethet az egyes fogászati beavatkozások következtében megváltozott szájmiliő, például a (fogak külső oldalára) rögzített fogszabályzó felragasztása miatt zavart szenvedhet a bilabiális hangok képzése (Farronato et al. 2012). A módosulások, változtatások mértékének és az adott egyén szájképletei adaptív képességének, illetve az egyén akaraterejének következtében pillanatnyi, átmeneti vagy akár hosszú időre rögzült hangképzési zavar jöhet létre. A legkisebb művi beavatkozás is – ha olyan helyen kell elvégezni – a beszédet hallgató számára is észlelhető artikulációs eltérést eredményez. A kivehető fogszabályzók behelyezése után hallható kiejtési különbségek gyakoribbak. Ez nem csupán a rés vagy zár képzését érinti; például az alsó állcsont fogmedrének, fogsorának kivehető készülékes kezelésekor a nyelv oldala „elmenekül" a fogszabályzó készülék műanyag felületétől, és ez az állapot hosszú távon neuromuszkulárisan is rögzülhet. Ez az ún. elkent beszéd a fogszabályzó készülék hordási idején túl is megmaradhat. Mindezeket fontos időben felismerni, mert a jól időzített multidiszciplináris terápia (fogszabályozás, logopédia, foniátria) sok problémától kímélheti meg a beteget.

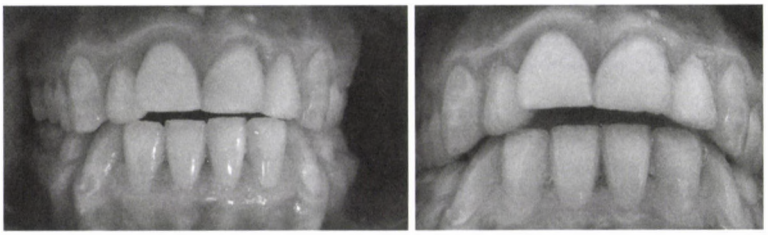
Artikulációs zavart okozhat a kezdeti időszakban a rögzített fogszabályozó készülék is, bár ez nagymértékben függ attól, hogy a fogszabályzó a fogak külső vagy belső felszínére kerül, és eszerint a labiális, a labiodentális vagy az alveoláris, posztalveoláris hangok képzését befolyásolja. Az esetek túlnyomó többségében az ilyen hátterű kiejtési zavar átmeneti (a külső felszínre ragasztottnál egy hét, a belsőnél két-három hét).

Úgy tűnik, hogy hangképzési szempontból a legnagyobb adaptivitást a kivehető fogszabályzók igénylik a használójuktól, mivel az ilyen típusú készülékkel nemcsak a fogakat, de a fogmedernyúlványt is lefedjük a műanyag fogszabályzó rész segítségével. Ez egyfelől nagyon hasznos (stabilitás éjszakai hordás esetén is), másfelől éppen emiatt nehéz megszokni, különösen ha sohasem volt az adott terület mesterségesen lefedve. A fogszabályzó kezelés során létrehozott túlzott mértékű ún. harapásemelés (összezáráskor az alsó és felső állcsont egymáshoz viszonyított vertikális távolsága nő) az új harapási magasság megszokásáig szintén problémát okozhat (Fábián et al. 1997).

„Fonetikatudatos" fogszabályozás

Ismerve a fogszabályozási eltéréseket és a hozzájuk gyakran társuló hangképzési hibákat (Khinda et al. 1999), célzottan lehet olyan fogszabályzó készüléket választani, amelynek artikulációt nehezítő mellékhatása kifejezetten segíti az adott eltérés fogszabályozási és neuromuszkuláris korrekcióját (Yashiro et al. 1999).

A 4. ábrán egy olyan beteg szájfotója látható, akinek elülső nyitott harapása nyelvlökéses nyeléssel társul. Az alsó és a felső fogsor között levő vertikális rés miatt a beszélő interdentális selypességgel képezi a sziszegő réshangokat és az affrikátákat (Lebrun 1985). Ezekben az esetekben nehéz diagnosztizálni, hogy melyik probléma keletkezett előbb – azaz a nyitott harapás miatt alakult ki nyelvlökéses nyelés, vagy a nyelvlökés miatt deformálódott a fogazat, és feljlődött ki a nyitott harapás. Egy, a nyelv felőli oldalra felragasztott, rögzített fogszabályozó készüléket helyeztünk fel a betegnek (5. ábra).



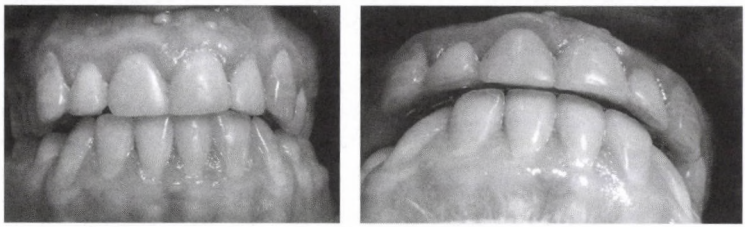
4. ábra
Elülső nyitottharapás jelentős réssel



5. ábra
A felső fogív belső (nyelv felőli) felületére ragasztott készülék

Az ilyen típusú készülék nehezíti egy kicsit az artikulációt, mivel csökkenti a nyelvteret, és az eddig előre benyomuló nyelvet dorzális irányba tereli. A készülék eleve átprogramozza a nyelv nyugalmi helyzetét és mozgástartományát, mivel kismértékben irritálja a nyelv elülső részét, így a páciensnek nincs is kedve előre pozicionálni a nyelvét. Ennek megfelelően a mássalhangzók képzési helye átmenetileg glottális irányba mozdul, és az erőtlenebb nyelvzár miatt még a zöngésség fenntartása is átmenetileg gyengébbé válik. A fogszabályozó készülék segítségével sikerült a metszőfogak területén levő

rést megszüntetni, ezáltal nemcsak esztétikailag, hanem az ejtés szempontjából is kedvező eredményt elérni (6. ábra).



6. ábra
A korrigált frontfogharapás képe a fogszabályzó levétele előtt

Az interdentalis képzés megszűnésével a sziszegő réshangok és affrikáták képzése jóval tisztábbá válik, rövidebb képzési idő jellemzi őket. A posztalveoláris mássalhangzók zörejgóca magasabb frekvenciasávba tolódik és határozottabbá válik, mint a fogszabályozás előtti ejtéskor.

Összefoglalás

A fogazat, illetve a fogsorzáródás módja, minősége hatással van a beszédhangok frekvencia- és időviszonyainak változására, a labiodentális, alveoláris és posztalveoláris mássalhangzók zörejgócainak frekvencia- és intenzitásszerkezetére. Ezért a fogszabályzó készülékes beavatkozások és a fogszabályzók tervezése alkalmával nemcsak a rágófunkció helyreállítását és az esztétikum megtartását, hanem a zavartalan artikuláció intaktságát is érdemes szem előtt tartani. A megfelelő fogszabályzó készülék alkalmazásával nemcsak a fogszabályozási célokat, de a tökéletesebb beszédprodukciót is könnyebb elérni. A fogszabályzó készülékek szájon belüli elhelyezkedésének, hangképzésben betöltött esetleges mellékhatásainak és az egyes beszédhangok képzési helyének együttes ismerete megkönnyíti az adott eltérés korrekcióját, hiszen már maga a fogszabályzó pozíciója is kedvezően tudja befolyásolni beszédprodukciót. A különféle fogszabályzó készülékek és az egyes beszédhangok képzésének kapcsolatát érdemes alaposabban vizsgálni több adatközlő segítségével. Ugyanakkor nehéz különféle adatközlőket összehasonlítani, mivel nincs két egyforma eset, illetve nem egyforma a fogszabályozási megoldásuk, nincs megfelelő közös halmaz sem az ejtés objektív összevetésére. A viszonyítási alap mindig a beszélő kiindulási artikulációs gesztusának akusztikai következménye lehet a fogszabályzó kezelés megkezdése előtt. Ehhez kell majd hasonlítani az egyes fogszabályozó készülékek következményeként kialakuló átmeneti ejtésmódosulásokat, esetleges beszédhibákat, valamint a kezelés végén, a készülék levétele utáni, az esetek nagy részében kedvezőbb artikulációs állapotot. Ezekről az eredményekről egy későbbi tanulmány számol be.

Irodalom

Bloomer, Harlan 1963. Speech defects in relation to orthodontics. *American Journal of Orthodontics* 49: 920–929.

Fábián Tibor – Götz György – Kaán Miklós – Szabó Imre 1997. *A fogpótlástan alapjai*. Semmelweis Egyetem, Budapest.

Farronato, Giampietro – Giannini, Lucia – Riva, Rinaldo – Galbiati, Giovanni – Maspero, Cinzia 2012. Correlations between malocclusions and dyslalias. *Eur. J. Paediatr. Dent.* 13. 13–18.

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

Khinda, Vrethala – Grewal, Nanda 1999. Relationship of tongue-thrust swallowing and anterior open bite with articulation disorders: a clinical study. *Journal of Indian Society of Pedodontics and Preventive Dentistry* 17. 33–39.

Laine, Tellervo 1992. Malocclusion traits and articulatory components of speech. *European Journal of Orthodontics* 14. 302–309.

Lebrun, Yvan 1985. Tongue thrust, tongue tip position at rest, and sigmatism: a review. *Journal of Communication Disorders* 18. 305–312.

Mason, Robert M. 2011. Myths that persist about orofacial myology. *International Journal of Orofacial Myology* 37. 26–38.

Moyers, E. Robert 1973. *Handbook of orthodontics*. Year Book Medical Publishers, Chicago.

Surján László – Frint Tibor (szerk.) 1982. *A hangképzés és zavarai, beszédzavarok*. Medicina Könyvkiadó, Budapest.

Vallino, Linda D. – Tompson, Bryan 1993. Perceptual characteristics of consonant errors associated with malocclusion. *Journal of Oral Maxillofacial Surgery* 51. 850–856.

Yashiro, Kohtaro – Takada, Kamura 1999. Tongue muscle activity after orthodontic treatment of anterior open bite: a case report. *American Journal of Orthodontics and Dentofacial Orthopedics* 115. 660–666.

A PEDAGÓGUSJELÖLTEK KAZINCZY-VERSENYE

A felsőoktatási intézmények Kazinczy Ferenc Szép magyar beszéd versenyének 40. országos döntője

A felsőoktatási intézmények Kazinczy Ferencről elnevezett Szép magyar beszéd versenyének 40. országos döntőjét 2012. március 30-án a Károli Gáspár Református Egyetem Bölcsészettudományi Karának Magyar Nyelvtudományi Tanszéke rendezte. A verseny szervezői a Nemzeti Erőforrás Minisztérium, a Kazinczy-díj Alapítványi Bizottság, az ELTE Fonetikai Tanszéke és a Károli Gáspár Református Egyetem Magyar Nyelvtudományi Tanszéke voltak. A verseny célja az anyanyelvi nevelés és a beszédkultúra fejlesztése. A szövegolvasó verseny megrendezésével arra szeretnék ösztönözni a résztvevőket, hogy növekedjék a magyar nyelv kiejtési szabályait figyelembe vevő, kifejező beszéd iránti igényesség.

Péchy Blanka (1894–1988) színművésznő Kazinczy Ferencről elnevezett díjat alapított 1963-ban olyan hivatásos beszélők (színészek, énekesek, a média szereplői), valamint pedagógusok és az ifjúság számára, akik példamutatóan szép nyelvhasználatuk (elsősorban kiejtésük) révén sokat tesznek a magyar nyelv ápolásának ügyéért. A Péchy Blanka halálát követően létrehozott *Kazinczy–díj Alapítvány – Péchy Blanka Emlékére* évente ad át *Kazinczy-érmet* a pedagógusjelöltek számára rendezett szépkiejtési verseny győzteseinek.

Az országos döntőn a hazai és a határainkon túli felsőoktatási intézmények azon hallgatói vettek részt, akik a helyi versenyeken a legjobbaknak bizonyultak, összesen negyvenegyen.

A versenyt dr. Sepsi Enikő, a Károli Gáspár Református Egyetem Bölcsészettudományi Karának dékánja nyitotta meg. A résztvevőket köszöntötte dr. Gósy Mária, az ELTE Fonetikai Tanszékének tanszékvezető egyetemi tanára, a bírálóbizottság elnöke és dr. Fóris Ágota, a Károli Gáspár Református Egyetem Magyar Nyelvtudományi Tanszékének vezetője. Tudományos előadást tartott dr. Markó Alexandra (ELTE Fonetika Tanszék) „Hangsúlyok a felolvasásban és a spontán beszédben" címmel. Ezután következett a szabadon választott szövegek bemutatása. A délutáni szekcióban a kötelező szöveget olvasták fel a versenyzők. A zsűritől kapott pontszámok összesítése után az ünnepélyes eredményhirdetés következett, majd Gósy Mária professzorasszony zárta a versenyt.

2012-ben az első tizenegy versenyző kapott Kazinczy-érmet, név szerint a következők: Lénárd Valentina (KRE TFK, Nagykőrös), Hajdú Fanny (ELTE

BTK, Budapest), Tomcsányi Árpád (PPKE BTK, Piliscsaba), Lakos Boglárka (Corvinus Egyetem, Budapest), Szabó Adrienn (ELTE TÓFK, Budapest), Szabó Ágnes (ELTE BTK, Budapest), Simon Viktória (KRE BTK, Budapest), Belágyi Hanna (NyME Benedek Elek Pedagógiai Kar, Sopron), Báthory Kinga (Eszterházy Károly Főiskola, Eger), Pólyi Izabella (Konstantin Filozófus Egyetem, Nyitra), Kovács Edit (Apor Vilmos Katolikus Főiskola, Vác).

A bírálóbizottság tagjai voltak: dr. Gósy Mária, elnök (ELTE BTK, MTA Nyelvtudományi Intézet), Barra Mária (Barra Kommunikációs Intézet), dr. Bóna Judit (ELTE BTK), dr. Havas Judit (Petőfi Irodalmi Múzeum), dr. Markó Alexandra (ELTE BTK), dr. Merényi Hajnalka (ELTE TÓK), dr. Molnár Ildikó (Mozgássérültek Pető András Nevelő- és Nevelőképző Intézete), Nábrádiné dr. Jurasek Ibolya (PPKE, Vitéz János Kar), Sólyom Réka (KRE BTK).

A Károli Gáspár Református Egyetem négy versenyzőt indíthatott, közülük ketten nyertek érmet: Lénárd Valentina a nagykőrösi Tanítóképző Főiskolai karról, felkészítője dr. Madarász Imréné, a Bölcsészettudományi Karról pedig Simon Viktória, felkészítője Sólyom Réka.

Részletes információk: www.kre.hu/nyelveszet.

Fóris Ágota

A MAGYAR ÉS A NEMZETKÖZI TUDOMÁNYOS ÉLET ESEMÉNYEI

Beszédtudomány. Az anyanyelv-elsajátítástól a zöngekezdési időig Szerkesztette Markó Alexandra. ELTE Bölcsészettudományi Kar–MTA Nyelvtudományi Intézet, Budapest, 2012. 398 lap.

A recenzens számára is megtiszteltetés olyan tudományos munkáról beszámolni, amely nemcsak minősége és sokszínűsége miatt érdekes, de abból a szempontból is, hogy Gósy Mária professzor munkássága előtt tiszteleg.

A tanulmánykötet a beszédtudomány hazai kutatásaiból válogatva, aktív kutatók írásait gyűjti össze a legkülönfélébb témakörökben. A cikkek átgondolt, alapos munkáról tanúskodó tanulmányok. A szövegek válogatásakor a szerkesztő és a szerkesztőbizottság célja nem egy adott tudományterület vagy részterület minél teljesebb ismertetése volt, hanem a szaktudományban aktuális elméleti és módszertani kérdések körüljárása, mindezeket pedig témájukban Gósy Mária munkásságához igazítva – kollégái és tanítványai tollából. Ez a mozaikszerű szerkesztés nemcsak arra jó, hogy megmutassa a beszédtudomány kutatásainak széles palettáját, hanem ennek köszönhetően az érdeklődő kétféle szerkesztési elv szerint is olvastatja a művet: az első felosztásban témák szerinti csoportosításban (*Régi fogalmak új megvilágításban; Életkori sajátosságok a beszédben; A beszédfeldolgozás aspektusai; Zönge, zöngésség, zöngétlenedés; Prozódiai sajátosságok;* valamint *A beszéd információtechnológiai vonatkozásai*), a második felosztás értelmében pedig tudományterületek szerinti csoportosításban (*Fonológia; Fonetika; Pszicholingvisztika; Beszédtechnológia; Retorika, írásjeltan*) – a tájékozódásban a kétféle tartalomjegyzék segít. Így válik a kötet az első pillantásra szétszaladó témák több rétegben felépített struktúrájává, lehetővé téve a tanulmánykötet nem pusztán lineáris elolvasását.

A két tartalomjegyzék után a szerkesztő előszavát, majd a nyelvtudomány két kiemelkedő művelőjének, Kenesei Istvánnak és Nyomárkay Istvánnak a születésnapi köszöntőjét olvashatjuk.

A *Régi fogalmak új megvilágításban* című részben Siptár Péter a degeminációfogalmának újraértelmezése, Varga László pedig a magyar mellékhangsúly léte

és jelentősége mellett érvel. Gyarmathy Dorottya tanulmánya az újraindítások körüli kettősséget igyekszik egyértelművé tenni.

Az *Életkori sajátosságok a beszédben* című fejezetben a korai életszakasz nyelvelsajátítását vizsgáló ritka tanulmányok közül hármat is olvashatunk: Vesna Mildner és Martina Bajzec arról számol be, hogy a horvát anyanyelvű gyermekek iskolaérett korukra nyelvi szempontból egységesülnek: megszűnnek a nemek különbségéből adódó nyelvhasználati eltérések. Deme Andrea az iskola előtt álló fiúk és lányok vokálisainak formánsszerkezetéről közöl precíz és átfogó elemzést. Neuberger Tilda különböző életszakaszokban mérte fel a kommunikáció grammatikai szerkesztettségét, Bóna Judit pedig az idősek kommunikációs sajátosságainak vizsgálatába egy egészen új és érdekes szempontot emel be: a matuzsálemek nyelvhasználatát, amelyben száz évesnél idősebb személyek spontán beszédét elemzi és veti össze fiatalabb idősek beszédével.

A beszédfeldolgozás aspektusai részben Imre Angéla a GMP-teszt (a beszédészlelési és a beszédmegértési teljesítmény vizsgálatára Gósy Mária által kifejlesztett és sztenderdizált eljárás) sikeres, német nyelvre történő adaptációját és első eredményeit mutatja be. Menyhárt Krisztina a gyermekek által közkedvelt rajzfilmek kommunikációját vizsgálja percepciós szempontból, illetve hasznos tanulságokat fogalmaz meg a kisgyermekek tévézési szokásaival kapcsolatosan. Horváth Viktória érdekes tanulmányt közöl a hezitálások percepciójáról: a beszélők (sok esetben tévesen) vélekednek úgy saját magukról, hogy keveset hezitálnak, továbbá a hezitálások megfigyelésében és felfedezésében nagy egyéni különbségeket mutattak. A. Jászó Anna a retorikai megközelítésről mint szövegelemző módszerről értekezik, Navracsics Judit és Sáry Gyula cikke pedig a kétnyelvűek írott alapú mondatfeldolgozását hasonlítja össze több szempont szerint.

A *Zönge, zöngésség, zöngételenedés* részben Damir Horga a zavarmentes artikuláció és az artikuláció mesterséges akadályozásának különböző eseteit vizsgálta meg, és leírja, hogy az adatközlők zöngeképzési mechanizmusuk új helyzethez történő igazításával törekszenek a helyes artikulációra. Gráczi Tekla Etelka és Kohári Anna a több felpattanással megvalósuló explozívák VOT-mérésének metodikai kérdéseit igyekeznek tisztázni tanulmányukban. Ruth Huntley Bahr és Stacy N. Kile mesterségesen zöngétlenné alakított végű szavak esetében vizsgálta meg különböző anyanyelvű gyermekek percepcióját.

A *Prozódiai sajátosságok*at tárgyaló fejezetben további érdekes tanulmányokkal találkozhatunk. Keszler Borbála régi szövegekben elemzi az intonációs jelöléseket, bizonyítva ezzel az élőbeszéd vizsgálatának létjogosultságát a szövegekben. Markó Alexandra vizsgálata szerint a felolvasás és a spontán beszéd hangsúlyai akusztikai sajátosságaikban különbözőek, és mindez a percepcióra is hatással van. Váradi Viola a spontán beszédben megnyilvánuló bekezdéshatárok akuszti-

kai-fonetikai korrelátumait kutatja, Gocsál Ákos pedig új irányt jelöl ki az alaphangmagasság mérésére, a mérés további pontosításának érdekében.

A hatodik fejezetben, amely a beszéd információtechnológiai vonatkozásáról szól, Beke András a gépi beszélőfelismerés új eljárásának segítségével (kevert Gauss-modellek) az eddiginél jobb eredményeket mutat be. Olaszy Gábor a gépi beszéd előállításának eddigi tapasztalatait összegzi, illetve írja le a jövő kutatásainak irányát, szükségességét.

A kötet kívülről színekben és formákban gazdag, küllemének változatossága a beszédtudomány sokféleségével harmonizál, noha a külső még nem sejteti a mű igazi értékeit. A borító színvilága, a feltüntetett kulcsszavak alapján sokkal inkább hihetjük azt, hogy hétköznapi felhasználásra készült ismeretterjesztő könyvvel, esetleg tankönyvvel van dolgunk. A könyvet átlapozva azonban gyorsan feltűnik az olvasó számára, hogy komoly tudományos munkáról van szó, kiváló kutatók keze nyomán, igényes szerkesztéssel.

A kötet bármely cikkét tanulmányozva elmondható, hogy a beszédtudományi kutatások mindennapi aktualitását és innovációját hordozza magában. A tanulmányok egytől egyik színvonalas, alapos munkák eredményei, amelyek nemcsak Gósy Mária professzor munkásságának méltó továbbvivői és üdvözlői, hanem a modern beszédtudomány egy pillanatképének is nívós alkotóelemei.

Németh Mátyás

14. Nemzetközi Klinikai Fonetikai és Nyelvészeti Konferencia – 14th Meeting of the International Clinical Phonetics and Linguistics Association

Cork, 2012. június 27–30.

Az 1991-ben alapított Nemzetközi Klinika Fonetikai és Nyelvészeti Társaság (International Clinical Phonetics and Linguistics Association – ICPLA) kétévente szervez szakmai találkozót, amelynek célja a beszéd és nyelv zavaraira fókuszáló fonetikai és nyelvészeti kutatások ösztönzése, a kutatás és gyakorlat közti átjárás elősegítése. A 2012-es évi konferencia helyszíne a több mint másfél évszázados múltra visszatekintő University College Cork, a tudományos programot szervező helyi bizottság tagjai az egyetem jelentős kutatási tapasztalattal rendelkező oktatói: Fiona Gibbon, Alice Lee, Paul Fletcher, Ciara O'Toole, Nicola Bessell, Clara Egan, Nicole Kennedy, Pauline Frizzelle és Marcin Szczerbinski. A szakmaközi

párbeszéd elősegítését a 4 plenáris előadásnak, 9 tematikus panelnek, 60 szekcióelőadásnak és 160 poszterbemutatónak teret engedő tudományos program szolgálta. A konferencia nagy érdeklődés mellett megrendezett eseménye David Crystal köszöntése volt, aki a klinikai nyelvészet területén végzett munkája elismeréseként kitüntetésben részesült. A díjátadást a klinikai nyelvészet történetét áttekintő szakmai beszélgetés követte.

A konferencia nyitóelőadását Alan Titley (University College Cork) tartotta az ír nyelv múltjáról, jelen helyzetéről és jövőjéről (*What is Irish?*). Faraneh Vargha-Khadem (University College London) központi témája a FOXP2 gén mutációjához kötődő strukturális és funkcionális változások, valamint az ezek nyomán felrajzolható, a beszéd, illetőleg a nyelv alapjául szolgáló neurális háló bemutatása volt (*Dissecting the neural network involved in speech and language through structural and functional brain imaging: FOXP2 and the KE family*). Richard G. Schwartz (University of New York) a nyelvi zavarokat a pszicholingvisztika oldaláról közelítve meg, cochleáris implantált gyermekek lexikális hozzáférési sajátosságairól számolt be (*Language processing in children with cochlear implants: Lexical access*). A konferencia negyedik, záró plenáris előadásában Roelien Bastiaanse (University of Groningen) nyelvközi adatok elemzésével afáziások igehasználati sajátosságait értelmezte (*Verbs in aphasia: What is the problem?*).

A speciális szekciók keretében elhangzott előadások a beszéd és nyelv zavarait főként a diagnosztika oldaláról közelítették meg, ezen kívül az atipikus mintázatban megfigyelhető variabilitás természete, valamint a terápia tudományos megalapozottságának szorgalmazása szerepelt még a programban.

A szekció-előadások és poszterbemutatók a közvetítő nyelvek számát, a vizsgálat/elemzés módszerét és tárgyát, valamint az atipikus mintázatot eredményező kóroktani hátteret, illetőleg speciális környezeti feltételeket tekintve egyaránt nagyon változatosak voltak. A korábbi találkozókon megfigyelhető törekvéseken kívül talán kettőt lehetne megemlíteni, mely a kutatások száma alapján kiemelt figyelmet kapott. Klinikai területen is hangsúlyosabbá vált az igény a „hallgatóközpontú" adatgyűjtéstől a „beszélőközpontú" irányába történő elmozdulásra, kiejtéshű adatrögzítés, illetőleg az észlelési adatokat kiegészítő eszközalapú mérések szorgalmazásával. A másik törekvés, a megsegítés hatékonyságát tartva szem előtt, az evidenciaalapú terápiás gyakorlat szorgalmazásában fogalmazódott meg.

Az idei évben 14 hazai kutató két előadással, illetőleg 4 poszterbemutatóval vett részt a konferencián. Gósy Mária, Bóna Judit, Beke András és Horváth Viktória (Budapest) a kitöltött szünet sajátosságait, illetőleg azok életkorral összefüggő változásait vizsgálta. Gyarmathy Dorottya és Gósy Mária (Budapest) olvasási nehézséget mutató gyermekek dichotikus hallásvizsgálatának eredményeiről számolt be. Markó Alexandra, Bajnócziné Szucsák Klára és Gráczi Tekla Etelka a hangte-

rápia hatékonyságát vizsgálta akusztikai paraméterek, illetőleg észlelési adatok elemzésével a tekintetben, vajon a kliensnek a beszélés mechanizmusára vonatkozó ismeretei befolyásolják-e a terápia kimenetét. Hoffmann Ildikó, Bánréti Zoltán, Szűcs Márta és Mészáros Éva (Szeged) kutatása különböző súlyossági fokú demenciát mutató Alzheimer-kóros betegek rekurzív képességére fókuszált. Zajdó Krisztina (Győr) az anyai beszéd hatását vizsgálta a magánhangzó fonológiai hosszúsága és a nyelvre jellemző zöngekezdési idő elsajátítása tekintetében. S. Tar Éva (Budapest) a zöngés/zöngétlen kontraszt elsajátítottságát elemezte atipikus nyelvfejlődést mutató gyermekek képmegnevezéses produkciójában.

A következő ICPLA konferenciára 2014-ben Svédországban kerül sor.

Tar Éva

Workshop Fluent Speech

Utrecht, Hollandia, 2012. november 12–13.

2012 novemberében első alkalommal rendeztek workshopot az utrechti egyetemen a beszéd folyamatosságát megakasztó jelenségekről. A kétnapos rendezvény célja az volt, hogy a nyelvészet különböző területein dolgozó, a spontán beszédben előforduló megakadásokat kutató szakemberek lehetőséget kapjanak egymás eredményeinek megismerésére, új szakmai kapcsolatok létesítésére. A workshop alatt ötven szerző tizenhat szóbeli és tíz poszteres prezentációját tekinthették meg a résztvevők. Az előadók Európából és az amerikai kontinensről érkeztek: Angliából, Belgiumból, Dániából, Finnországból, Hollandiából, Kanadából, Magyarországról, Németországból, Skóciából, Spanyolországból és az USA-ból. A poszter előadások azonos súllyal bírtak, mint a szóbeli prezentációk; a poszterek a teljes workshop ideje alatt láthatók voltak, így a mindkét napi poszter szekció mellett a kávészünetekben is megtekinthették őket az érdeklődők.

Az előadások a megakadásjelenségekkel kapcsolatos kutatások széles körét fedték le. Foglalkoztak a megakadásokkal mind a produkció, mind a percepció szempontjából, elemezték őket egynyelvű és kétnyelvű beszélők spontán beszédében is. A meghívott előadók között volt Rob Hartsuiker (Belgium), aki *Where do disfluencies come from?* című előadásában Levelt beszédprodukciós modellje alapján mutatta be a megakadások eredetét. Norman Segalowitz (Kanada) *Toward a framework for the systematic study of L1 and L2 fluency: What are the important questions?* címmel tartott előadásában az anyanyelvi és idegen nyelvi

produkció folyamatosságának vizsgálatát hasonlította össze. Az első nap tudományos programját John De Jong (Hollandia) szellemes módszertani előadása zárta *(Rasch measurement for testing subjects, data and hypotheses)*. Joan-Carles Mora és Roger Gilabert (Spanyolország) *Individual factors in utterance and perceived fluency: some empirical issues* előadásában a beszélők közötti különbségek befolyásoló hatását elemezte az idegen nyelvi beszédprodukcióra. Martin Corley (Skócia) azt mutatta be, hogy a bizonytalanságból fakadó megakadásjelenségek milyen hatással vannak a beszédfeldolgozásra *(Where do disfluencies go to?)*.

Magyarországot négy előadás képviselte. A rendezvény második szóbeli előadása volt Gósy Mária (MTA Nyelvtudományi Intézet és ELTE) és Váradi Viola (ELTE) prezentációja az önkorrekciók tipizálásáról és időzítési sajátosságairól. Deme Andrea és Markó Alexandra (ELTE) a nyújtások jellemzőit mutatták be gyermekek és felnőttek beszédében. Gyarmathy Dorottya és Neuberger Tilda (MTA Nyelvtudományi Intézet) a téves szótalálások és a téves kezdések javításairól tartottak előadást. Bóna Judit (ELTE) idősek és fiatalok beszédét vetette össze háromféle beszédtípusban (narratíva, tartalomösszegzés és társalgás) a beszédtempó, a szünettartás és a megakadások szempontjából.

A szervezők gondoskodtak arról, hogy a résztvevőknek kötetlen beszélgetésre is alkalmuk legyen. Az első nap estéjére szervezett vacsora és a konferencia zárásaként tartott koccintás is közelebb hozta egymáshoz a különböző országokban dolgozó kutatókat. A tervek szerint a workshopot kétévente rendezik meg az utrechti egyetemen.

Bóna Judit – Váradi Viola

A KÖTET SZERZŐI

Név	Intézmény	E-mail
Baditzné Pálvölgyi Kata	ELTE Spanyol Nyelv és Irodalom Tanszék	b.palvolgyikata@gmail.com
Bartha Krisztina	Partiumi Keresztény Egyetem	krisztinabartha@yahoo.co.uk
Beke András	MTA Nyelvtudományi Intézet	beke.andras@nytud.mta.hu
Bereznai, Anja	ELTE BTK Alkalmazott Nyelvészeti Doktori Program	anja@bereznai.de
Bóna Judit	ELTE Fonetikai Tanszék	bona.judit@btk.elte.hu
Font-Rotchés, Dolors	Universitat de Barcelona, Department of Catalan Philology, Applied Phonetics	dolorsfont@ub.edu
Fóris Ágota	KRE Magyar Nyelvtudományi Tanszék	foris.agota@kre.hu
Gósy Mária	MTA Nyelvtudományi Intézet	gosy.maria@nytud.mta.hu
Götz Gergely	Bakáts Dentál Bt.	gotzgergely@t-online.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
Kiss Gábor	BME Távközlési és Médiainformatikai Tanszék	kiss.gabor@tmit.bme.hu
Krepsz Valéria	ELTE Fonetikai Tanszék	krepszvaleria@gmail.com
Markó Alexandra	ELTE Fonetikai Tanszék	marko.alexandra@btk.elte.hu
Németh Mátyás	ELTE BTK Alkalmazott Nyelvészeti Doktori Program	nemeth.matyas@outlook.com
Neuberger Tilda	MTA Nyelvtudományi Intézet	neuberger.tilda@nytud.mta.hu
Olaszy Gábor	BME Távközlési és Médiainfomatikai Tanszék	olaszy@tmit.bme.hu

Név	Intézmény	E-mail
Siptár Péter	ELTE Elméleti Nyelvészeti Tanszék, MTA Nyelvtudományi Intézet	siptar.peter@nytud.mta.hu
S. Tar Éva	ELTE Bárczi Gusztáv Gyógypedagógiai Kar	evatar@barczi.elte.hu
Váradi Viola	ELTE Fonetikai Tanszék	varadiviola@gmail.com
Vicsi Klára	BME Távközlési és Médiainformatikai Tanszék	vicsi@tmit.bme.hu